Simon Eckermann

# Health Economics from Theory to Practice

## Optimally Informing Joint Decisions of Research, Reimbursement and Regulation with Health System Budget Constraints and Community Objectives

Adis

# Health Economics from Theory to Practice

Simon Eckermann

# Health Economics from Theory to Practice

## Optimally Informing Joint Decisions of Research, Reimbursement and Regulation with Health System Budget Constraints and Community Objectives

Adis

Simon Eckermann
Health Economics
University of Wollongong
Wollongong
New South Wales
Australia

ISBN 978-3-319-50611-1 ISBN 978-3-319-50613-5 (eBook)
DOI 10.1007/978-3-319-50613-5

Library of Congress Control Number: 2016963793



Printed on acid-free paper

This Adis imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Foreword

Economic evaluation is a lot more sophisticated now than it was 30 years ago. Then, it provided a powerful but very simple framework for systematically assimilating and comparing the costs and benefits of health-care interventions. The idea was revolutionary. It turned prevailing thought on its head. Health care is rationed whether the system is predominantly public or private. This idea was a surprise to many and resisted by most. That one should consider the cost-benefit of health care when deciding how to allocate resources was seen as anathema. Health economics was a contradiction in terms. We had no right to put a price on life. This was too big an assault on people's 'cherished ideals'. Clinicians typically asserted that the economist's role was to find them the resources they needed to get the job done and then stand out of the way. This view had popular support.

The simplicity of the economic framework played an essential role in overcoming this opposition. The logic – impressed upon me by my teachers and which I repeat endlessly in my own teaching that scarcity of resources = inevitability of choice = opportunity cost – was impeccable. The case for economic evaluation was crystal clear and hard to challenge though this did not stop people trying. The discipline's founding figures gleefully used the logic of economic evaluation and the notion of opportunity cost to reveal the flaws in the arguments employed by clinicians and bureaucrats alike. To name just two, Alan Williams articulated the shortcomings of the view that clinical freedom was sacrosanct. Gavin Mooney revealed how those who said you could not value life frequently did just that in their policy decisions. The introduction of rollover structures on farm vehicles and the decision not to introduce childproof lids on drug containers being two examples that implied a value to life in a way that was neither transparent nor consistent. Resistance to economic evaluation was not a question of value or ethics as it was an attempt to dodge accountability.

These early efforts broke the ground for those of us who followed. We had much easier paths to take. My peers and I, including the authors of this book, readily found work undertaking health economic evaluations. We gained extensive practical experience in multiple settings evaluating different sorts of health intervention addressing different sorts of health problems. The breadth of this experience soon

revealed the shortcomings of our methods: shortcomings that were readily apparent to those whose services we evaluated. The terrain soon changed. Where I could once hold forth on *why* one should do economic evaluation, I soon had to talk about *how* to do economic evaluation and then soon after that on *how to do it* better.

Fortunately, with such a large empirical programme available to us, we also had ample opportunity to explore the many theoretical and methodological issues that arose.

Finessing the methods for measuring and valuing the benefits of heath interventions was one important and popular area of work. We compared multi-attribute scales and tested their validity and reliability. We considered the dimensions of health included in different scales and critiqued the potential for bias when popular scales did not cover outcomes that were important for particular areas of health practice such as prevention and palliative care. We examined the values that people attached to different health states, tested framing effects and other forms of bias and we compared the values provided by people from different nations and social backgrounds. As a result of all this foundational work, concepts such as quality-adjusted life years once an abstract and experimental concept moved to become a common method of valuing health outcomes used in health systems around the world. Costs per QALY now appear not infrequently on the front pages of national newspapers in the United Kingdom in discussions of the recommendations of the National Institute of Health and Care Excellence: an astonishing indicator of the status of the methodological advances that have been made and of how mainstream economic evaluation has become.

The development of methods for addressing and reporting uncertainty was another area that attracted a lot of keen young research interest. This led to the development of analytic advances to describe uncertainty jointly in terms of both cost and effectiveness, and it included graphical advances in how to depict uncertainty to decision makers.

We have also seen advances in modelling techniques that improve the way we are able to extrapolate lifetime costs and health gains expressed as QALYS from trial-based data with short time frames and/or intermediate outcomes, which is necessary if we are to compare the results of evaluations for the full range of very different health technologies. And as another indicator of the maturity of this science, there are now guidelines for practice that ensure a degree of comparability among studies and standards for assessing the quality of the work.

For many of us, it was enough to pick off one of these methodological areas to explore alongside a busy agenda of practical applied evaluations. My early experience was in evaluating the cost-effectiveness and quality of services received by people with learning disabilities newly discharged from large institutions into the community. One of the motivations behind this effort was the drive to ‘normalise’ the living experiences of people with learning disabilities: to provide the sort of life that most of us take for granted – going shopping, cooking one’s own meal, eating with friends around a dining table rather than a refectory, choosing one’s own clothes and choosing the time when one got up and when one went to bed, perhaps even earning some money. While economics helped us conceptualise how to

evaluate this change, the benefits of this shift in the locus of care did not fit comfortably into the outcome frameworks being developed to measure health-related quality of life.

Later, after moving to Australia, I became more interested in public health rather than health care, and this opened up new challenges. The most interesting public health interventions were those that sought to improve population health and reduce health inequalities by changing the properties of whole systems, such as whole communities, schools or worksites, and not the properties of individuals, at least not directly. Systems change is non-linear. It is not necessarily dose-respondent. Its timing can be difficult to predict and so difficult to measure. If effective, the outcomes are both multiple and multiplied as reinforcing feedback amplifies the impact of the intervention. This affects how one should evaluate cost-effectiveness and when one should evaluate it. It touches on the need for new methods from macroeconomics, complexity science, developmental evaluation and network analysis. It is a fertile ground for economic evaluation.

Alone among his generation of health economists, Simon Eckerman was not satisfied with picking off one methodological challenge alongside his busy work programme in applied economic evaluation. Simon saw the pressing need to address all shortcomings simultaneously if one was to generate estimates of value that were meaningful and useful for policy. It is not enough to have highly sophisticated methods for describing the uncertainty that comes from measurement error in some of the parameters in an evaluation if the outcome measure one uses systematically excludes aspects of benefit that are relevant to the intervention being examined. Similarly, there is little point finessing methods for dealing with particular types of health-care intervention such as surgery and medicine, if those methods are biased against other forms of health care that compete for a share of the budget, perhaps geriatric care or palliative care. The results had to be consistent to guide resource allocation decisions across research, reimbursement and regulation.

Thus, Simon's professional career has combined practical economic evaluations with an extensive methodological research agenda to rethink how we evaluate and compare diverse efforts to improve health. The aim is always to develop a framework capable of generating robust estimates of cost-effectiveness: estimates that stand up to changes in context, measurement error, sampling bias and the like.

The results of that effort are distributed among research papers, across many years, in high-class journals such as *Social Science and Medicine*, *Health Economics*, *Pharmacoeconomics*, *Medical Decision Making*, and *International Journal for Technology Assessment in Health Care*. Now, the cumulative insights generated by this large body of work and that of co-author Nikki McCaffrey in relation to palliative care evaluation in two chapters have been collected into one volume and reorganised and re-worked to provide a complete narrative that yields deeper insight into the arguments contained in that research. Covered here are ideas that have been exposed in numerous workshops, tested in various policy forums, examined in conferences of health economists.

While I am not sure that I necessarily agree with all of the arguments made, I am highly sympathetic to the effort. There are some advantages of the partial approaches

that Simon and Nikki challenge and critique, where simplicity resonates with decision makers. The results might not be as robust as Simon and Nikki would like, but they may be robust enough where decision maker needs are partial. I wrote earlier that systems change is usually non-linear. Rather it is discontinuous, occurring in phase transitions. Effects flat line for long periods before jumping alarmingly when a tipping point or threshold is reached. In the evaluation of many health interventions, linear approximations may have served us well so far, and can be easier to apply and therefore more widely used. But to be able to assess this, we need the sorts of methodological critique that Simon and Nikki outline here to allow us to make the comparisons.

This book is a mark of the maturity of the field. However for some it will not be an easy book to read. It is an ambitious book. It covers a broad and diverse terrain in optimising across research, reimbursement and regulatory decisions. It is challenging and in certain parts (particularly value of information methods in Chap. 5) has sections which quickly become technically advanced. While methods are presented from first principles and contain helpful diagrams, full understanding of readers in such sections either require a high level of prior mathematical ability or a deep commitment to learn. It should repay the effort though.

Alan Shiell
Professor Public Health Economics
Department of Public Health
School of Psychology and Public Health
La Trobe University
Melbourne 3086, VIC, Australia
14 December 2016

# Text Background and Author Acknowledgement

How did this text develop? Who might benefit from reading it?

The 'Health Economics from Theory to Practice' course underlying this text has been developed from first principles over the last 11 years and caters to any level of background in health economics. Many variants-related courses from 1 to 5 days long have been run for a range of clinical, health services, policy, industry and HTA practitioners and students with no to advanced background since 2005.

However, the antecedents of materials for the first Health Economics from Theory to Practice course in 2005 really began in 1993 with a research question developed under the guidance of Gavin Mooney. The research question arrived at was in relation to whether hospital efficiency measures could avoid incentives for cost-minimising quality of care, cost shifting and cream skimming and create appropriate incentives for quality of care in practice. This was explored and developed further in undertaking Gavin's Tromso graduate diploma course in 1993–1994 which in turn lay the pathway to a PhD thesis (Eckermann 2004) where the net benefit correspondence theorem arrived at a unique solution. I thank Kevin Fox and Knox Lovell for their PhD supervision in developing that thesis and their mentorship more generally.

The Health Economics from Theory to Practice course itself developed from a chance meeting with Bernie O'Brien in 2001 and an invitation to present at a course he was running while on sabbatical in Australia. That course predominantly presented materials that Bernie had developed in his long and fruitful collaborations with Andy Willan. Talking to and working with Bernie in preparation for that and subsequent courses and listening to his presentations had a profound effect on me, when I was lucky enough to spend time with him in Australia and Canada from 2001 to 2003. I also owe a great debt to Bernie for introducing me to Andy Willan in 2002 as well as other colleagues.

Since 2000, variants of these course materials have been developed with colleagues and students I have been fortunate enough to have been associated with including:

John Simes and students of the Decision Analysis course in the Masters of Public Health at Sydney University taught with each year from 2000-2005 - John I thank for his mentorship and collaborations in research and teaching during that period and for encouraging me to develop the original Health Economics from Throery to Practice Course in 2005;

Professor Willan with Health Economics from Theory to Practice courses run 14 times from 2005-2014 in Australia, the UK and Canada;

A one day course 'tasting of health economics from theory to Practice' run for clinicians and researchers in SA in 2008;

Health economic methods for health technology assessment, a two day course run for ARCS in 2009 and 2010 with Michael Coory;

Health economic principles and research methods, a 5 day course run for University of Wollongong each year from 2010-2014, and;

The most recent Health Economic from Theory to Practice courses run with Dr McCaffrey in Tasmania in 2014 and Sydney in 2015 and 2016.

Hence the current course and this text have benefited greatly from interaction with various mentors, colleagues and participants over the past two decades or more. The course for me has acted as a regular sounding board for methods developed as well as constructive feedback in helping improve their presentation and applied use but also many collaborations and further research with those teaching and attending the courses. Similarly, I thank Andy Briggs and many colleagues at the Oxford Health Economics Research Centre (Alistair Gray, Oliver, Jose, Boby particularly) for a delightful sabbatical (with record length seminar of almost 3 h!) in 2004 as well as many subsequent memorable visiting seminars at Oxford and Glasgow and running of the Health Economics from Theory to Practice course in Oxford in 2009.

I especially thank those colleagues who have taught with me as part of the course faculty since 2005, where, along with Nikki McCaffrey, guest lecturers have included Tim Coelli, Brita Pekarsky, Jon Karnon and Andy Briggs. However, my most significant debt without a doubt is to Andy Willan who has been there since the beginning of the course and aided at many levels in encouraging and both contributing to and leading the publication of many of the central methods papers.

Naturally I also owe a general debt of gratitude to students who have previously undertaken the Health Economics from Theory to Practice course (some up to 5 times) in aiding shape a course where methods are developed from first principles and for which no prior knowledge is required.

Indeed, what has become clear in running the Health Economics from Theory to Practice course over the years is that those with no background at all in health economics are often the most comfortable. Those with a prior background can often be somewhat coy, particularly on the first day, and tend to admit they have had to question what they thought was solid ground in light of what has been exposed as biased methods by the course material.

This is particularly in relation to:

(i) Use of relative risk in indirect comparison and translation of evidence (odds ratios are required in these cases; see Eckermann et al. 2009, 2011);
(ii) Value of information research design locally, but particularly globally, in providing a first best solution maximising the globally expected value relative to cost of research designs and evidence translation while enabling feasible early

adoption and overcoming strategic and technical difficulties to best support joint research, reimbursement and implementation processes (Eckermann and Willan 2007, 2008a, b, 2009, 2013; Eckermann et al. 2010; Willan and Eckermann 2010, 2012);

(iii) Robust presentation and summary measures with multiple strategy comparisons, where net benefit and cost-effectiveness curves become unreliable and other presentation (the cost disutility plane) and summary measures (expected net loss curves and frontiers) have distinct benefits (Eckermann et al. 2008, Eckermann and Willan 2011);

(iv) Natural extension of multiple strategy comparison to comparison of health system or efficiency of providers in practice consistent with maximising net benefit; and

(v) Threshold values for effects reflecting opportunity costs of reimbursement and displacement actions where the health shadow price (Pekarsky 2012, 2015) provides support for optimal research, reimbursement and implementation decisions for new technologies (Eckermann and Pekarsky 2014)

I trust that previous students who have undertaken the course have benefited from, while no doubt in many cases at least initially being challenged by, the broad overarching scope of the principles and methods taught. I thank all those who have participated and those who have given feedback and collaborated. I hope that this text will aid those who have previously undertaken the course, those who might do so in the future, those who might like to teach such materials and those who simply would like a text to guide them towards unbiased methods for optimising related health economic decisions.

In preparation of this book over the past 3 years, I would like to thank those who have contributed both directly and indirectly to its production.

Dr. McCaffrey most directly as a collaborator on Chaps. 4 and 10 in relation to palliative care evaluation and multiple domain methods, respectively, and indeed is leading author of Chap. 10.

I acknowledge and thank Professor Willan for checking Chaps. 5, 6 and 7 and offering suggestions for Chap. 5 on value of information methods. As above I also acknowledge and thank him for the many research collaborations that lead to the methods extensively referred to in Part II (Chaps. 5, 6 and 7) on value of information methods for optimising joint research and reimbursement decisions. More generally still his collaboration on methods optimising across research, reimbursement and regulation decisions represented and referred to in each section and providing the overarching backbone on the book.

I thank and acknowledge Brita Pekarsky for checking Chap. 11 on economically meaningful threshold values and the health shadow price, her suggestions for improving that chapter and her seminal work in that research area, collaborations, conversations and deep understanding more generally of issues addressed in the text and friendship over the past 25 years.

I also thank Brita for checking equations and formatting the submitted text.

I gratefully thank and acknowledge the guidance, expertise and aid in helping construct policy examples in Chap. 12 of:

Lyn Phillipson on age and dementia friendly community programs in section 12.2;
Richard Fleming on design of dementia friendly aged care facilities in section 12.3; and in section 12.4 Justin Sinclair, David Caldicott and Lynnaire Sheridan on optimal policies for medicinal cannabis programs and Marie Ranson in consideration of a policy relevant promising reformulation of 5-FU.

I acknowledge and thank the University of Wollongong and particularly John Glynn and Charles Areni in supporting my undertaking of the text on top of my usual policy, methodological and grant-related research and associated administrative and teaching activities.

Finally I thank Nicola and our two beautiful children Pascal and Ruby for their support over the 3 years it took to write the text.

The book received no external funding and the authors have no pecuniary interests associated with the methods and research presented. Naturally any errors or omission are the responsibility of the authors, myself in all chapters and Nikki McCaffrey in collaboration in Chap. 4 and in leading Chap. 10.

Simon Eckermann

For Bernie and Gavin – the best mentors any health economist could ask for and the most delightful of human beings.

Acknowledgement from Dr. McCaffrey

I'd like to acknowledge and thank Simon for providing me with the opportunity to contribute to the Health Economics from Theory to Practice course and this text and for his inspiration, encouragement and support in developing the methods for conducting multiple outcomes cost-effectiveness analysis. Many thanks go to Professor David Currow for his guidance and mentorship and for providing clinical advice on Chaps. 4 and 10. Further, I acknowledge and thank Professor Jon Karnon for his technical guidance on Chap. 10.

I'd like to thank the Australian government Department of Health and Ageing National Palliative Care Program, the Palliative Care Clinical Studies Collaborative and Flinders University for supporting my contribution to the text and courses. Finally, I'd like to thank my husband, Kevin McCaffrey, for his unwavering support with my demanding academic career.

Nikki McCaffrey

## References

Eckermann S. Hospital performance including quality: creating incentives consistent with evidence-based medicine. PhD Dissertation, UNSW, Sydney. http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/.

Eckermann S, Pekarsky B. Can the real opportunity cost stand up: displaced services, the straw man outside the room. PharmacoEconomics. 2014;32(4):319–25.

Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16:195–209.

Eckermann S, Willan AR. The option value of delay in health technology assessment. Med Decis Making. 2008a;28:300–5.
Eckermann S, Willan AR. Time and EVSI wait for no patient. Value Health. 2008b; 11:522–6.
Eckermann S, Willan AR. Globally optimal trial design for local decision making. Health Econ. 2009;18:203–16.
Eckermann S, Willan A. Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies. PharmacoEconomics. 2011;29(7):563–77.
Eckermann S, Willan A. Optimal global VOI trials: better aligning manufacturer and decision maker interest and enabling feasible risk sharing. PharmacoEconomics. 2013;31:393–401.
Eckermann S, Briggs A, Willan A. Health technology assessment in the cost-disutility plane. Med Decis Making. 2008;28:172–81.
Eckermann S, Coory, M, Willan AR. Indirect comparison: relative risk fallacies and odds solution. J Clin Epidemiol. 2009. 62:1031–6. doi: 10.1002/hec.1353.
Eckermann S, Karnon J, Willan A. The value of value of information: best informing research design and prioritization using current methods. PharmacoEconomics. 2010;28(9):699–709.
Eckermann S, Coory M, Willan A. Consistently estimating absolute risk difference when translating evidence to jurisdictions of interest. PharmacoEconomics. 2011;29(2):87-96.
Pekarsky BAK. Trusts, constraints and the counterfactual: reframing the political economy of new drugs. PhD Dissertation. 2012. University of Adelaide. http://digital.library.adelaide.edu.au/dspace/handle/2440/79171.
Pekarsky BAK. The new drug reimbursement game: a regulator's guide to playing and winning. New York: Springer; 2015.
Willan AR, Eckermann S. Optimal clinical trial design using value of information with imperfect implementation. Health Econ. 2010;19:549–61.
Willan AR, Eckermann S. Value of information and pricing new health care interventions. PharmacoEconomics. 2012;30(6):447–59.

# Abbreviations

| | |
|---|---|
| ABS | Australian Bureau of Statistics |
| AE | Allocative efficiency (AE = EE/TE) |
| AIHW | Australian Institute of Health and Welfare |
| AN | Adopt with no trial |
| AT | Adopt and trial |
| CBD | Cannabidiol |
| C-DU | Cost disutility (plane, space with multiple effects) |
| C-E | Cost-effectiveness (plane) |
| CEA | Cost-effectiveness acceptability (curve, frontier; plane and surface with multiple effects) |
| CEAC | Cost-effectiveness acceptability curve |
| CEAF | Cost-effectiveness acceptability frontier |
| CEAP | Cost-effectiveness acceptability plane |
| CEAS | Cost-effectiveness acceptability surface |
| CHD | Coronary heart disease |
| CIHR | Canadian Institutes of Health Research |
| COPD | Chronic obstructive pulmonary disease |
| CRS | Constant returns to scale |
| DEA | Data envelopment analysis |
| DRG | Diagnosis-related group |
| DT | Delay and trial |
| DU | Disutility ($E^{DU}$, effect framed from a DU perspective, e.g. mortality) |
| DM | Decision making |
| ECV | External cephalic version |
| EE | Economic efficiency (EE = TE × AE) |
| ENG | Expected net gain (EVSI less expected costs) |
| ENB | Expected net benefit |
| ENL | Expected net loss (curves and frontier with one effect; planes and contour with multiple effects) |
| EOL | Expected opportunity loss |
| EVPI | Expected value of perfect information |

| | |
|---|---|
| EVSI | Expected value of sample information |
| GAP | Good agricultural practice |
| GERD | Gastro-oesophageal reflux disease |
| GMP | Good manufacturing practice |
| H2RA | Histamine2-receptor antagonists |
| HMO | Health maintenance organisation |
| HRQOL | Health-related quality of life |
| HTA | Health technology assessment |
| ICER | Incremental cost-effectiveness ratio |
| IHPA | Independent Hospital Pricing Authority (Australia) |
| INB | Incremental net benefit |
| INMB | Incremental net monetary benefit |
| LIPID study | Long-Term Intervention with Pravastatin in Ischaemic Disease study |
| MAUI | Multi-attribute utility instrument |
| NB | Net benefit |
| NBCT | Net benefit correspondence theorem |
| NCB | Net clinical benefit (absolute incremental effect) |
| NHMRC | National Health and Medical Research Council (Australia) |
| HHPA | National Health Performance Authority (Australia) |
| NICE | National Institute for Health and Care Excellence (UK) |
| NL | Net loss |
| NNT | Number needed to treat |
| NW | North-west (quadrant on the C-E plane) |
| OECD | Organisation for Economic Co-operation and Development |
| OR | Odds ratio |
| PBAC | Pharmaceutical Benefits Advisory Committee |
| PBMA | Programme budgeting and marginal analysis |
| PPI | Proton pump inhibitor |
| QALY | Quality-adjusted life year |
| QIC | Quality inclusive cost (QIC = $C + \lambda \times \text{DU}$) |
| QOC | Quality of care |
| RCT | Randomised controlled trial |
| RD | Risk difference (difference in absolute risk) |
| RR | Relative risk |
| SAKGNP | Stephanie Alexander Kitchen Garden National Program |
| SDM | Societal decision making |
| SE | Scale efficiency |
| SW | South-west (quadrant on the C-E plane) |
| SWTP | Societal willingness to pay |
| TE | Technical efficiency |
| THC | Tetrahydrocannabinol |
| UK | United Kingdom |
| US | United States |
| VOI | Value of information |
| VRS | Variable returns to scale |
| WHO | World Health Organization |

# Contents

# Chapter 1
# Introduction

## 1.1 Overview

This text aims to provide a robust set of health economic principles and methods for informing societal decisions in relation to research, reimbursement and regulation. We do not aim or claim to be comprehensive in covering all methods. We do aim to provide a theoretical and practical framework that navigates to avoid common biases and suboptimal outcomes observed in practice of health economic analysis and highlight methods that address these problems.

Our goal is to facilitate constrained optimisation of health system related community outcomes or net benefit from a societal perspective given budget constraints, existing technology and processes available for technology and program evaluation and assessment. This is shown to require methods which efficiently inform health system decision making across research, reimbursement and regulation decisions. Importantly, such joint consideration includes identifying an efficient process to maximise the potential that arises from research in relation to existing and new technology under uncertainty (Eckermann and Willan 2007, 2009, 2013; Eckermann 2010) and associated opportunity costs of adoption and financing actions undertaken with reimbursement and pricing decisions (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014; Eckermann 2015).

Joint coverage and comparability principles (introduced in greater detail in Chap. 2) for assessing incremental costs, effects and cost effectiveness or net benefit with appropriate threshold values for effects (see Sect. 2.7 and Chap. 11) are central to robust and unbiased health economic decision making, analysis and methods considered throughout this book. This is the case whether considering:

(i) Cost effectiveness analysis and adoption decisions in Chaps. 2, 3 and 4, Part 1;
(ii) Joint research and adoption decisions in Chaps. 5, 6 and 7, Part 2;
(iii) Joint research, adoption and regulation of providers and health systems in practice with robust multiple strategy, provider and domain methods consistent with maximising net benefit in Chaps. 8, 9 and 10, Part 3; and



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_1

(iv) Optimising joint research, reimbursement (adoption and financing) and regulation with appropriate health shadow pricing and threshold effect valuation for net benefit in integrating new with existing technology options for budget-constrained optimisation of decision making and policy analysis in Chaps. 11 and 12, Part 4.

The methods developed and presented to optimise societal decision making recognise that health economics research, analysis and societal decision making does not occur in a vacuum. Hence, optimising across joint societal research, reimbursement and regulation decisions explicitly allows for funder, provider and manufacturer interaction including degree of provider implementation conditional on strength of evidence (Willan and Eckermann 2010, 2012; Eckermann and Willan 2013) and consideration of incentives created by institutional arrangements and policies (Eckermann 2004; Pekarsky 2012, 2015; Eckermann and Coelli 2013; Eckermann and Pekarsky 2014; Eckermann 2014b; Eckermann and Sheridan 2016).

The importance of political economy considerations underlying the health system is also centrally recognised in framing societal decision making as budget-constrained optimisation of net benefit, from the community perspective that health systems should serve. This aims to avoid fundamental biases that arise where the political economy underlying societal decision making for the health system can stray from an underlying objective of community values and rather reflect partial perspectives such as manufacturer-vested interests and/or clinical values (Mooney 2012). As Gavin Mooney often stated, health systems should serve the communities they care for, and central to this is having an underlying societal objective function that reflects community preferences.

## 1.2 An Appropriate Underlying Objective Function

An appropriate objective for societal decision making represents the fundamental structural issue for health economics in establishing what is valued and should be included in determining appropriate principles, evaluation approaches, methods and metrics to inform health system decision making. The underlying objective should be central to decision making whether in relation to system architecture and institutions, strategies, technology, research or implementation options. Health economic evaluation and analysis should always consider what is the objective function and from whose perspective. Having an inappropriate objective or not considering or losing sight of community objectives will inevitably result in perverse incentives, while also facing the very real dangers of partialisation where objectives become too reductionist.

Maximising incremental net benefit from a societal decision making perspective, as the value of incremental effects of strategies (whether alternative health promotion programs, screening or diagnostic pathways or interventions) less their incremental cost across the health system has been suggested in health technology

assessment as a robust and appropriate metric to inform societal decisions in health care (Claxton and Posnett 1996; Stinnet and Mullahy 1998; Willan and Lin 2001; Briggs et al. 2002; Eckermann 2004; Willan and Briggs 2006; Drummond et al. 2005). More generally, Graham (1981, 1992) has previously shown that maximising net benefit allows constrained optimisation in investment choices across public and private sector investment decisions, provided threshold values for effects reflect opportunity costs associated with best alternative actions. However, critical sets of questions remain in health care for use of the net benefit metric as a robust objective function for optimisation, with those posed and addressed in this text including:

(i) Is there adequate coverage (scope[1] and duration) as well as comparability to obtain unbiased estimates of incremental costs and effects for robust net benefit assessment?
(ii) What should societal threshold value of effects for net benefit assessment for new and existing be, given opportunity costs of adopting and financing across potential investment options with characteristic health system inefficiency and constrained budgets?
(iii) Do efficiency comparisons and funding of health-care providers in practice create incentives consistent with budget-constrained maximising of health system net benefit of the community?

In relation to (i), coverage issues and biases were in large part ignored with the clinical origins of cost effectiveness analysis (Weinstein and Stason 1977) and health technology assessment (HTA) processes having an almost sole emphasis on randomised control trial evidence in avoiding selection bias. Such emphasis on evidence comparability without considering evidence coverage has inadvertently led to many structural, coverage and method biases being inappropriately considered as uncertainty (Briggs et al. 2012). That is, in many cases, failure to consider biases aside from selection bias associated with non-randomised evidence has resulted in a failure to appropriately recognise or control for other biases before considering uncertainty. While lack of randomisation is emphasised and recognised as a form of selection bias, biases should also be recognised as highlighted in Chaps. 2 and 3 when associated with:

[1] Adequate coverage of the scope of outcomes in cases such as evaluating palliative care (policies, strategies, interventions or providers) points to the need for methods for comparing multiple domains and outcomes of interest such as carer, family and patient impacts on process of death (finalising personal and financial affairs, dying in a setting of choice, etc.). These domains are not able to be summarised by a common metric such as quality-adjusted life years (QALYs), which in integrating patient survival with patient quality of life prevent integration of processes related to death. Further, even for cases where outcomes could be summarised with QALYs, issues arise in relation to the subjectivity of health-related utility weights and local nature of any preference weights used with QALYs and hence lack of universality of QALYs. Consequently, it should be clear that approaches that enable multiple outcome domains to be compared will often be necessary or valuable. Robust methods which appropriately and flexibly allow for multiple domain consideration with comparison on multiple strategies and domains on the cost-disutility plane are highlighted in Chaps. 4 and 10.

(i) Inadequate or inconsistent coverage of the scope and/or duration of incremental effects and costs (O'Brien 1996);
(ii) Partial analysis of cost and effects such as the box method (Briggs et al. 2002) and cost minimisation analysis (Briggs and O'Brien 2001); and
(iii) Selection biases in indirect comparison and evidence translation with use of relative risk, as a nonsymmetric metric, where inconsistent results arise with alterative framing of binary outcomes (survival, no survival; progression, no progression, etc.) whenever evidence synthesis or translation is needed (Eckermann et al. 2009, 2011).

Addressing such biases before considering uncertainty is critical to better informing societal decision making, given modelling uncertainty around biased estimates systematically misleads decision making and is open to abuse by vested interests. Indeed, modelling uncertainty of cost, effectiveness or cost effectiveness before establishing unbiased estimates is akin to wearing rose-coloured glasses to avoid squinting when one wants to see the true colour of something.

The importance of avoiding biases before considering uncertainty is hammered home when considering the Arrow-Lind theorem (Arrow and Lind 1970) underlying public investment decisions. The theorem clarifies that societal decision making in large government with risk spreading and diversification characteristic of societal investment across many decisions and populations should be foremost interested in expected cost effectiveness, given preferences asymptote towards risk neutrality in such cases. Chapter 8 highlights this as particularly key for multiple strategy comparisons and the many decisions over time made by jurisdictional governments and their institutions and regulatory bodies such as the PBAC, NICE, etc. across large populations.

Methods enabling joint satisfaction of coverage and comparability principles, and avoid biases, are emphasised as the foundation blocks for improving societal decision making in Part 1 (Chaps. 2, 3 and 4). Coverage and comparability principles are emphasised in Chap. 2 where the need to jointly consider costs and effects for within trial cost effectiveness analysis is highlighted following the death of cost minimisation and thinking outside the box papers (Briggs and O'Brien 2001; Briggs et al. 2002). Chapter 3 highlights common problems where selection biases are inadvertently introduced in processes of evidence synthesis, extrapolation and translation to jurisdictions of interest. More importantly, solutions are identified in each case enabling unbiased consistent estimates of cost effectiveness for decision making (Briggs et al. 2002; Eckermann et al. 2008, 2009, 2011).

Chapter 4 extends these considerations to complex community-based settings such as those of health promotion and palliative care with network multiplier impacts and multiple domain comparisons, respectively. Coverage and comparability principles in evaluating across population network impacts over time and multiple domains in these settings are further highlighted and illustrated for community-ageing policies in Chap. 12 and multiple domain comparisons in Chap. 10.

Hence, coverage and comparability principles and methods to avoid bias are established up front as the backbone for robust cost effective analysis in informing reimbursement decisions with processes of health technology or program assessment in any jurisdiction of interest. They are also established as central principles throughout the text to robustly inform related decisions including:

(i) Joint research and reimbursement locally and globally with value of information methods illustrated in Part II (Chaps. 5, 6 and 7) which avoid partial hypothesis test problems of conventional methods.
(ii) Comparison of more than two strategies, multiple providers and/or outcomes with analysis methods and summary measures in Part III (Chaps. 8, 9 and 10) which simply and effectively avoid inferential and conflation issues otherwise arising.
(iii) Regulation of budget-constrained threshold values, pricing and system efficiency in practice and associated research and reimbursement decisions and policy and political economy challenges in providing a pathway to optimisation while avoiding silo mentalities, in policy and practice in Part IV (Chaps. 11 and 12).

Coverage and comparability principles become harder to satisfy with health promotion and prevention programs in complex community settings, such as schools as highlighted in Chap. 4. However, Chap. 4 also points to the distinct potential for costless expansion of effects from health promotion and prevention strategies in such community settings where there are network impacts and community ownership of strategies. Success of strategies in complex community settings requires assessing the acceptance and longer-term embedding and ownership of promotion and prevention approaches in targeted communities, as well as impacts on individual behaviour (Hawe and Shiell, 2000; Moore et al. 2006). Impacts arising from community behavioural change with health promotion and prevention are typically well beyond short-term evaluation timeframes, broader than health alone, and have diffuse network-related impacts across populations at a community level. These community-level network impacts across diffuse populations over time are critical to assessing the long-term ownership, effectiveness and cost effectiveness of health promotion strategies in practice. However, they are not captured by individual-focussed evaluation methods, with atomistic adding of impacts on individuals in highly targeted populations and often very short-term impacts. Hence, conventional cost effectiveness analysis models based on patient level evidence alone struggle to robustly estimate the direction let alone extent of long-term societal incremental costs, effects or cost effectiveness expected with health promotion and prevention programs.

Measuring network and multiplier effects over time from initial investment flowing into community activities are shown to provide adequate coverage and quantitative indicators of community ownership, engagement with and building of social networks and capital and sustainability of community health programs over time (Hawe et al. 2009; Shiell et al. 2008). Further, such multiplier impacts lend themselves to triangulation with qualitative assessment. It should therefore not be surprising that coverage and comparability principles point to different approaches from those to evaluate individual level therapies in evaluating community health promotion and prevention programs in complex community settings. This is illustrated

evaluating a kitchen-garden health promotion program in primary schools in Chap. 4 (Eckermann et al. 2014).

Across health system settings (whether health promotion and prevention, diagnostic, curative, rehabilitative or palliative treatment and care), the aim is to enable robust and efficient decision making with a principled approach to identifying appropriate methods for evaluation of effectiveness and cost effectiveness. The same underlying coverage and comparability principles and community objectives actively inform fit for purpose methods and approaches for undertaking health economic analysis, such as multiplier and network methods in evaluating health promotion in complex community settings.

## 1.3 Principles for Constrained Optimisation Across Health Promotion, Prevention and Care Settings

To enable budget constrained optimisation across health prevention and promotion, diagnostic, curative, rehabilitative and palliative settings health economics needs robust principles and unbiased while flexible methods to inform societal decision making across joint research, reimbursement and regulatory decisions.

In identifying robust and principled health economic methods for constrained maximisation across these health care setting and joint decisions, we bring together:

(i) The seminal research of Bernie O'Brien and colleagues (Andy Willan, Andy Briggs and others) in highlighting the need to move beyond partial clinical and economic consideration to jointly consider costs and effects (O'Brien 1996; Briggs and O'Brien 2001; Briggs et al. 2002; Willan and Briggs 2006).
(ii) Decision analytic principles of coverage and comparability shown as more generally required to avoid biases and inferential fallacies in evidence synthesis, translation and extrapolation to inform societal decision making in any given jurisdiction(s) of interest (Eckermann et al. 2009, 2011).
(iii) Methods for robustly evaluating health promotion strategies in community settings, where following Shiell and Hawe's research, community-level social capital and network multiplier impacts of strategies in practice (Hawe and Shiell 2000; Shiell et al. 2008; Hawe et al. 2009) are key.
(iv) Value of information methods enabling optimisation of joint research and reimbursement decisions allowing for key decision contexts (Eckermann and Willan 2007, 2008a, b, 2009, 2011, 2013; Eckermann et al. 2010; Willan and Eckermann 2010, 2012).
(v) Robust methods for regulating to create appropriate economic incentives for net benefit maximisation with multiple provider efficiency (Eckermann and Coelli 2013) as well as multiple strategy (Eckermann et al. 2008; Eckermann and Willan 2011) and multiple outcome (McCaffrey et al. 2015) comparisons.
(vi) Budget-constrained threshold values for effects reflecting opportunity cost of adopting and financing new technologies given alternative research and reim-

bursement options and decision contexts faced in any jurisdiction of interest, highlighting the health shadow price research of Pekarsky (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014).

The principles and research methods developed for the Health Economics from Theory to Practice course underlying this text have been aimed at providing a robust framework to jointly address these six areas. The process of bringing these areas together explicitly addresses structural issues of who and what health systems should serve and reflect – community objectives and values – to underlie processes of evidence-based societal decision making. The political economy of health systems, their institutions and decision making and actions are hence central to health economics principles and practice in optimising societal decision making. This is particularly important given the characteristics of health-care transactions and health systems and the importance of appropriate community values and objectives in achieving both equity and efficiency across health systems. Characteristic asymmetry of information (Akerlof 1970; McGuire et al. 1988) between providers and patient populations with associated bounded rationality (Simon 1957) of patients in complex decision making under uncertainty (Arrow 1963) arises in agency relationships with health care. These characteristics are prevalent across health systems and settings.

In many health-care settings, information asymmetries are often extreme between provider and patient agents in health-care interactions. They are also likely to be present ex post (after treatment) as well as ex ante (before treatment) given the incremental impact of any actual strategy or pathway is relative to counterfactual alternative strategies or pathways (McGuire et al. 1988: 43–44). These information conditions in turn lead to the need for health-care providers to act as agents for patients in enabling efficient decision making in these settings. However, they also create the conditions for supplier-induced demand to arise (and associated detrimental cost and health outcomes from overtreatment) where providers have incentives (financial and/or professional) to induce such demand. Understanding these characteristics and the need for appropriate incentives for health-care providers in practice is key to establishing policy and regulatory frameworks for efficiency as well as equity in health-care institutions and their performance monitoring and funding arrangements. The theoretical underpinnings of these factors for the health economics discipline point to the need for universal public health-care provision on equity but also population health, health system cost and efficiency grounds in providing appropriate incentives for providers, as Chap. 12 highlights.

Empirically, there is ample evidence for the importance of universal health care and payment arrangements and provider incentives for appropriate care, rather than perverse incentives for supplier-induced demand in populations able to access care, in order to enable health system efficiency as well as equity of access. This is particularly borne out in contrasting the joint cost and outcomes of the US health system relative to universal access systems in places such as Canada, the UK, France, Australia and indeed the vast majority of Organisation for Economic Co-operation and Development (OECD) countries. In 2013 the US health system cost on average double the proportion of GDP of universal publicly provided systems in the OECD

(18 vs. 9%) at $8505 per capita, in spite of the US systems' lack of universal access to health care. Further, despite this much higher health expenditure, the USA had some of the worst population-level health outcomes in the OECD, with life expectancy in 2013 lower than any country with annual health expenditure of more than US$2000 per capita, as well as the worst performing country in terms of health improvements over the past 50 years (OECD 2013). Such clear health system inefficiency with higher costs and worse health outcomes as well as inequity in the absence of universal health-care access (Davis et al. 2014) in large part arises as those who have access; the highest income quintile(s) are over-serviced, while those who don't have access or have very limited access, are underserviced. Over-servicing of those with access in the USA is reinforced by defensive medicine under threat of litigation and can be extreme, manifesting in unnecessary tests and subsequent unnecessary treatment of false positives (particularly for rare conditions) as well as cascading use of polypharmacy with symptom chasing of side effects.

Another related reason costs are significantly higher in the USA is due to increased complexity in administering a health system without universal access. This leads to higher costs in monitoring and assessing access and exclusion criteria and maintaining property rights across multiple systems and care provision (privately insured, Medicare for the elderly and Medicaid for some disadvantaged populations). There are also costs of the 'paper trails' between health-care providers, insurers and other funders. Historically, about one quarter of current US health-care system costs are associated with administration of their systems compared with approximately one tenth of health-care costs in universal access systems (Woolhandler et al. 2003). Further, for those with private insurance in the USA, alongside being over-serviced if they gain access to treatment, denial of treatment by health maintenance organisations (HMOs) simultaneously arose prior to Obamacare for necessary treatment of conditions which existed prior to insurance (pre-existing conditions required to be declared). Where pre-existing conditions are excluded, large administrative costs also arise in attempting to identify pre-existing conditions and associated litigation costs to patients and insurers, alongside leading to worse health outcomes from needed treatment not being provided for pre-existing conditions.

The bottom line is that publicly funded universal access health systems are both theoretically expected to be, and with empirical evidence observed to be, less costly while having better access and population-level health outcomes, and hence more efficient, than privately funded systems (OECD 2013; Davis et al. 2014). However, the equity and efficiency advantages of publicly provided universal health-care systems still depend on such systems providing appropriate incentives for providers and reflecting the objectives of the community they serve. If public health systems are to optimise outcomes for community benefit with constrained budgets and resources, then community objectives need to be reflected in decision making and robustly regulated to reflect health system level opportunity costs for those objectives. This is particularly the case in research and assessment of the types of strategies available with existing and new technology that are appropriate to use in different parts of the health system. Naturally this is also the case in the

way therapies and strategies are used in practice, in coordinating between parts of the system and across populations over time. These are important health economic questions this text aims to shed some light across, highlighting common biases and problems of often partial and silo-based approaches while identifying simple methods to jointly, robustly and efficiently address research, reimbursement and regulation decisions.

In combining these areas of health economic research, we will be drawing on many sources, attempting to bring together best approaches from the west (evidence-based medicine in diagnosis and treatment) with the east (preventative medicine and health promotion allowing for sociological and societal determinants) in addressing the full spectrum of options across health settings. HTA infrastructure focussed on assessing patentable medications, devices and diagnosis and testing strategies to the exclusion of unpatentable options acts to create distinct barriers to appropriate research into and adoption of unpatentable options and best expansion and contraction of existing programs. Such a system denies appropriate consideration of better use of existing programs and non-patentable alternative strategies (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014), for example, community health promotion, rehabilitation or function decline prevention programs, investing in better infrastructure for information flows and care coordination, overcoming barriers and supporting enablers for better implementation of strategies and modifying methods of care. Systems focused on new patentable interventions delay or completely stall the evidence for, and ability to appropriately compare and defend, current programs and non-patentable options against their displacement in areas including:

(i) Expanding use of 'off-patent' medication and its better use in indicated populations, e.g. use of existing statins.
(ii) Non-patentable alternative modality areas such as rehabilitative care in coronary heart disease (CHD) and chronic obstructive pulmonary disease (COPD) population or palliative care support at home or in institutional settings as alternatives to therapies such as radiotherapy and chemotherapy and associated medications in cancer populations.
(iii) Health promotion and primary prevention measure community-based approaches in community settings such as community gardens and kitchens in schools (Eckermann et al. 2014) or other community settings, walking paths and more generally age- and dementia-friendly facilities, programs and policies (Kalache 2013).
(iv) The use of natural plant varieties and extracts at factor costs in treatment of common conditions. For example, medicinal cannabis exploiting entourage benefits of CHD-, terpene- and THC-rich varieties (Wagner and Ulrich-Merzenich 2009; Russo 2011; Gallily et al. 2015) titrated up to individual patient needs and tolerance in palliative pain management populations (Johnson et al. 2010; Carter 2011).

Such options are explored at length in Chap. 12 considering promising policy, research, reimbursement, pricing and practice options, in response to the health,

aged care and wider social system challenges of baby boomer ageing. Historical evidence-based medicine and HTA approaches to research, reimbursement and pricing do not facilitate optimisation of health outcomes from constrained budgets until such non-patentable options are appropriately explored and compared alongside patentable technologies. Indeed, unless such options are adequately explored, HTA and EBM processes can be rightly accused of creating institutional barriers that promote selection bias in alternatives considered, which unduly privilege allocation of constrained research and reimbursement funding to new patentable technology. This lack of appropriate coverage of options leads to ill inform societal decision making in relation to reimbursement (adoption and financing) of new technology without appropriate consideration of opportunity costs (best alternative adoption and financing actions) associated with current programs and technology and budget constraints. That is, comparison with best alternative actions in adoption, namely the most cost effective expansion of existing programs and technology, and best alternative action in financing, contraction of least cost effective programs as the research of Pekarsky (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014) highlights, and as explored at length in Chap. 11.

Note that this does not imply that unexplained or anecdotal evidence of benefits from practice of health promotion and preventative strategies, herbal medicine or other eastern approaches and therapies such as tai chi, iridology, foot reflexology and acupuncture should be accepted on face value. It points to the need to undertake research to trial and test in practice whether and why such benefits arise in order to advance the health system toolkit and appropriate use. Rather than ignore such strategies and therapies, they should be researched where appropriate as promising approaches in the same way that promising new patentable therapies are – ideally with globally optimal trials. As Chaps. 6 and 7 highlight, explicitly allowing for evidence translation in optimising global trial design provides a first best option globally for translatable evidence as part of expected net gain maximisation, but also the ability to adopt and trial in optimising joint research and reimbursement decisions. Until promising non-patentable options have research which is resourced to compare with that of promising patentable options, a systemic institutional bias arises in processes of EBM. As Pekarsky (2015: 34) notes, both Arrow (1963) and Tirole (1988) conclude that the failure of the market to provide an incentive to invest in innovation for non-patentable strategies provides the economic case for public sector investment in researching and adopting such non-patentable strategies.

These fundamental coverage issues for avoiding selection bias in optimisation are further explored in developing methods which facilitate robust evidence generation and consideration of appropriate comparators and multiple modalities in:

(i) Chapter 4, highlighting methods for appropriate cost effectiveness evaluation of health promotion and prevention strategies.
(ii) Chapters 6 and 7, identifying method for optimal global trial design which enable feasible evidence collection for adopting and trialling with existing or promising new strategies while maximising global value relative to cost of trial designs and decision making.

(iii) Chapters 8 and 10, highlighting robust multiple strategy and outcome comparison methods on the cost disutility plane with expected net loss and frontier methods to best summarise cost effectiveness analysis in informing reimbursement and later research decisions.
(iv) Chapter 9, illustrating the net benefit correspondence theorem method for practice comparisons (Eckermann 2004; Eckermann and Coelli 2013), uniquely enabling comparison of the efficiency of providers in practice consistent with maximising net benefit, which in making coverage and comparability conditions explicit also provide a robust framework to avoid cost-shifting and cream-skimming incentives.
(v) Chapter 11, where health shadow price methods developed by Pekarsky (2012, 2015; Eckermann and Pekarsky 2014) are shown to provide appropriate incentives to collect evidence on best expansion and contraction of existing programs and technology alongside evidence of new technology. The health shadow price is also shown as key to establishing appropriate pricing of new technology and a pathway to allocative efficiency and budget-constrained optimisation with related research, reimbursement and regulation decisions.

In joining together these parts, and key principles and appropriate methods for individual and community approaches across settings, it is also important to note that in each case, as well as in combination, they require a longer-term attention span and wider perspective than short-term political or market-based reductionist approaches typically allow. We trust that those readers who stay the course will understand why partial and reductionist approaches are dangerous and obtain the fullest picture we can muster for key links between principles and methods for optimising research, reimbursement and regulation decisions. As a result, this text is not a cookery book telling you what to do in the next part of the recipe; however, the following mud map of chapters may aid those wanting to dive into a particular area to have some understanding of the whole. Figure 1.1 provides a map of the big picture, depicting decision making cycles for optimal joint research, reimbursement and regulation of practice and pricing decisions locally and globally, referencing related book chapters.

The robust process of problem definition, synthesis of cost, effects and costs effectiveness evidence and translation to inform net benefit estimation in jurisdictions of interest and assessment of optimal joint research and reimbursement decisions locally and globally allowing for relevant opportunity costs is iterative. Note that locally there are absorbing states for decision making cycles with rejection of strategies where the incremental net benefit (INB) is negative at the relevant jurisdiction shadow price for effects or sufficient evidence with adoption now optimal if INB is positive, while expected net gain from feasible research is not. Monitoring and regulation in practice is nevertheless indicated with adoption, while a lower price or changed evidence or conditions have potential to allow a strategy to become potentially optimal where INB is currently negative. More generally, the potential arises at the end of each research cycle to redefine questions in light of changing factors such as additional appropriate comparators and target populations in addition to updating evidence of relative treatment effects, baseline risks in translating evidence. Local factor prices and the health shadow price and associated INB and expected net gain

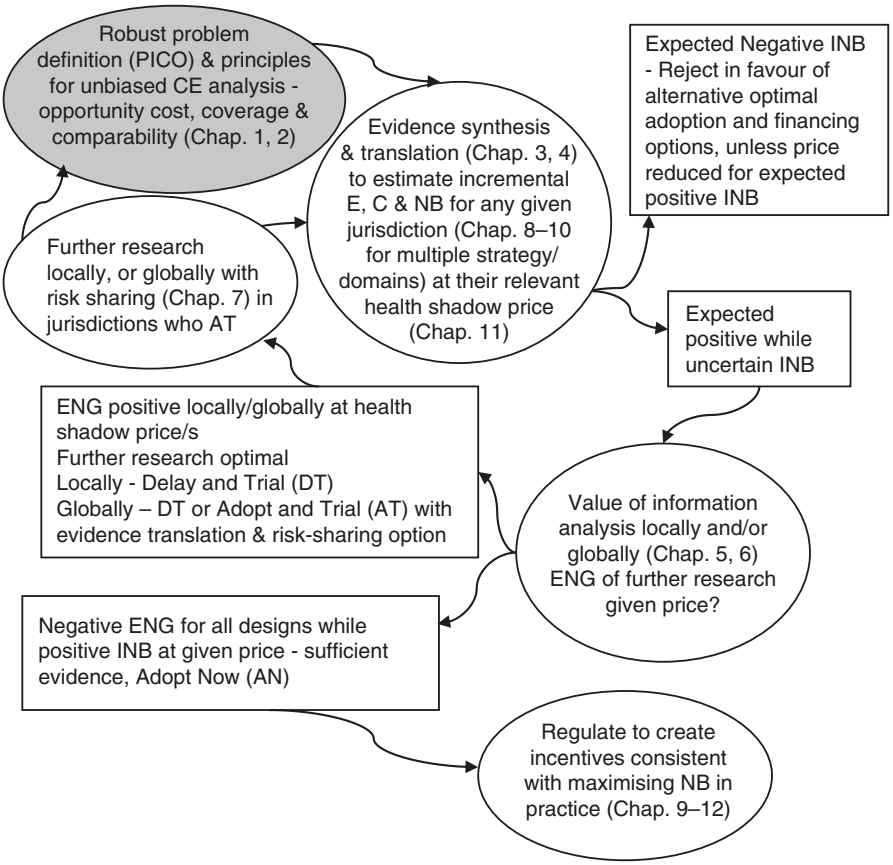


**Fig. 1.1** Optimal decision making cycles for joint research, reimbursement and regulatory processes locally and globally

(ENG) measures can also change to reflect local conditions. The appropriate health shadow price for any given set of decision contexts flowing through to appropriate threshold values for effects also clarifies that research into best use of existing programs and technologies is a priority for appropriate pricing of new technology as well as its own sake in optimising budget-constrained decision making.

## 1.4 Overview of Chapters

This chapter has introduced some central issues for health economic analysis to enable addressing constrained optimisation of societal decision making objectives informed by community values in a principled and robust way, whether in technology, program, policy or practice comparisons and in health promotion, preventative, curative, rehabilitative or palliative settings.

Chapter 2 further cements coverage and comparability principles as the robust foundation underlying unbiased decision making and health economic analysis and starts to consider robust approaches and methods to inform unbiased cost effectiveness analysis and adoption decisions. Satisfying coverage and comparability principles to avoid biases in undertaking cost effectiveness analysis is shown to require jointly considering adequate scope and duration of downstream cost and health effect impacts across strategies compared and relative treatment effect(s).

The advantages that the net benefit metric has over incremental cost effectiveness ratios (ICERs) – in summarising cost effectiveness evidence for such decisions – are illustrated and shown as particularly important when allowing for decision uncertainty. Useful presentation and summary measures for comparison under uncertainty of costs and effects of two strategies, the incremental cost effectiveness plane and net benefit and cost effectiveness acceptability curves, are introduced and illustrated for trial-based analysis. These presentation and summary measures are shown to be simply constructed in appropriately allowing for joint cost and effect distributions non-parametrically with bootstrapping and parametrically with Fieller's method. The need for joint consideration of costs and effects in avoiding bias and inferential fallacies when informing decisions under uncertainty is highlighted with consideration of seminal papers including 'The death of cost minimisation' (Briggs and O'Brien 2001) and 'Thinking outside the box' (Briggs et al. 2002). These papers also begin to point to more general problems of bias with reductionist approaches, a theme which is expanded on in:

(i) Chapter 3 for modelled cost effectiveness analysis;
(ii) Chapters 5, 6 and 7 for value of information (VOI) analysis;
(iii) Chapters 4, 8 and 10 for multiple strategy and outcome comparisons;
(iv) Chapter 9 in efficiency measurement across providers in practice consistent with maximising net benefit; and
(v) Chapters 11 and 12 in appropriately considering alternative actions for identifying the opportunity costs of investing in, and pricing of, new technology.

Chapter 3 highlights some further common problems and dangers in inherently inconsistent and biased methods for modelled cost effectiveness analysis where coverage and comparability principles are violated with choice of methods and metrics in synthesising, translating and extrapolating evidence. These are illustrated with inferential fallacies and inconsistencies arising with use of relative risk in indirect comparison (Eckermann et al. 2009) and translation of evidence (Eckermann et al. 2011). They are also illustrated with parametric methods in extrapolation of costs, effects and cost effectiveness inconsistent with indication or associated factors over time such as compliance, resistance and side effects, as well as in inconsistent extrapolation across cost and effects. More importantly, methods which solve these problems are identified in each case. Odds ratio methods are shown to enable unbiased consistent estimates with alternative framing of outcomes in indirect comparison and translation (Eckermann et al. 2009, 2011). Decision analytic modelling approaches with extrapolated treatment effects conditional on indication, continuation rules and compliance and side effect profiles in surviving populations in practice are indicated to allow unbiased and consistent extrapolation of costs, effects and cost

effectiveness. Solutions to these problems also serve to illustrate the need for complementary approaches to health economic evaluation with trial-based and model-based evaluation to allow evidence relevant to decision making in a jurisdiction of interest such as that of the Pharmaceutical Benefit Advisory Committee (PBAC) in Australia (Commonwealth of Australia 2016), where the seminal paper 'Frankenstein's Monster or the Vampire of Trials' (O'Brien 1996) takes centre stage.

In general, marrying coverage and comparability principles are required to avoid biases in divining how cost effectiveness presentation and summary measures under uncertainty can be robustly applied. These principles are illustrated with two-strategy comparison for modelled analysis in Chap. 3, while robust presentation and summary measures for more than two-strategy comparison are identified and illustrated in Chap. 8 for multiple strategies and additionally with multiple outcomes in Chap. 10. Problems of partialisation and failure to reflect community values are also pointed to as particularly important considerations in prevention and health promotion strategies in complex community settings, issues which Chaps. 4 and 12 expand on.

Chapter 4 explores some of the challenges faced when undertaking health economic analysis in comparing prevention and health promotion strategies in complex community settings such as schools and palliative care settings with multiple domain comparisons, and some principled approaches and methods developed to address these challenges. Evaluating community-based primary prevention programs makes clear that the principles and evaluation approach to health system decision making need to consider community population impacts over time. Conventional within-study cost effectiveness and extrapolated modelling methods are shown to struggle within typical short-term evaluation time frames to appropriately assess or tractably capture or model community acceptance or the diffusion of impacts over time in populations across community networks conditional on health promotion strategy acceptance. Hence, the need is shown for alternative evaluation methods in navigating coverage (scope and duration) and comparability of the acceptance, diffusion and incremental impact of prevention and health promotion strategies. The research of Shiell and Hawe, pointing to the value of assessing network multiplier impacts from investment on community activity, is illustrated as a more robust and appropriate approach to informing decision makers of the long-term acceptance and success of community-based health promotion and prevention interventions. In modelling terms, such multipliers and their trajectory over time represent the key prognostic factors, or surrogates, for long-term acceptance and success of community-based health promotion and prevention programs and their network impacts over time. Multiplier methods for assessing complex interventions are illustrated in evaluating the Stephanie Alexander Kitchen Garden National Program (SAKGNP), a health promotion and primary prevention program undertaken in primary schools (Eckermann et al. 2014; Yeatman et al. 2014).

The research of McCaffrey et al. (2010, 2013, 2015) is also highlighted in Chap. 4 as enabling robust comparison of multiple outcome domains under uncertainty and illustrated in greater detail with associated methods in Chap. 10. Multiple outcome domain comparisons are shown to be valuable in many settings to consider diffuse outcomes beyond single health metrics that inform wider community utility functions

but also alternative values and domain aspects of utility within health. This is particularly the case in areas such as palliative care where domains such as finalising affairs and process of death are not amenable to being integrated with survival time and hence unable to be incorporated into quality-adjusted life years. Further, even within a quality-adjusted life year (QALY) framework, significant value to decision makers in many circumstances arises from being able to explicitly present multiple events or effects underlying QALY estimates and robustly consider their joint uncertainty. Such analysis allows the potential for baseline risk of effect and/or utility weights for states or domains of effect to differ across populations and jurisdictions, as well as over time.

In Chap. 5, optimal decision making in relation to evidence-based reimbursement of technologies based on their incremental cost effectiveness (net benefit) under uncertainty, is shown to be inextricably linked to research decisions. Frequentist approaches to trial design such as the use of type I error, type II error and minimum significant difference to power hypothesis tests don't consider or reflect the expected value or expected cost of information and hence are unable to efficiently design trials or optimally inform such decisions. Bayesian methods are shown to enable joint optimisation of research and reimbursement decisions with robust estimation of expected value and cost of further research to decision makers' conditional on critical decision contexts given prior uncertainty in incremental net benefit and as a function of trial size and designs.

Nevertheless, to estimate the distribution of INB and undertake meaningful value of information (VOI) analysis in any jurisdiction of interest, unbiased estimates of incremental costs and effects (following Chaps. 2, 3 and 4) and a meaningful threshold value for effects are required. Hence, Chaps. 2, 3 and 4 should be considered alongside Chap. 11 in deriving a robust estimate of where the INB distribution lies given local decision contexts before undertaking VOI analysis such as that in Chaps. 5, 6 and 7. That is, an unbiased estimate of the expected value of INB is primary to informing societal decision making under the Arrow-Lind theorem (Arrow and Lind 1970) before consideration of uncertainty. While estimating expected INB is the key information decision makers require to assess reimbursement decisions, it also informs the location of tail distribution and associated estimation of expected value of sample information (EVSI) and any opportunity cost of delaying a decision to adopt while research is undertaken.

Value of information (VoI) principles and methods enabling optimisation of expected net gain (expected value less costs) of local trial design and decision making are identified and illustrated in Chap. 5. Importantly, central limit theorem (CLT)-based VOI methods presented are shown to be both:

(i) Simply applied in estimating expected value of actual trial designs (expected value of sample information) given estimates of mean cost and effects of their variance and covariance; and
(ii) Allow for relevant decision contexts that jurisdictions face in estimating expected value and cost to make these decisions locally, including rate of recruitment, follow-up and analysis time, opportunity cost and option value of delay and imperfect implementation.

Hence these CLT methods are shown to satisfy Occam's Razor in realtion to VOI methods (Eckermann et al. 2010), enabling simple optimisation of ENG under relevant decision contexts, providing the necessary and sufficient conditions to locally inform decisions including:

(i) Is further research for a specific HTA potentially worthwhile?
(ii) Is a given research design worthwhile?
(iii) What is the optimal research design?
(iv) How can funding best be prioritised across alternative research proposals?

The ability to optimise ENG, while allowing for key decision contexts in addressing these questions, is particularly suggested to better inform research grant allocation bodies who have mission statements emphasising 'value for research dollar', 'efficiency in research design' and 'research making a difference to practice'. From a researcher perspective, research designed to address decision making (DM) uncertainty and relevant DM contexts in maximising value relative to cost or ENG from limited budgets connect with decision making and funding bodies underlying objectives. Hence, VOI principles and methods enabling optimising ENG return on research should also increase research chances of success given the centrality of these factors to research funder aims, mission and objective statements, as well as budget-constrained expected impact on policy and implementation.

In Chap. 6, the methods for optimally and efficiently informing joint research and reimbursement decisions locally identified in Chap. 5 are extended to allow for optimal global trial design and local decision making across jurisdictions (Eckermann and Willan 2009). The ability to adopt and trial in jurisdictions as part of a global trial is shown to be particularly advantageous in moving from the local to global setting and avoiding opportunity costs of delay while obtaining best evidence globally. Optimally designed global trials allow promising technologies to be adopted early in jurisdiction to avoid opportunity costs of delay for societal decision makers and manufacturers alike. Such advantages of jurisdictions adopting and trialling with promising therapies arise provided evidence translates from other jurisdictions who undertake research, creating appropriate requirements and incentives for evidence coverage, which are made explicit in optimal global trial design and early adoption assessment.

In Chap. 7, the methods for optimal societal decision maker trial design and joint research and reimbursement decision making in Chaps. 5 and 6 are extended to allow for pricing under uncertainty locally (Willan and Eckermann 2012) and risk sharing in the case of jurisdictions who adopt as part of an optimal global manufacturer trial design (Eckermann and Willan 2013). Optimally designed global trials with explicit consideration of evidence translation in trial design are shown to allow earlier adoption of promising programs or technologies while this evidence is collected, with the ability to feasibly adopt and trial with translatable evidence. Further, the greater strength of evidence from larger trials expected *a priori* is also expected to result in improved implementation (Willan and Eckermann 2010). Such optimal global trials also provide the ability to feasibly and meaningfully risk share for jurisdictions who adopt and trial, with price changes able to be informed by

prospective randomised controlled trial (RCT) evidence from the global trial and local evidence of performance of the new technology in practice (Eckermann and Willan 2013).

The ability to feasibly adopt and trial and risk share in such optimally designed global trials better aligns societal decision maker and manufacturer research interests for best research design globally and evidence translation across jurisdictions. Further, such optimal global trial designs also provide an option for feasible collection of RCT evidence for existing strategies which have already been adopted, key to informing opportunity cost and health shadow price assessment, as highlighted in Chap. 11 and policy options in Chap. 12.

Chapter 8 moves beyond two-strategy cost effectiveness analysis to allow for multiple strategy cost effectiveness comparison, presentation and summary measures. Such multiple strategy comparisons are increasingly important with multiple treatment modalities, diagnostic and treatment options and combinations of customised strategies such as genetic testing and initiatives towards individualised care. When comparing multiple strategies, presentation on the cost disutility plane and use of expected net loss (ENL) curves and frontiers are shown to overcome limitations of methods for two-strategy comparison on the C-E plane with CEA and NB curves (Eckermann and Willan 2011; Eckermann et al. 2008).

When comparing multiple strategies, the optimal strategy for comparison is not fixed, as in two-strategy comparisons, but rather changes across replicates and/or threshold values. Flexible axes on the cost disutility plane explicitly addresses this, overcoming problems of fixed axes on the cost effectiveness plane and associated confounding of graphical inference (Eckermann and Willan 2011). Similarly, the expected net loss (ENL) statistic and associated summary measures enable flexible while consistent comparison of differences in expected net benefit with the optimal strategy in any given replicate at any threshold value. Hence, ENL curves and frontiers are shown to overcome problems of CEA curves and frontiers not presenting differences in expected net benefit and the fixed nature of the comparator with incremental net benefit statistics.

Consequently, for multiple strategy comparisons, ENL curves and frontiers are illustrated to fully inform asymptotically risk neutral societal decision making under the Arrow-Lind theorem. If societal decision makers are somewhat risk averse, ENL curves and the ENL frontier provide primary evidence of expected values for making decisions which can be supplemented by appropriate uncertainty evidence. Such evidence is highlighted as needing to be derived from bilateral CEA curves between potentially optimal strategies of interest to prevent confounding of probabilities from other strategies. Further, the ENL frontier identifies both the strategy minimising ENL (equivalently maximising expected net benefit (ENB)) across strategies at any given threshold value and the per-patient potential value of future research. That is, the ENL curve also represents the expected opportunity loss that could be avoided with perfect information and hence the expected per-patient value of perfect information. Thus, the ENL frontier makes explicit the link between optimal reimbursement and research, further supporting the joint nature of research and reimbursement decisions locally and globally, as highlighted in Chaps. 5, 6 and 7.

Chapter 9 shows how the advantages of comparing multiple strategies consistent with maximising net benefit on the cost-disutility plane in Chap. 8 naturally extend to efficiency measures across providers in practice consistent with maximising net benefit. The net benefit correspondence theorem (NBCT) providing the robust theoretical framework underlying methods in Chaps. 8, 9 and 10 is derived. In efficiency comparisons in practice, the NBCT is shown to uniquely provide explicit and joint consideration of the value and costs of quality of care in efficiency measures consistent with maximising net benefit. This overcomes problems of conventional efficiency measures in practice, such as cost per case-mix-adjusted admission in hospitals, implicitly including cost of quality while ignoring the value of quality of care and hence creating incentives for cost minimising quality of care.

More generally, the one-to-one correspondence of the NBCT underlying efficiency comparison with radial properties in cost-disutility space is shown to provide distinct advantages over alternative specifications (Eckermann 2004; Eckermann and Coelli 2013) in enabling:

(i) Identification of net benefit maximising peers over threshold value for effects where they maximise NB;
(ii) Practice and policy relevant net benefit (economic) efficiency of providers and decomposition into technical, allocative and scale efficiency consistent with maximising net benefit; and
(iii) Shadow price for service quality across industry behaviour without requiring prices for admissions.

Importantly, coverage and comparability conditions of the NBCT are also shown to provide an accountable framework to prevent cost-shifting and cream-skimming incentives in practice (Eckermann 2004; Eckermann and Coelli 2013). These explicit coverage and comparability conditions continuously support evidence-based approaches to joint accountability for cost and quality including risk adjustment and standardisation methods (Eckermann et al. 2009, 2011) and data linkage and/or modelling of expected effects beyond service to a common meaningful time point (e.g. 30 days or 1 year beyond admission in hospital).

The NBCT as a generalised method can more generally be applied to efficiency measure with these advantages in any health, care, service or industry setting where maximising net benefit is the appropriate economic objective. Further, radial properties on the cost-disutility (C-DU) plane enable robust comparison, presentation and summary measures for as many domains of effect as appropriate, as highlighted in Chap. 10 following the research of McCaffrey et al. (2013, 2014, 2015).

Chapter 10 shows how the framework presented in Chaps. 8 and 9, for optimal comparison across multiple strategies or providers' costs and effects on the incremental cost-disutility plane and expected net loss curves and frontiers, naturally extends to multiple effect comparisons. Radial properties on the cost disutility plane enable robust comparison of multiple outcomes under uncertainty, providing distinct advantages over cost consequences analysis. That is, allowing for uncertainty in joint consideration of multiple effects and cost effectiveness analysis summary measures to avoid inferential problems of partial analysis with single effect comparison and summary measures. Summary measures of ENL planes and surfaces

and cost effectiveness analysis (CEA) planes developed by McCaffrey (McCaffrey et al. 2010, 2013, 2015) are shown to have further distinct advantages over conventional methods in presenting cost effectiveness across multiple effects and potential threshold values for multiple domains of effect.

Chapter 11 addresses the requirement of net benefit assessment across joint research, reimbursement and regulatory decisions (Chaps. 2, 3, 4, 5, 6, 7, 8, 9 and 10), for threshold values for effects that reflect the opportunity costs (best alternative actions) for relevant decision contexts in the jurisdiction of interest to enable budget-constrained optimisation. The research of Pekarsky (2012, 2015) is highlighted in deriving the health shadow price of reimbursement (adoption and financing) actions for investments with net incremental cost (NW quadrant on the C-E plane) under characteristic health system allocative and displacement inefficiency conditions. The best alternative action to adopting a new technology financed with displacement of programs (ICER = $d$) is adopting the most cost effective expansion of existing programs (ICER = $n$) financed by contraction of the least cost effective existing program (ICER = $m$) leading to a health shadow price for effects of

$$\beta_c = \left(1/n + 1/d - 1/m\right)^{-1}$$

where the subscript $c$ refers to the prevailing economic context in the jurisdiction of interest.

Importantly, this health shadow price reflects conditions of allocative inefficiency ($n < m$) and displacement inefficiency ($d < m$) characteristic of current health systems that can be improved with optimal decision making. The implications of the health shadow price in providing a pathway to allocative efficiency and addressing market failure in provision of evidence for $n$, $d$ and $m$ are discussed following Eckermann and Pekarsky (2014).

Shadow prices are also considered for the less usual case, on the south-west (SW) quadrant, where new investment is expected to lead to health system cost savings while being potentially less effective. The opportunity cost of decisions to invest in new technology on the SW quadrant, expected to generate net funding relative to current practice over time while some potential health loss, is shown to differ qualitatively as well as quantitatively (Eckermann 2015). If the budget is free to contract, strategies on the SW quadrant should be compared with the least health-reducing way of generating funds for the health budget, and hence the health shadow price reflects an ICER of m. However, where budgets are fixed from going down as well as up, then funds raised by such cost-saving technologies are required to be spent on adoption of other programs. In that case, the health shadow price on the SW quadrant in generating funds ($\beta_f$) is shown to be derived equating returns of funding generated with the cost-saving technology and adopted with ICER $a$ and that of the best alternative fund generating and adoption actions, leading to:

$$\beta_f = \left(1/a + 1/m - 1/n\right)^{-1}$$

Where adoption is efficient ($a = n$), then $\beta_f = m$, while if adoption is inefficient ($a > n$) as with threshold based on displaced services, then $\beta_f$ is greater than m, or

can even be required to be dominant to be optimal, as illustrated for the UK with current adoption thresholds. In general, a kink in the economically meaningful threshold value is shown to arise under characteristic health system conditions of allocative and displacement inefficiency, where the threshold value is higher in the SW relative to NE quadrant (Eckermann 2015). The extent of this kink reflects the degree to which there is allocative and displacement inefficiency.

Chapter 12 highlights application of health economic principles and methods to address the challenge of budget-constrained successful ageing of baby boomers with publicly provided universal access health systems (Eckermann 2014a, b; Eckermann et al. 2016; Eckermann and Sheridan 2016) in Australia and internationally. This points to the need for reform that addresses historical inefficiencies across the spectrum from prevention to palliative care including:

(i) Community age and dementia-friendly policies to successfully age while minimising the need for aged care and nursing home facilities in line with community health promotion and prevention considerations from Chap. 4;
(ii) Dementia-friendly aged care and nursing home design and care, illustrating the key need for better use of factor priced environmental design approaches to better care for and meet community needs and preferences;
(iii) Effective factor-priced promising palliative care options to address palliative care primary preferences for key palliative domains – finalising affairs in community of choice while minimising family and carer distress – identified in Chaps. 4 and 10 – in particular policies for optimising net benefit of medicinal cannabis cultivation and program provision and a promising reformulation of 5-FU; and
(iv) Extending NBCT efficiency measures from Chap. 9 to funding mechanisms in providing active incentives for budget-constrained health and aged care system net benefit optimising quality of care rather than for minimum cost per service quality of care, cost shifting and cream skimming with current case-mix funding methods.

Implications are drawn in each case for optimal policy direction and options and methods that should be adopted to support better joint research, reimbursement and regulatory societal decisions made locally and internationally.

Chapter 13 concludes, showing how a principled approach to health economic evaluation and research can optimise community objectives under resource and budget constraints, but only where key bigger picture structural issues are jointly addressed for research, reimbursement and regulation (pricing, performance monitoring and funding). Optimisation and robust analysis with health economic-related decision making requires satisfying coverage and comparability principles in addressing research, reimbursement and regulatory decisions in HTA and practice. The need to systematically address critical weaknesses of the current political economy in research, reimbursement and regulation biasing towards new technology and away from better use of existing technology is identified. The failure of community preferences to be reflected in resource allocation and policy making in key areas such as palliative and end of life care are also highlighted.

Optimal global trials with coverage of evidence translation and the ability to adopt and trial with use of the NBCT to monitor performance in practice while providing evidence to enable robust risk share are suggested as a first best solution that overcomes many otherwise intractable joint decision making and political economy issues. In addition to optimising joint research, reimbursement and regulation decisions across jurisdictions globally, such designs in avoiding opportunity costs of delay while providing best evidence for decision making also better align societal decision maker and provider (non-patented programs, strategies or technologies) or manufacturer (patented products) interests. Critically such global trials would also enable an optimal pathway to providing evidence for existing or new technologies required to inform health shadow prices in any jurisdiction and optimise budget constrained reimbursement decisions and pricing of new technology given best alternative actions. Importantly, this provides a pathway towards allocative and displacement efficiency with appropriate research, reimbursement and regulatory incentives.

## References

Akerlof GA. The market for "lemons": quality uncertainty and the market mechanism. Q J Econ. 1970;84(3):488–500.

Arrow KJ. Uncertainty and the welfare economics of medical care. Am Econ Rev. 1963;LIII(5): 941–73.

Arrow KG, Lind RC. Uncertainty and the evaluation of public investment decision. Am Econ Rev. 1970;60:364–78.

Briggs A, O'Brien B. The death of cost minimization analysis. Health Econ. 2001;10:179–84.

Briggs A, O'Brien B, Blackhouse G. Thinking outside the box: recent advances in the analysis and presentation of uncertainty in cost effectiveness studies. Annu Rev Public Health. 2002;23:377–401.

Briggs A, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM modelling good research practices task force working group-6. Med Decis Mak. 2012;32(5):722–32.

Carter GT, Flanagan AM, Earleywine M, et al. Cannabis in palliative medicine: improving care and reducing opioid-related morbidity. Am J Hosp Palliat Care. 2011;28:297–303.

Claxton K, Posnett J. An economic approach to trial design and research priority-setting. Health Econ. 1996;5(6):513–24.

Commonwealth of Australia. 2016. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee, Version 5. Accessed 20 Oct 2016 at: https://pbac.pbs.gov.au/.

Davis K, Stremikis K, Squires D, Schoen C. Mirror, mirror on the wall: how the performance of the U.S. health care system compares internationally: 2014 update. New York: The Commonwealth Fund pub; no. 1755. 2014. Accessed 20 Oct 2016 at: http://www.commonwealthfund.org/~/media/files/publications/fund-report/2014/jun/1755_davis_mirror_mirror_2014.pdf.

Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 2005.

Eckermann S. Hospital performance including quality: creating incentives consistent with evidence-based medicine. PhD Dissertation. Sydney: UNSW; 2004. http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/.

Eckermann S. Over and underservicing: further reasons to scrap the GP co-payment. The Conversation. 2014a; http://theconversation.com/over-and-under-servicing-further-reasons-to-scrap-the-gp-co-payment-30199.

Eckermann S. Avoiding a health system hernia and the associated outcomes and costs. ANZJ Public Health. 2014b;38(4):303–5.
Eckermann S. Kinky thresholds revisited: opportunity costs differ in the NE and SW quadrants. Appl Health Econ Health Policy. 2015;13:7–13. doi:10.1007/s40258-014-0136-3.
Eckermann S, Briggs A, Willan A. Health technology assessment in the cost-disutility plane. Med Decis Mak. 2008;28:172–81.
Eckermann S, Coelli T. Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice. Soc Sci Med. 2013;76:159–68. http://dx.doi.org/10.1016/j.socscimed.2012.10.020
Eckermann S, Coory M, Willan AR. Indirect comparison: relative risk fallacies and odds solution. J Clin Epidemiol. 2009;62:1031–6. doi:10.1002/hec.1353.
Eckermann S, Karnon J, Willan A. The value of value of information: best informing research design and prioritization using current methods. Pharmacoeconomics. 2010;28(9):699–709.
Eckermann S, Coory M, Willan A. Consistently estimating absolute risk difference when translating evidence to jurisdictions of interest. Pharmacoeconomics. 2011;29(2):87–96.
Eckermann S, Dawber J, Yeatman H, Quinsey K, Morris D. Evaluating return on investment in a school based health promotion and prevention program: the investment multiplier for the Stephanie Alexander Kitchen Garden National Program. Soc Sci Med. 2014;114:103–12. doi:10.1016/j.socscimed.2014.05.056.
Eckermann S, Pekarsky B. Can the real opportunity cost stand up: displaced services, the straw man outside the room. Pharmacoeconomics. 2014;32(4):319–25.
Eckermann S, Sheridan L. Supporting medicare health, equity and efficiency in Australia: policies undermining bulk billing need to be scrapped. App Health Econ Health Policy. Published online 20 July 2016;14:511–4. Available at http://link.springer.com/article/10.1007%2Fs40258-016-0258-x.
Eckermann S, Sheridan L, Ivers R. Which direction should Australian health system reform be heading? ANZJPH. 2016;40(1):7–9.
Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16:195–209.
Eckermann S, Willan AR. The option value of delay in health technology assessment. Med Decis Making. 2008a;28:300–5.
Eckermann S, Willan AR. Time and EVSI wait for no patient. Value Health. 2008b;11:522–6.
Eckermann S, Willan AR. Globally optimal trial design for local decision making. Health Econ. 2009;18:203–16.
Eckermann S, Willan A. Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies. Pharmacoeconomics. 2011;29(7):563–77.
Eckermann S, Willan A. Optimal global VOI trials: better aligning manufacturer and decision maker interest and enabling feasible risk sharing. PharmacoEconomics. 2013;31:393–401.
Gallily R, Yekhtin Z, Ondřej HL. Overcoming the bell-shaped dose-response of cannabidiol by using cannabis extract enriched in cannabidiol. Pharmacol Pharm. 2015;6:75–85.
Graham DA. Cost-benefit analysis under uncertainty. Am Econ Rev. 1981;71:715–25.
Graham D. Public expenditure under uncertainty: the net- benefit criteria. Am Econ Rev. 1992;82:822–46.
Hawe P, Shiell A. Social capital and health promotion: a review. Soc Sci Med. 2000;51(6):871–85.
Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. Am J Community Psychol. 2009;43(3–4):267–76.
Johnson JR, Burnell-Nugent M, Lossignol D, Ganae-Motan ED, Potts R, Fallon MT. Multicenter, double-blind, randomized, placebo-controlled, parallel-group study of the efficacy, safety, and tolerability of THC: CBD extract and THC extract in patients with intractable cancer-related pain. J Pain Symptom Manag. 2010;39(2):167–79.
Kalache A. The Longevity Revolution: Creating a society for all ages. Adelaide Thinker in Residence 2012-2013. Adelaide: Government of South Australia; 2013. Available at: http://www.flinders.edu.au/sabs/fcas-files/Publications/The%20Longevity%20Revolution.pdf.

McCaffrey N, Karnon J, Currow D, Eckermann S. The old dog and the poor relation: informing funding decisions in palliative care with cost-consequence analysis in the cost-disutility plane. Palliative Med. 2010;24(4):S207.
McCaffery N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Is home-based palliative care cost effective? An economic evaluation of the Palliative Care Extended Packages at Home (PEACH) pilot. BMJ Support Palliat Care. 2013;3(4):431–5.
McCaffrey N, Skuza P, Breaden K, Eckermann S, Hardy J, Oaten S, Briffa M, Currow D. Preliminary development and validation of a new end-of-life patient-reported outcome measure assessing the ability of patients to finalise their affairs at the end of life. PLoS One. 2014;9(4):e94316.
McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Better informing decision making with multiple outcomes cost-effectiveness analysis under uncertainty in cost-disutility space. PLoS One. 2015;10(3):e0115544.
McGuire A, Henderson J, Mooney GH. The economics of health care: an introductory text. London: Routledge & Kegan Paul; 1988.
Mooney G. The health of nations: towards a new political economy. London: Zed books; 2012.
Moore S, Haines V, Hawe P, Shiell A. Lost in translation: a genealogy of the "social capital" concept in public health. J Epidemiol Community Health. 2006;60(8):729–34.
O'Brien B. Economic evaluation of pharmaceuticals. Frankenstein's monster or vampire of trials? Med Care. 1996;34(12 Suppl):DS99–108.
Organisation for Economic Co-Operation and Development. Health at a glance 2013: OECD indicators, OECD, Publishing. 2013. http://dx.doi.org/10.1787/health_glance-2013-en, http://www.oecd.org/els/health-systems/Health-at-a-Glance-2013.pdf. Accessed 23 May 2014.
Pekarsky BAK. Trusts, Constraints and the Counterfactual: Reframing the Political Economy of New Drugs. PhD Dissertation. 2012. University of Adelaide. http://digital.library.adelaide.edu.au/dspace/handle/2440/79171.
Pekarsky BAK. The new drug reimbursement game: a regulator's guide to playing and winning. New York: Springer; 2015.
Russo EB. Taming THC: potential cannabis synergy and phytocannabinoid-terpenoid entourage effects. Br J Pharmacol. 2011;163(7):1344–64.
Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. Br Med J. 2008;336(7656):1281–3.
Simon H. A behavioral model of rational choice. In: Models of man, social and rational: mathematical essays on rational human behavior in a social setting. New York: Wiley; 1957.
Stinnett AA, Mullahy J. Net health benefits: a new framework for analysis of uncertainty in cost-effectiveness analysis. Med Decis Making. 1998;17(4):483–9.
Tirole J. The theory of industrial organisation. Cambridge: MIT; 1988.
Wagner H, Ulrich-Merzenich G. Synergy research: approaching a new generation of phytopharmaceuticals. Phytomedicine. 2009;16:97–110.
Weinstein MC, Stason WB. Foundations of cost effectiveness analysis for health and medical practices. N Engl J Med. 1977;296(13):716–21.
Willan A, Briggs A. The statistical analysis of cost-effectiveness data. Chichester: Wiley and Sons; 2006.
Willan AR, Eckermann S. Optimal clinical trial design using value of information with imperfect implementation. Health Econ. 2010;19:549–61.
Willan AR, Eckermann S. Value of information and pricing new health care interventions. Pharmaco Econ. 2012;30(6):447–59.
Willan AR, Lin DY. Incremental net benefit in randomized clinical trials. Stat Med. 2001;20:1563–74.
Woolhandler S, Campbell T, Himmelstein DU. Costs of health care administration in the United States and Canada. NEJM. 2003;349:768–75.
Yeatman H, Quinsey K, Dawber J, Nielsen W, Condon-Paoloni D, Eckermann S, Morris D, Grootemaat P, Fildes D. Combining realism with rigour: evaluation of a National Kitchen Garden program in Australia primary schools. Evaluation Journal of Australasia. 2014;14(2): 17–24.

# Part I
# Principles and Practice for Robust Net Benefit Analysis Informing Optimal Reimbursement (Adoption and Financing) Decisions Across Individual and Community-Focused Programs Using Trial, Model and Network Multiplier Methods

This text aims to identify health economic principles and methods for optimising societal decision making across health system reimbursement (adoption and financing), research and regulation questions in the context of community objectives and budget constraints, as the book's title suggests and Chap. 1 introduced in overview. Part I (Chaps. 2, 3, and 4) considers the simplest cases for such analysis where only a reimbursement question is considered, and comparison is restricted to that between two strategies (typically a new therapy and current usual practice) in relation to their costs and a single measure of effect.

Decision analytic principles of coverage and comparability are shown as the key joint considerations for unbiased analysis and decision making in any given jurisdiction/s with such comparisons. In Chaps. 2 and 3, the importance of decision analytic coverage and comparability principles are highlighted for trial and model based analysis of incremental cost-effectiveness or net benefit. In Chap. 4, coverage and comparability conditions are shown to be just as, if not more, important in evaluating community interventions such as health promotion and palliative care program. Coverage of network impacts across communities over time is key to assessing success and cost-effectiveness but also allowing fair comparison with evaluation of individual patient-focused interventions (e.g. medical procedures or medications for disease management, primary or secondary prevention of symptoms and in rare cases cure) considered in Chaps. 2 and 3.

Chapter 2 introduces decision analytic coverage and comparability principles alongside net benefit assessment of joint health system costs and effects along alternative treatment pathways to allow unbiased cost-effectiveness analysis informing societal decision making. While trial-based cost-effectiveness analysis methods are outlined in Chap. 2, they are never the less pointed to as usually requiring adaption to local, rather than trial, conditions in informing unbiased societal decision making for any given jurisdiction of interest.

The seminal paper of Bernie O'Brien 'Frankenstein's Monster or the Vampire of Trials' (O'Brien 1996) takes centre stage in Chap. 3, highlighting the weaknesses of randomised trial and model-based health economic analysis in isolation and need

for joint consideration of coverage and comparability principles to avoid dangerous biases that otherwise arise. Randomised trial-based analysis can satisfy comparability of evidence, provided appropriate interventions or treatment strategies of interest are compared within trial and randomisation is adequate. However, the coverage of relative treatment effects for clinically important side effects and resource use is often absent, and trial evidence and protocol conditions will usually differ from that of decision making jurisdiction/s of interest on one or more elements of population risk, practice and preferences. Hence, trial evidence like puncture marks of a vampire are localised to protocol conditions by arm and usually require translation, with trial-based analysis methods in Chap. 2 rarely able to directly inform INB assessment and usually require translation to avoid the Vampire of trials. Jointly satisfying comparability and coverage principles usually require a marriage between trial-based analysis and modelling, the Vampire and Frankenstein in O'Brien (1996), to enable robust analysis in such settings.

Model-based analysis, can address translation of trial evidence of treatment effect to the absolute effect expected in a jurisdiction of interest, and more generally coverage of incremental effects important to incremental net clinical benefit assessment, as well as resource use and net benefit. Such coverage can also be in terms of adequate duration as well as the scope of effect in robustly estimating incremental net benefit, and hence modelling also involves extrapolation methods. However, in synthesising, translating and extrapolating, evidence modelling needs to be critically aware of common biases that arise where comparability issues are not addressed. Chapter 3 highlights common biases and Frankenstein's monster issues that arise in cost minimisation analysis, indirect comparison, evidence translation and extrapolation and more importantly illustrates robust unbiased methods for evidence synthesis, translation and extrapolation that avoid these biases.

In Chap. 4, evaluation in complex community setting with health promotion or palliative care are shown to require moving beyond conventional individual based within study analysis, to satisfy coverage and comparability principles or more generally enable adequate assessment of the success or otherwise of alternative options. In health promotion settings, jointly satisfying coverage and comparability principles is shown to primarily require assessment of community acceptance, ownership and network impacts of programs or strategies, following the research of Hawe and Sheill. Undertaking network analysis such as multiplier impacts of program investment on the value of health promotion activity generated across communities over time are shown as key to assessing community ownership but also assess long-term success and cost effectiveness of such programs. These multiplier assessments become particularly powerful where triangulated with short-term assessment of attitude and behavioural impacts in target population and qualitative assessment of enablers and barriers to community acceptance. Triangulation across these methods is illustrated in a primary school setting with evaluation of the Stephanie Alexander Kitchen-Garden National Program (SAKGNP) in Australia.

In palliative care settings, assessment of incremental net benefit and community preferences are shown to require consideration of multiple effect domains unable to be integrated with patient survival in quality adjusted life year (QALY) assessment.

Indeed, domains such as finalising personal and financial affairs, undertaking palliative processes of death in community setting of choice (most often at home) and family distress and carer impacts are primary considerations in such palliative settings. Hence, robust methods for multiple domain comparisons (considered in detail in Chap. 10) are pointed to as critical to enable appropriate decision making that reflect key domains for palliative preferences in such settings and associated decisions to optimise incremental net benefit across alternative interventions and strategies.

Coverage and comparability principles introduced in Part I are also pointed to as key building blocks to later more complex consideration of:

(i) Joint optimal research and reimbursement decisions in allowing for NB evidence under uncertainty (Part II, Chaps. 5, 6, and 7)
(ii) Robust methods for multiple strategy, provider efficiency in practice and multiple domain of effect comparisons to best inform societal decision making in those more complex settings (Part III, Chaps. 8, 9, and 10)
(iii) Budget-constrained optimal joint research, reimbursement and regulation decisions with the health shadow price of Pekarsky (2012, 2015) providing a pathway for optimising budget-constrained use of existing programs and new therapies (Part IV, Chaps. 11 and 12), while recognising a starting point of characteristic inefficiencies in current health system practice

## References

O'Brien B. Economic evaluation of pharmaceuticals. Frankenstein's monster or vampire of trials? Med Care. 1996;34(12 Suppl):DS99–108.

Pekarsky B. The new drug reimbursement game: a regulator's guide to playing and winning. London: Springer; 2015.

Pekarsky B. Trusts, constraints and the counterfactual: reframing the political economy of new drugs. PhD Thesis University of Adelaide, 2012.

# Chapter 2
# Principles and Practice for Trial-Based Health Economic Analysis

## 2.1 Overview

In this chapter, key principles and practice for health economic analysis to undertake robust within-study cost effectiveness analysis are identified and illustrated. Principles are introduced considering the decision analytic basis for comparing alternate strategies in defined patient populations and their costs and effects along treatment pathways. Decision analytic principles for robust cost effectiveness analysis are shown to require joint coverage and compatibility of cost and effect evidence to allow unbiased estimation, the predominant consideration in informing societal decision making under the Arrow-Lind theorem (Arrow and Lind 1970).

These principles are initially applied in this chapter to within-study cost effectiveness analysis for two-strategy comparisons, before being extended to more complex analysis in later chapters. In this simplest two-strategy within-study case, evidence of joint incremental cost and effects can be directly presented from trials on to the incremental cost effectiveness plane. Nevertheless, this only provides unbiased cost effectiveness analysis estimates to inform societal decision making where trial coverage and comparability of relevant incremental effects and costs along alternative treatment pathways are satisfied. Satisfying coverage and comparability conditions to inform unbiased cost effectiveness estimation and decision making more generally requires unbiased methods for evidence synthesis, translation and extrapolation relevant to the context of the jurisdiction to which decisions relate (see Chap. 3). The primary importance under the Arrow-Lind theorem of establishing unbiased cost-effectiveness estimates prior to considering societal decision making under uncertainty in informing joint reimbursement and research decisions (see Chaps. 5, 6 and 7) is nevertheless clarified.

Partialisation problems of the box method when attempting to present cost effectiveness evidence under uncertainty are shown as able to be overcome with non-parametric methods (bootstrapping) and parametric methods (Fieller's method). Joint consideration of cost and effect uncertainty with these methods enables



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_2

within-study cost effectiveness uncertainty to be appropriately considered with bivariate distributions on the incremental cost effectiveness plane where within-study analysis is directly applicable to societal decision making (does not require evidence synthesis, translation or extrapolation as per Chap. 3). Similarly, for two strategy comparisons bivariate distributions can in turn be simply summarised for societal decision making with cost effectiveness acceptability and net benefit curves. They respectively directly inform societal decision makers of the probability of, and incremental expected net benefit from adopting strategies, conditional on societal threshold values for effects. Principles and methods are illustrated with reference to the seminal 'Thinking outside the box' paper (Briggs et al. 2002) and the LIPID study of statin use (Glasziou et al. 2002).

The importance of net benefit as a robust metric to jointly allow for costs and effects in decision making under uncertainty while avoiding ordering problems inherent with incremental cost effectiveness ratios (ICERs) is highlighted, following Willan and Briggs (2006). Incremental net benefit metrics as the value relative to a comparator of incremental effects, less incremental costs also make explicit the need for economically meaningful threshold values for effects (Graham 1981, 1992) conditional on decision context for investment. Methods for determining economically meaningful threshold values that reflect opportunity costs (alternative best actions) conditional on local contexts (health system allocative and displacement inefficiency) are introduced with the health shadow price (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014). However, this health shadow price and economically meaningful threshold values for effects are not fully considered, allowing for all relevant decision contexts, and emprically, until Chap. 11.

Nevertheless, summary measures for two-strategy comparisons of net benefit and cost-effectiveness acceptability curves introduced are shown to appropriately condition across potential threshold values in the absence of knowledge by analysts of the relevant empirical threshold value and related decision contexts in any given jurisdiction. That is, they present the probability of maximising net benefit (CEA curves) and expected incremental net benefit (INB curves) as a function across plausible ranges for threshold values. Similarly, robust summary measures for multiple strategy and effect comparisons (expected net loss curves and frontiers in Chap. 8 and planes and surfaces in Chap. 10) are presented as functions of plausible threshold values for effects in informing related reimbursement and research decisions.

Hence, jointly satisfying coverage and comparability principles and evaluating costs and effects together with net benefit analysis is illustrated not only as key for robust two-strategy within-study comparison but also as a foundation later for robust more complex analysis. Coverage and comparability principles with consistent and joint consideration of cost and effects along alternative pathways are later shown to also be critical building blocks for robust methods of:

(i) Evidence synthesis, translation and extrapolation (O'Brien 1996; Eckermann et al. 2009, 2011) that are usually required to robustly inform societal decision making within any jurisdiction, as highlighted in Chap. 3;

(ii) Joint research and reimbursement decisions when considering cost effectiveness evidence of promising strategies under uncertainty (Chaps. 5, 6 and 7); and
(iii) Cost effectiveness analysis with more than two strategies (Eckermann et al. 2008; Eckermann and Willan 2011) in Chap. 8 and more than two outcomes (McCaffrey 2013; McCaffrey et al. 2015) in Chap. 10.
(iv) Comparisons of providers, strategies and health systems in practice (Eckermann 2004, 2009; Eckermann and Coelli 2013) in Chaps. 10 and 11.

## 2.2 Principles for Robust Health Technology Assessment

In undertaking economic evaluation, public health systems are responding to scarcity of resources in attempting to satisfy health needs across populations over time. Processes of health technology assessment attempt to inform choices between alternative strategies in treating defined patient populations based on comparing their relative costs and value of effects. Trade-offs arise in two strategy comparisons unless one strategy has lower costs and higher effects (dominates the other strategy) or equivalently the other strategy is dominated (has higher costs and lower effects). Where trade-offs arise, assessing value can be viewed as a set of scales (Fig. 2.1) weighing up the value of net incremental effects relative to net incremental costs.

Note, however, that calibration of such scales is required to represent value in trade-offs between incremental costs and effects. Hence, in making a decision about whether to invest in, or reimburse (adopt and finance) a strategy that has higher expected net costs, decision maker threshold values for effects need to reflect opportunity costs of adopting and financing actions to optimise health effects within any constrained budget (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014). That is, threshold values for effects in reimbursement decisions should reflect highest value alternative adoption and financing actions. In Sect. 2.10, we start to consider how threshold values reflecting opportunity cost should

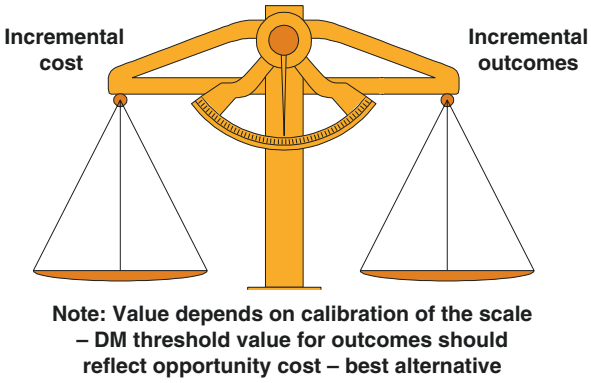


**Fig. 2.1** Cost effectiveness analysis – weighing up value of incremental impacts

be appropriately determined, a critical issue which we later return to in detail in Chap. 12 allowing for relevant decision contexts faced by jurisdictions in their health systems (allocative and displacement inefficiency particularly). Suffice to say from the beginning that one should always be mindful of the opportunity cost – the best alternative action(s) – that such threshold values should reflect to enable resource-constrained optimisation in decision making for any given health system or jurisdiction of interest.

## 2.3 Decision Analytic Approaches to Robust Analysis

A decision analytic approach (Fig. 2.2) provides a systematic and explicit way to estimate incremental effects, resource use and costs of alternative strategies and points to principles for undertaking robust analysis.

Each patient in a target population travelling down care pathways (whether prevention, diagnosis, treatment, rehabilitation or palliative care, etc.) associated with alternative interventions or strategies has a cost and effect associated with that pathway. Principles of comparability and coverage are highlighted in such decision trees. For any given target patient population unbiased estimation of incremental effects, resource use and costs require that their comparable relative impacts are adequately captured along treatment pathways. To support comparability when estimating relative effects, resource use and costs, randomised control trial evidence for compared strategies compared is ideally available to avoid selection biases (both

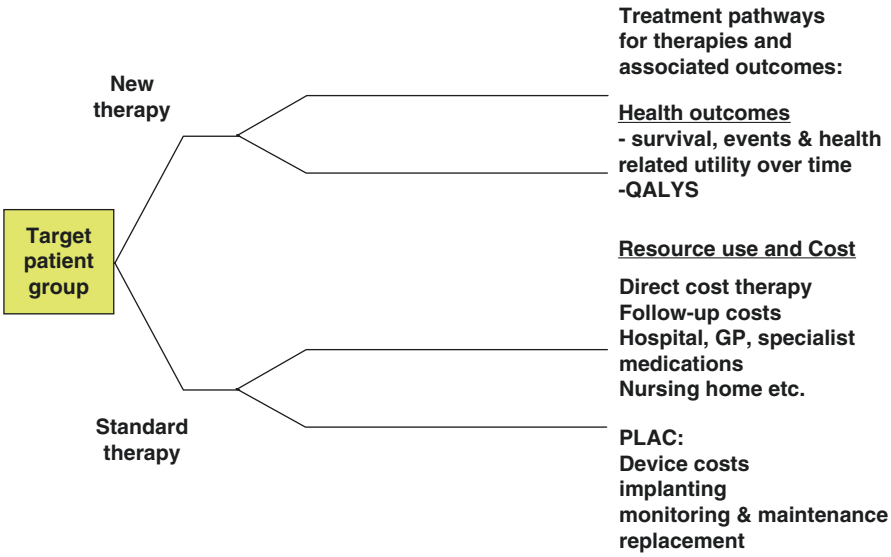


**Fig. 2.2** Decision analytic principles – coverage and comparability in capturing costs and outcomes (Eckermann, 2nd April 2014)

observed and non-observed factors) in allocating patients between arms. However, adequate coverage of the scope and duration of effects and associated costs along treatment pathways is also required to avoid selection biases arising in the coverage of effects, resource use and costs included in incremental cost effectiveness comparison.

Without randomised control trial evidence, the potential arises for systematic biases in relation to non-observed as well as observed factors associated with selection of patients by arm, in estimating net incremental cost and effects. However, not having adequate coverage of the scope and duration of effects and associated costs of treatment also leads to systematic bias in estimating net incremental costs and effects, for example, if the health impacts and cost of treatment associated with side effects are not included or study duration does not capture downstream cost and effect impacts of differences in rates of sequalae. Hence, decision analytic principles underlying health economics highlight the need for adequate and consistent coverage (scope and duration), as well as comparability of evidence, to robustly estimate relative and absolute incremental effects, resource use and costs for defined patient populations across alternate pathways.

For two strategy comparisons robust estimation of incremental costs, effects and their joint consideration, incremental cost effectiveness analysis (or equivalently incremental net benefit analysis as we later see in this chapter) requires:

(i) Unbiased estimation of treatment effects on health affects resources relative to an appropriate comparator (compatibility)
(ii) Sufficient length of follow-up and scope of resource use and health effects to capture incremental costs and effects (coverage)

Joint consideration and satisfaction of these coverage and comparability principles is key to preventing biases in cost-effectiveness analysis.

Importantly, the Arrow-Lind theorem (Arrow and Lind 1970) highlights the primary importance of avoiding biased cost effectiveness estimates before considering cost effectiveness uncertainty for societal decision making to be best informed in processes of health technology assessment. The Arrow-Lind theorem establishes that societal risk preferences asymptote towards risk neutrality with risk spreading across large populations and multiple decisions. Hence, minimising bias should predominate over increasing precision as the primary focus of cost effectiveness analysis in health technology assessment processes. Consequently, repeated decision making across large populations informed by bodies such as The National Institute for Health and Care Excellence (NICE) in the UK and the PBAC in Australia should primarily be interested in avoiding biases in estimating expected incremental cost, effect and their joint consideration, cost-effectiveness.

This is highlighted in Fig. 2.3, where unbiased estimation of incremental effects, costs and INB is the primary foundation to robustly informing optimal decision making cycles locally.

Note that this does not mean that uncertainty is not important. Considering uncertainty of INB is the key consideration in creating appropriate incentives for

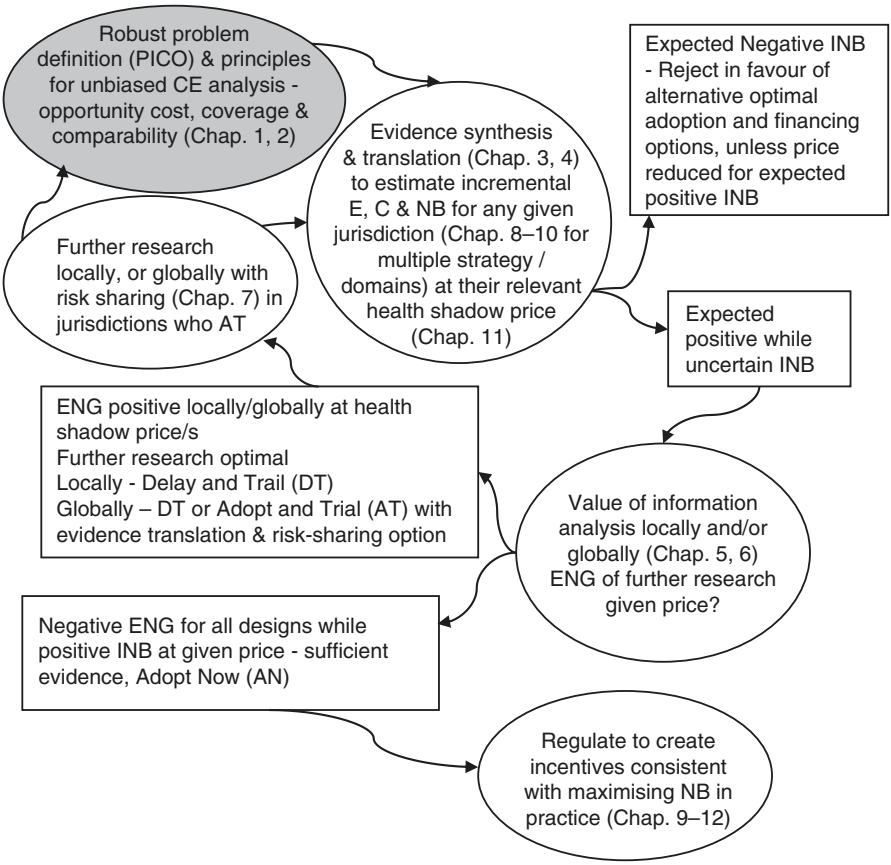


**Fig. 2.3** Optimal decision making cycles for joint research, reimbursement and regulatory processes locally and globally

adequate research (Chaps. 5, 6 and 7) and more generally for optimal joint research, reimbursement and pricing decisions in evaluation, policy and practice (Chaps. 8, 9, 10, 11 and 12). However, meaningful consideration of such uncertainty and associated decision making requires unbiased estimation of incremental costs, effects and net benefit, as considered in this chapter for within-trial evidence and in Chap. 3 when translating trial evidence to jurisdictions of interest.

The alternative, modelling uncertainty with biased methods, is akin to looking with rose-coloured glasses at a light that you primarily need to identify the central colour of, because you might be able to see the edge shapes better.

Biased cost effectiveness analysis cannot be justified for reimbursement decisions given an underlying objective and decision context for HTA informed by the Arrow-Lind theorem. Given many reimbursement decisions made across large populations, the Arrow-Lind theorem makes clear the need for unbiased methods to maximise expected net benefit of such decisions. Further, for research decisions, location of the

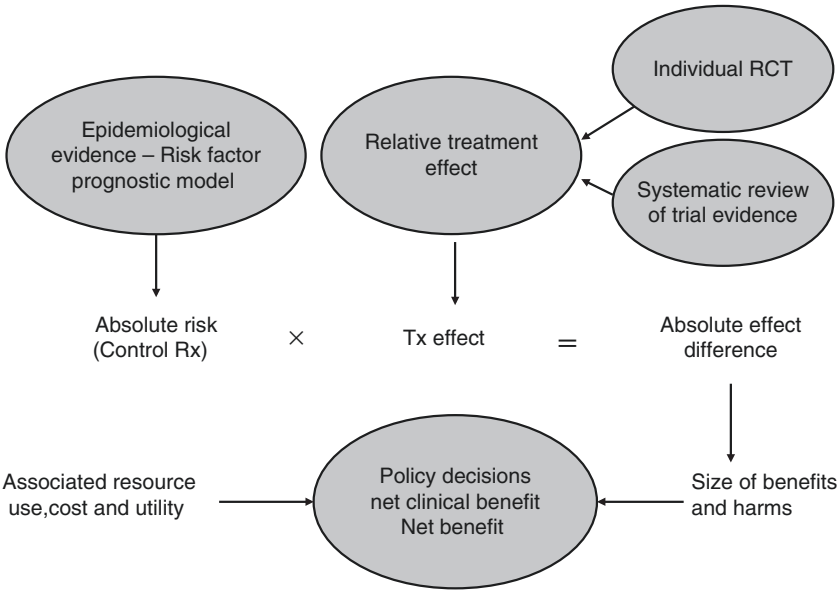


**Fig. 2.4** HTA processes informing decision making in a jurisdiction of interest – e.g. PBAC in Australia

distribution of INB is also fundamental for any jurisdiction(s) to robustly compare the expected value and cost of further research locally (DT vs. AN) and globally (AT vs. AN) allowing for key decision contexts (Chaps. 5, 6 and 7). While rose-coloured glasses might make a shape marginally more discernible at the edges, they end up changing the whole colour (shifting the location of the whole distribution). Satisfying coverage and comparability principles for unbiased cost effectiveness analysis provides the key to robust reimbursement and research decisions and their joint optimisation.

Consequently, the starting points for a building block to consider any such more complex methods are principles and methods for unbiased cost effectiveness analysis. Minimising bias by jointly satisfying comparability and coverage principles for effects and costs along relevant pathways of care is paramount to robust within-trial analysis (this chapter), inform decision making in any jurisdiction of interest (Chap. 3) or any more complex forms of analysis.

Figure 2.4 highlights the decision analytic principles of coverage and compatibility in practice and points towards methods required to enable robust unbiased analysis satisfying these principles (this chapter), but more generally for bodies such as the PBAC in Australia to best inform cost effectiveness decisions for their relevant jurisdiction. In particular, the need to move beyond within-trial-based analysis developed in this chapter to methods and metrics for consistently synthesising and translating trial evidence to inform clinical and health economic policy decisions in any given jurisdiction of interest (Chap. 3).

Policy decisions from a community perspective need to consider net clinical benefit of strategies expected in a given patient population trading off expected harms and benefits. For two-strategy comparisons incremental net clinical benefit is in many settings ideally measured with incremental quality adjusted life years (QALYs) allowing for relative mortality, morbidity and side effect impacts over time. However, note that in areas such as palliative care, multiple additional key domains of effect not able to be integrated with survival such as finalising personal and financial affairs in process of death, family and carer distress and carer burden and preference for place of palliative care and death are primary concerns, as highlighted in Chap. 4 and multiple domain methods in Chap. 10. Incremental net benefit (Graham 1981, 1992; Claxton and Possnet 1996; Stinnett and Mullahy 1998) simply extends assessment of absolute incremental effect or net clinical benefit ($\Delta E$) to additionally allow for impacts on resource use and net incremental cost ($\Delta C$). Incremental net monetary benefit (INMB) considers the value of net incremental effects at a threshold value ($\lambda$) for effect, less net incremental cost:

$$\text{INMB} = \lambda \Delta E - \Delta C.$$

Incremental net benefit (INB) can also be expressed in terms of effects as incremental net effect benefit (INEB): $\text{INEB} = \Delta E - \Delta C/\lambda$.

Nevertheless, for health economics analysis and to avoid issues that arise with INEB where a 0 threshold value for effects is considered, we will stick to INMB in considering INB.

## 2.4 Why Use Incremental Net Benefit and Not Incremental Cost Effectiveness Ratios

During the late 1980s and 1990s, the incremental cost effectiveness ratio (ICER) was proposed and became a popular way of summarizing cost effectiveness evidence to inform health technology assessment. The ICER represents the incremental cost (including direct cost and downstream costs associated with effects) divided by incremental effect of a strategy relative to a comparator.

Formally, for intervention $i$ (e.g. the treatment arm of a trial) and comparator $c$ (e.g. control arm of a trial), an estimate of the ICER for intervention $i$ relative to comparator $c$ can be estimated from evidence for mean costs and effects as

$$\text{ICER}_{i,c} = \frac{\text{Cost}_i - \text{Cost}_c}{\text{Effect}_i - \text{Effect}_c} = \frac{\Delta C_{i,c}}{\Delta E_{i,c}}$$

If the effect per patient were survival, then the ICER becomes incremental mean cost per survivor. If the effect were life years, then the ICER becomes incremental cost per life year. If the effect were QALYs, then the ICER estimate becomes incremental cost per QALY.

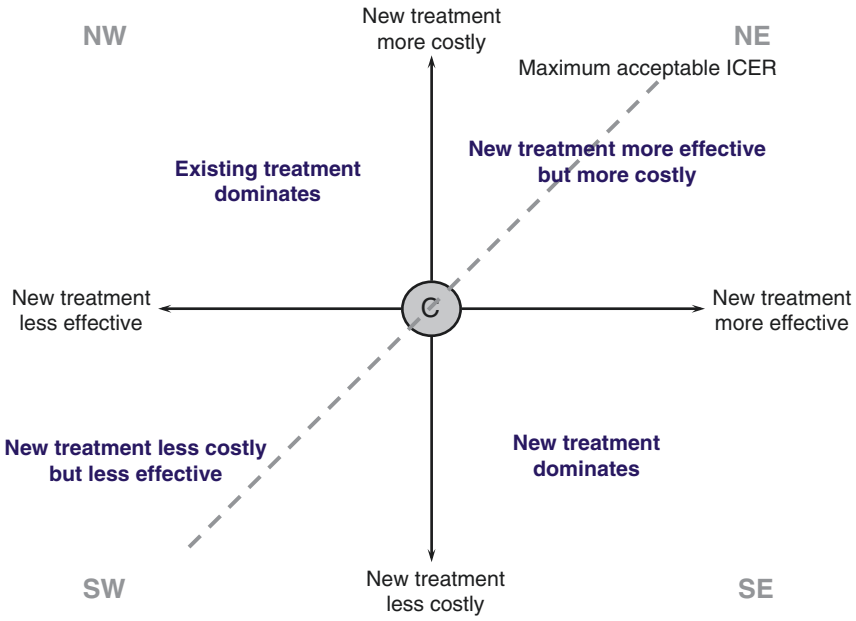


**Fig. 2.5** The incremental cost effectiveness plane

Following Willan and Briggs (2006), the ICER can alternatively be written as

$$\text{ICER}_{i,c} = NNT_{i,c} \times \Delta\text{Cost}_{i,c}$$

noting that

$$NNT_{i,c} = \frac{1}{\Delta E_{i,c}}$$

That is, the number needed to treat (NNT) to gain one unit of effect, an extra survivor, life year or QALY, is the inverse of change in effect per patient. Hence, it naturally follows that the incremental cost per unit effect is the average incremental cost per patient multiplied by the NNT (expected number of patients required to achieve one incremental unit of effect). Incremental costs, effects and the ICER for an intervention or strategy relative to a comparator are also simply and informatively presented on the incremental cost effectiveness plane (Fig. 2.5).

The incremental cost effectiveness plane presents incremental effects and costs of the intervention relative to a fixed comparator at the origin. By convention, incremental effects are presented on the horizontal axis and incremental costs on the vertical axis. These axes divide the incremental cost effectiveness plane into four quadrants which can be described by quadrants as in a compass, as northeast (NE), southeast (SE), southwest (SW) and northwest (NW) quadrants.

If the new therapy has expected positive incremental net clinical effect and lower net cost (allowing for costs associated with effects as well as direct costs of the intervention and comparator strategies) and lies in the SE quadrant ($\Delta E > 0$, $\Delta C < 0$), then the existing strategy is said to dominate the comparator. Conversely, if the new therapy has negative incremental net effect and higher net cost relative to the comparator strategy, and lies in the NW quadrant ($\Delta E < 0$, $\Delta C > 0$), then the existing strategy is said to be dominated by the comparator. Note that in either of these cases there is not a trade-off between incremental cost and effects in distinguishing which intervention is preferred and a threshold value for effects is not required to discriminate what should be the preferred intervention (at least not until uncertainty is considered).

In the NE and SW quadrants, trade-offs between incremental cost and effects arise, and a threshold value for effects is required to distinguish which strategy is preferred. Presenting evidence on the incremental cost effectiveness plane relative to a fixed comparator, the ICER at any point is represented by the slope of a line from the origin. That is, the slope of a line from the origin to any point on the plane represents the ICER or incremental costs divided by incremental effects.

Given the slope of any line through the origin represents the ICER, if one considers the maximum threshold value of the ICER on the NE quadrant for a given jurisdiction at a point in time (and implicitly for given decision contexts, see Chap. 11) as a constant (i.e. not altered by size of budget impacts), then a line from the origin on the NE quadrant with that slope can represent the threshold acceptable ICER. Under this assumption, for two-strategy comparison, a line with slope equivalent to the threshold ICER can distinguish which intervention is preferred in the NE quadrant given evidence of incremental expected costs and effects.

However, note that such analysis is not able to delineate preferred strategies for more than two strategy comparisons, as with multiple strategies there is no longer one fixed comparator (Eckermann et al. 2008; Eckermann and Willan 2011; Eckermann 2004), and requires alternate methods and summary measures as identified in Chap. 8. Further, the direction of budget impacts, additional cost (NE quadrant) or cost reduction (SW quadrant) is also shown to alter the subjective nature of opportunity cost (alternative adoption and financing vs. funding generation) and appropriate threshold values in the SW and NE quadrant (Eckermann 2015), as considered at length in Chap. 11.

Of more obvious and immediate importance, problems arise with ICER metrics when change in effect is 0 or crosses the horizontal axis across 0 effect. When change in effect is 0, the ICER directly or as NNT (inverse of incremental effect) multiplied by incremental cost per patient is undefined. This is the first of a series of problems with the ICER, which in general is not well ordered. As Willan and Briggs (2006) highlight, the ICER has:

(i) A discontinuity when $\Delta E$ changes sign. For example, with positive incremental cost, an ICER changes from approaching infinity when change in effect is small and positive in the NE quadrant to approaching negative infinity when change in effects is small and negative, in crossing to the NW quadrant.
(ii) The same negative sign in the NW and SE quadrants, but diametrically opposite implications with an intervention or strategy dominating the comparator (having higher effect and lower cost) in the SE quadrant while being

dominated by the comparator (having lower effect and higher cost) in the NW quadrant.

(iii) The same value in moving along any given ray from origin, while in the SE and NW quadrants, respectively, representing increasing domination of (SE) and domination by (NW) the comparator strategy.

The ICER as a result of (i) and (ii) requires separate statements and consideration of which strategy is preferred when effects are positive or negative, while (iii) implies that even within such separate statements, ICER ordering makes no sense where the ICER is negative. These ordering problems make the ICER highly problematic as a summary measure of cost effectiveness in interpreting or comparing point estimates, let alone under uncertainty. Additional knowledge of which quadrant incremental cost and effect estimates are in is required to allow any meaningful interpretation for decision making. Further, these ordering problems mean the ICER usually becomes untenable as a summary measure once cost effectiveness uncertainty is considered.

Hence, in general the ICER as a ratio measure does not have good statistical properties, where any evidence lies outside the NE quadrant.

The inherent and largely intractable ordering problems of the ICER as a ratio measure in attempting to inform cost effectiveness decision making are, however, simply circumvented by use of incremental net benefit metrics. Incremental net monetary benefit (INMB) as the value of incremental effects ($\lambda\Delta E$) less incremental costs ($\Delta C$), $\text{INMB} = \lambda\Delta E - \Delta C$, provides a continuous metric that does not face the decision ordering problems of the ICER as a ratio, while representing the same decision rule. That is, INMB being greater than 0 represents the same decision rule as the ICER being acceptable relative to a decision threshold value for effects for two strategy comparisons.

Formally, the cost effectiveness decision rule of

$$\Delta C / \Delta E < \lambda, \quad \Delta E > 0,$$

where $\lambda$ is the threshold value per unit effect, or the less often considered

$$\Delta C / \Delta E > \lambda \ \ for\, \Delta E < 0$$

can both be rewritten as

$$\text{INMB} = \lambda\Delta E - \Delta C > 0$$

As a linear combination of $\Delta E$ and $\Delta C$, INMB is continuous with regard to both and has linear properties in relation to their mean and variance. INMB also does not require separate consideration of whether incremental effect is positive or negative while representing the same decision rule with respect to cost effectiveness. Together these advantages of INMB overcome the statistical and interpretability problems of the ICER.

In contrast to the ICER, the INMB statistic is well ordered within and across quadrants, and its direction and extent reflect appropriate decision making, with INMB:

(i) Continuous when effect changes from being positive to negative or vice versa (the sign of $\Delta E$ changes around 0).
(ii) Unambiguously negative in the NW quadrant where the comparator dominates the new intervention and positive in the SE quadrant where the new treatment dominates.
(iii) Increasingly negative and positive, as appropriate in the NW and SE quadrants respectively, as one moves along a ray away from the origin. That is, INMB reflects increasingly being dominated or dominating in moving away from the origin along a ray in the NW and SE quadrants.

Hence, the direction and extent of gain or loss expected with decision making are reflected in INMB. Further, when we compare to multiple strategies in Chap. 8, INMB unlike the ICER has the property of being additively separable (Stinnett and Paltiel 1997). This implies that with comparison of multiple strategies, INMB ordering across strategies at a given threshold value does not change with choice of comparator, while such ordering can easily change with choice of comparator with the ICER.

## 2.5 Illustrating Principles Within Study: The LIPID Trial Case Study

The LIPID study represents a double blinded placebo-controlled randomised trial comparing pravastatin incremental to standard care undertaken in 9014 Australian patients with prior myocardial infarction (MI) or unstable angina pectoris (UAP). The health economic analysis undertaken on behalf of the LIPID study group (Glasziou et al. 2002; Eckermann and Kirby 2003) was motivated by concern about the long-term cost-effectiveness of statin use in Australia for these populations.

The LIPID trial design (Fig. 2.6) satisfies the key principles of coverage as well as comparability required for unbiased health economic analysis. Comparability is satisfied by the randomised double blinded nature of the placebo-controlled study. Coverage is addressed both in terms of duration of outcomes over the median 6-year follow-up and in terms of scope of outcomes with evaluation of mortality, hospital and medication use across all 9014 patients and sub-studies of ambulatory care use, medication dose and quality of life impacts on utility measures in more than 1100 patients.

LIPID study results are summarized for all-cause mortality by arm (pravastatin vs. placebo) over the trial follow-up, the primary within-study effect in Fig. 2.7.

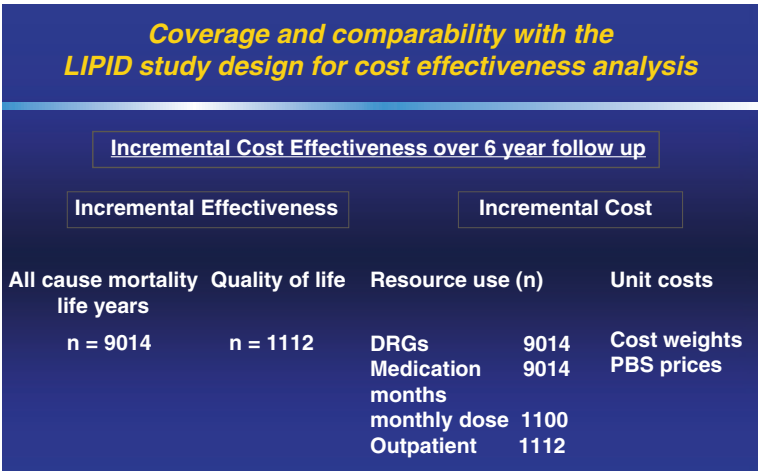


**Fig. 2.6** LIPID cost effectiveness study design

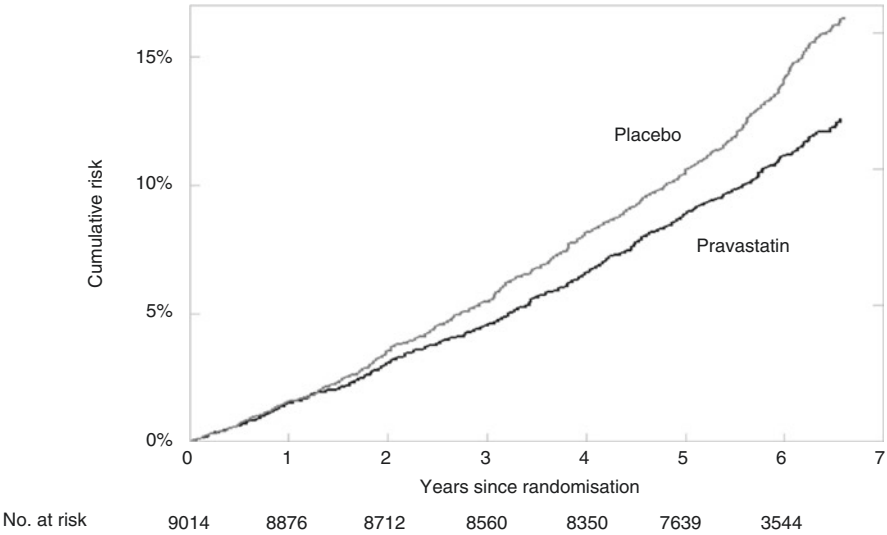


**Fig. 2.7** Lipid study all-cause mortality over study follow-up for pravastatin versus placebo (*Source*: Eckermann and Kirby (2003) on Behalf of the LIPID Study Investigators)

Pravastatin reduced all-cause mortality by 3.01% in absolute terms over a 6-year follow-up, which reflects a 22% relative risk reduction on a baseline risk of 14.1% in the placebo-controlled arm (Table 2.1).

A mean cost of pravastatin of $4913 per patient over a 6-year follow-up was somewhat offset by reduced hospitalisation and other medication costs, leading to an incremental cost of $3246 per patient. Given this mean incremental cost and

**Table 2.1** LIPID within-study incremental cost per life saved

| | |
|---|---|
| Relative risk reduction Mx | 22% (13–31) |
| Baseline (placebo risk) Mx | 14.1% |
| Absolute risk reduction Mx | 3.0% (1.6–4.4) |
| Cost pravastatin per patient | $4913 |
| Reduction in other medication | $360 (272–448) |
| Reduction in hospitalisation | $1385 (804–1966) |
| Incremental cost* | $3246 (2637–3854) |
| ICER ($ per life saved) | $107,730 |

*includes $22 of other cost offsets

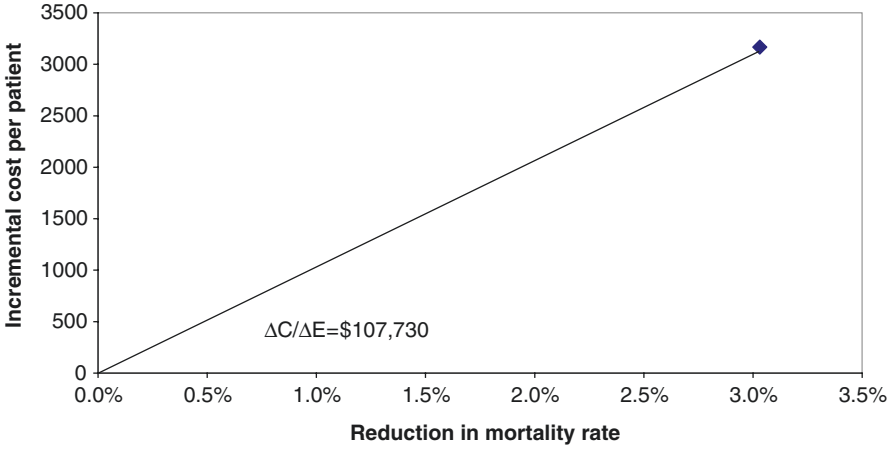


**Fig. 2.8** LIPID evidence on the incremental cost effectiveness plane

reduction in all-cause mortality rate, the incremental cost per additional survivor is estimated as $107,730 ($3246/0.0301) and presented on the incremental cost effectiveness plane as the slope of line from the origin (comparator) to the incremental effect and cost ($\Delta E$, $\Delta C$) point estimate (Fig. 2.8).

This point represents the within-study estimate for incremental costs and effects, and their joint consideration in relation to cost effectiveness is reflected in the ICER estimate, meaningful here noting that it lies on the NE quadrant. The trial population and practice in the LIPID control arm also represented secondary prevention of CHD in Australia at the time of analysis. Hence, for societal decision making in Australia, this also represented the expected incremental costs, effects and their joint consideration in the Australian population at the time analysis was undertaken. More generally, as Chap. 3 highlights, robust estimation of absolute incremental cost and effect requires evidence translation to reflect the baseline risk of the population in practice for the jurisdiction of interest where the decision is being made.

## 2.6 Representing Cost Effectiveness Uncertainty

To allow for uncertainty around incremental cost effectiveness ratios, a box method was initially proposed in health economics literature (O'Brien et al. 1994; Wakker and Klaassen 1995). The 'box method' literally drew a box around the point estimate with the boxes corner points representing the various lower and upper 95% confidence interval (CI) for costs and effects (see Fig. 2.9 for the case of LIPID).

The box method proposed that the 95% confidence interval around the point estimate for the ICER, for example, $107,730 per additional survivor in the case of the LIPID study, could be estimated from the ICER (slope) of lines from the origin to corners of the box representing:

(i) The lower 95%CI for costs and upper 95% CI for effects
(ii) The upper 95% CI for costs and lower 95% CI for effects

Hence, for the LIPID study, the box method would estimate the lower and upper 95% CIs around the point estimate of $107, 730 for the ICER as ranging from about $60,000 per life saved ($2637/0.0439) up to $235,000 per life saved ($3854/0.0164).

In their seminal paper 'Thinking outside the box', Briggs et al. (2002) show distinct problems arising with the box method approach in estimating such uncertainty around the ICER. They note the box method implicitly assumes that the upper and lower CI for cost and effects will occur together and contain 95% of the joint cost and effect distribution. In doing so, the box method fails to allow for the bivariate nature of the relationship (covariance) between incremental costs and effects in

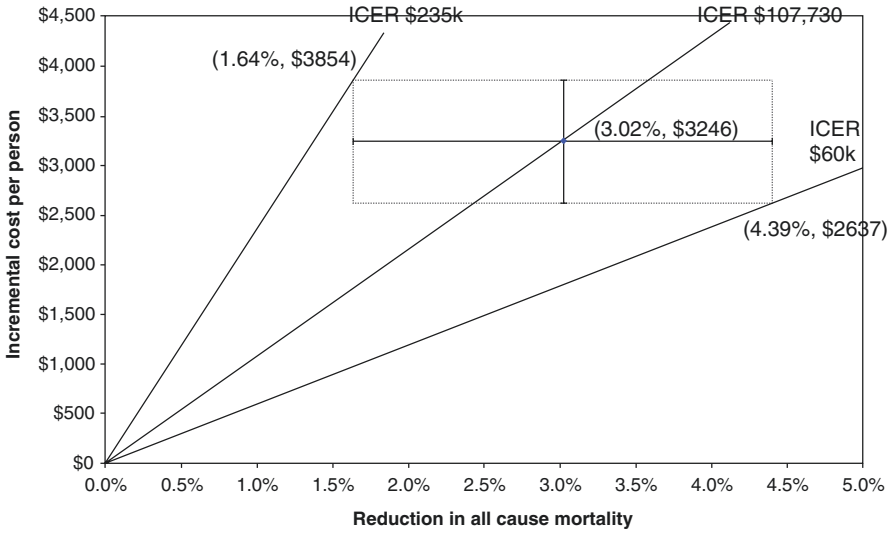


**Fig. 2.9** The 'box method' with LIPID evidence (*Source*: Adapted from Eckermann and Kirby 2003)

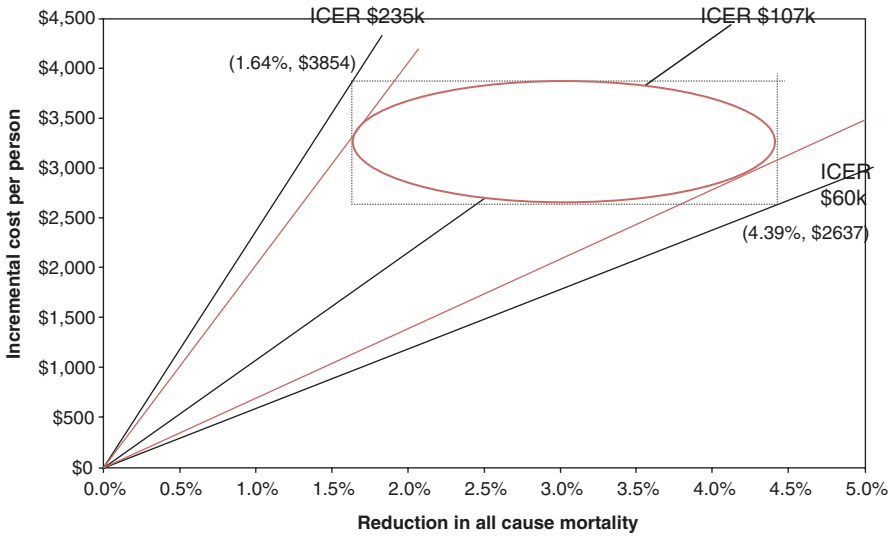


**Fig. 2.10** Box method versus bivariate distribution with no covariance

estimating their joint distribution. Hence, the box method effectively treats separately, or partialises, costs and effects and their distributions. This fails to appropriately reflect the joint nature of how costs and effects arise along treatment pathways and hence the joint distribution of incremental cost and effects under uncertainty.

In reality even if there were no covariance between incremental cost and effects, the box methods' extreme 95% CI highest cost and lowest effect and lowest cost and highest effect points would not be expected to arise together or the box shape around this includes 95% of the distribution. As Briggs et al. (2002) show if there were no covariance between incremental cost and effects, then a distribution radially radiating out from the point estimate is expected. Hence, if there were no covariance between incremental cost and effects (covariance = 0), then the joint distribution of costs and effects would result in a radial shape with narrower band for ICER 95% CI than the box methods in Fig. 2.9 suggest. Rather it would reflect a narrower radial distribution such as that in Fig. 2.10.

However, this does not imply the box method is necessarily conservative, as more generally the joint distribution of costs and effects is elliptical with the orientation and shape of the joint distribution determined by the sign and extent of covariance between incremental cost and effect.

Hence, while ICER uncertainty with the box method will be overestimated if there is no or a positive covariance between incremental costs and incremental effects, ICER uncertainty can be easily underestimated where there is a negative relationship (covariance) between incremental costs and effects. A negative relationship between incremental cost and effects (framed from a utility-bearing perspective on the CE plane, e.g. survival) causes radial joint distributions on the incremental CE plane, such as that in Fig. 2.10, to elliptically flatten out and orientate with a

SE direction. Negative covariance relationships consequently increase cost effectiveness (INB or ICER) uncertainty, widening 95% CIs for the ICER (or INB). Negative covariance between incremental cost and effects in practice typically reflects where incremental effects mainly relate to morbidity, given reducing morbidity (increasing effect) reduces downstream treatment costs while conversely increasing morbidity (reducing effect) increases downstream treatment costs.

However, the box method can significantly overestimate cost effectiveness uncertainty (NB or ICER 95% CIs) if there is a positive relationship between incremental costs and effects. For example, where net effects mainly relate to survival, given increased survival is expected to increase incremental downstream treatment costs of survivors, or equivalently reducing survival is expected to reduce downstream treatment costs of survivors. Such positive relationships between incremental costs and effects cause the distribution in Fig. 2.10 to flatten out and orientate with an NE-positive slope, narrowing cost effectiveness uncertainty from that with no covariance.

In summary, problems of the box method in estimating 95% CI for cost effectiveness arise in inappropriately combining partially determined separate cost and effect inference in attempting to inform cost effectiveness inference. Consequently, the box method does not appropriately allow for the linked relationship (covariance) between costs and effects along treatment pathways and the impact this has on the joint cost and effect distribution on the CE plane or cost effectiveness uncertainty.

Importantly, Briggs et al. (2002) in addressing problems of the box method identify and illustrate how these partial problems can be overcome with methods that jointly consider costs and effects – think outside the box. That is, with robust estimation methods allowing for the joint relationship and covariance of the bivariate distribution between costs and effects, either non-parametrically with bootstrapping or parametrically using Feiller's method.

Both bootstrapping and Fieller's methods enable incremental costs and effects and their joint distribution to be jointly considered allowing for their joint relationship along alternate treatment pathways (covariance). We first consider non-parametric bootstrapping and then turn our attention to Fieller's method.

## 2.7 Bootstrapping the CE Distribution

Bootstrapping is simply repeated resampling with replacement, a non-parametric method which can be used to build up a sampling distribution for joint incremental costs and effects and uncertainty around point estimates for related cost effectiveness summary measures (Briggs et al. 2002). In the case of a trial with Nt patients in the treatment arm and Nc patients in the control arm bootstrapping, the bivariate CE distribution can be summarized as a four-stage process where joint cost and effect patient level data are:

(i) Randomly resampled with replacement for Nc patients and their associated cost and effects from the control group: calculate mean cost and effects for this control group resample.

(ii) Randomly resampled with replacement for Nt patients and their associated cost and effects from the treatment group: calculate mean costs and effects in the treatment group.
(iii) Form a replicate from (i) and (ii) where calculate mean incremental effects and cost for treatment relative to control ($\Delta E$, $\Delta C$).
(iv) Repeat many times (1000 or more) to build up a bootstrapped sampling distribution around the point estimate.

In undertaking these four steps if the seed for random number generation is recorded this allow such resampled bootstrapping of the ICER distribution to be repeatable in various software packages. Importantly, whatever package is used, there should be an equal chance of resampling any individual in any draw when bootstrapping patients with random resampling with replacement. In practice if there are Nc patients (e.g. 200) in the control arm, then a random number between 0 and 1 generated by Rand(), for example, would require random patient assignment using formulae of the general form

round(rand() × Nc + 0.5); Nc + 0.5 = nc.

That is, if there were 200 patients: round(rand() × 200 + 0.5); 200.5 = 200.

This allows an equal chance for each patient to be resampled with any random number, choosing patient 1 for random values from 0 up to 1/200 (0.005), patient 2 from 1/200 (0.005) up to 2/200 (0.01), etc., and patient 200 with values from 199/200 (0.995) up to 1.

A bootstrapped sampling distribution around the point estimate for incremental cost and survival in the LIPID study is shown in Fig. 2.11 for 1000 replicates.

When bootstrapping the bivariate CE distribution, covariance between cost and effects is implicitly maintained as resampling patients retains the relationship between costs and effects for each patient. For two-strategy comparisons considered in this chapter where the comparator is fixed, bootstrapping such distributions allows simple unbiased estimation of the bivariate distribution and summary measures such as the probability of being cost effective (having positive net benefit). The probability of being cost effective at any given threshold value can be simply calculated as the proportion of the distribution with positive INB or equivalently in the acceptance region below (south east of) a threshold line through the origin whose slope reflects the threshold value.

For example, in the case of LIPID 2.5% (25/1000) of the distribution lies at or below $68,626 per life saved and 97.5% at or below $209,881 per life saved (or equivalently 2.5% above). Hence, a 95% CI for the ICER distribution is estimated from the bootstrapped distribution as between $68,626 per life saved and $209,881 per life saved.

Note, however, that while bootstrapping is simple to understand and useful for within-trial and illustrative purposes in establishing the need to jointly consider cost and effects, it does face at least one potentially significant drawback. The method is not exact, with estimates varying depending upon resamples in building up a sampling distribution and subsequent estimating uncertainty. This has led to bodies such as the PBAC being suspicious of such methods when applied and presented by

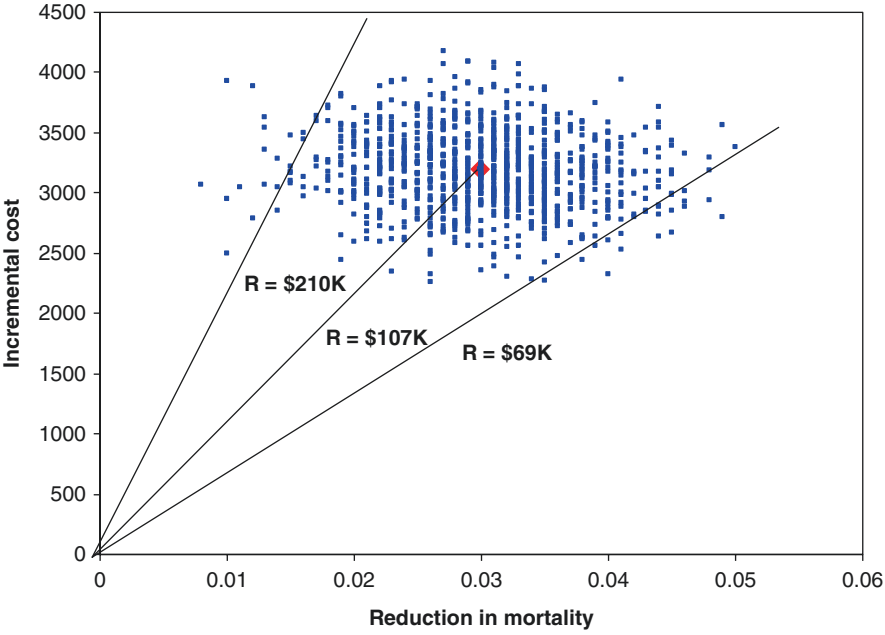


**Fig. 2.11** LIPID – bootstrapped distributions in bivariate space and 95% CI for cost per life saved (*Source*: Adapted from Eckermann and Kirby 2003)

groups with vested interests such as manufacturers. This is particularly the case where bootstrap estimates are presented as a black box without associated replicates or the proportion of times each individual is chosen across replicates.

However, there is a parametric method, Fieller's method, which addresses this concern, providing an exact closed from solution, under the central limit theorem (CLT) assumption of normality, for INMB, as a bivariate distribution.

## 2.8 Fieller's Method

Fieller's method fits a bivariate normal distribution to INMB from summary measures for mean incremental cost and effect, their variances and covariance. That is, Fieller's method uses the fact that INMB is linear in $\Delta E$, $\Delta C$ and $\lambda$ and hence the mean and variance of INMB depends only on the mean and variance of $\Delta E$, $\Delta C$ and their covariance. Formally

$$\text{INMB} = \lambda \Delta E - \Delta C$$

has a variance of

$$\lambda^2 \operatorname{var}(\Delta E) + \operatorname{var}(\Delta C) - 2\lambda \operatorname{cov}(\Delta E, \Delta C)$$

Hence, dividing the INB statistic at the threshold value through by its standard error (square root of the variance) results in a standard normal distribution. This can then be used to find the upper and lower 95% CI, but also the bivariate distribution more generally, dependant only on incremental costs and effects, their variance and covariance, following Willan and Briggs (2006). Further, Nixon et al. (2010) show that such CLT parametric methods outperform bootstrapping with small samples, as the asymptotic properties of the CLT kick in at smaller trial sample sizes than with bootstrapping.

In the case of LIPID, Fieller's method is simply applied with the cost and effect estimates ($3246 and 0.03013 increased survival over the 6-year median study follow-up) and their variance ($100,651 and 0.0000487) and covariance −0.209. This results in a 95% CI for cost per life saved from $68,732 to $204,889. However, the 95% CI for the ICER is a very crude summary of uncertainty of the CE distribution. In simply representing two extreme points on the CE distribution, the ICER 95% CI:

(i) Fails to capture implications across potential sets of decision maker threshold values;
(ii) Is highly reductionist in picking two arbitrary points, the 2.5% and 97.5% point on the ICER distribution; and
(iii) Lacks interpretability where either of these extreme points on the ICER distribution lies outside the NE quadrant.

Summary measures which inform societal decision makers of the whole distribution and across the range of potential decision making threshold values are more useful than throwing away evidence from all but two arbitrarily picked extreme points on the ICER distribution in informing cost effectiveness related decisions.

## 2.9 Useful Cost Effectiveness Summary Measures from Bivariate Distributions Conditioning on Threshold Values for Effect

Conditioning on threshold values per unit of effect, more useful and interpretable summary measures can be found across the full distribution at any threshold value. For two-strategy comparisons where there is only one comparator and one distribution to summarise on the CE plane, useful C-E summary measures informing decision makers across plausible threshold values include:

(i) The cost effectiveness acceptability curve – the probability that a treatment is cost effective (has highest net benefit) across plausible threshold values for a unit of effect; and
(ii) The incremental net benefit (INB) curve and 95% CI curves – the incremental net benefit expected and 95% CI for INMB across plausible threshold values for a unit of effect.

Figure 2.12 shows the LIPID CEA curve for the probability of pravastatin being cost effective in Australia at threshold values from A$0 to A$260,000 per life saved.

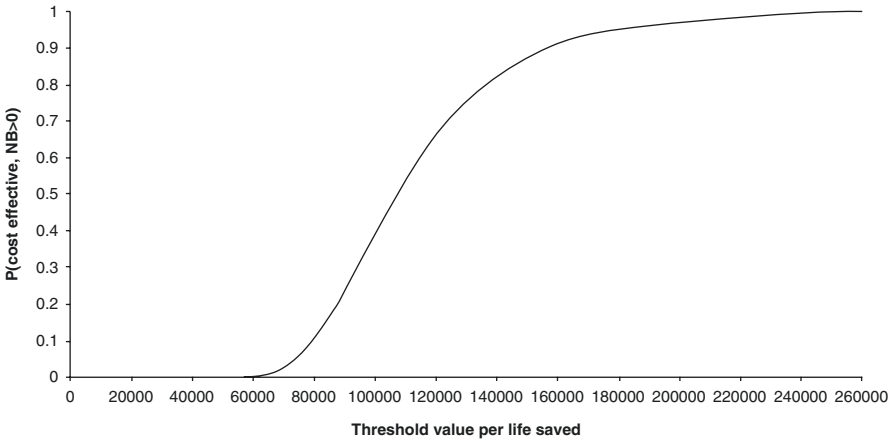


**Fig. 2.12** LIPID cost effectiveness acceptability curve ($/life saved) – Fieller's method

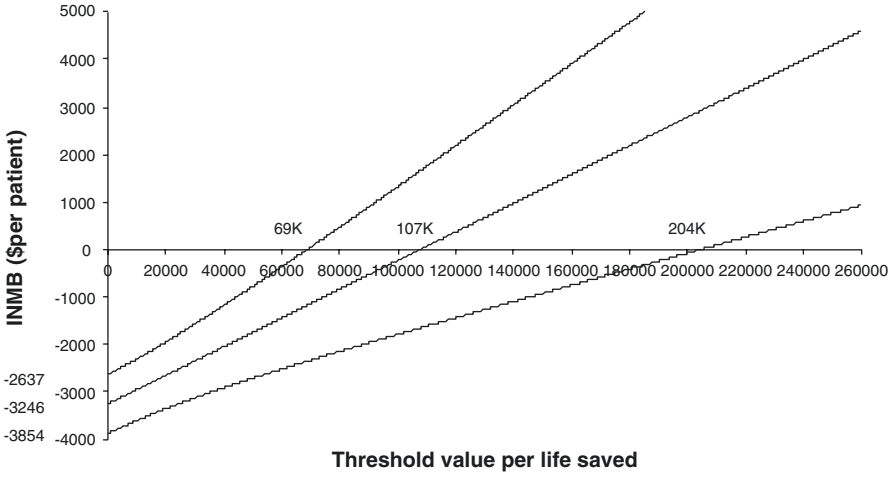


**Fig. 2.13** LIPID net benefit curve with 95% CI – Fieller's method

Similarly, an expected incremental net benefit curve conditional on potential threshold values and curves representing 95% confidence intervals around this expected NB line can also be presented. Figure 2.13 shows such INB curves for pravastatin relative to placebo from the LIPID study conditional on the same range of potential threshold values per life saved as Fig. 2.12.

Where threshold values for effect are 0, INMB (INMB = $\lambda\Delta E - \Delta C$) simplifies to $-\Delta C$. Hence, the expected value and 95% CI for INMB on the vertical axis in Fig. 2.13 with a 0 threshold effect value is simply negative incremental cost. In the case of LIPID, the point estimate for INMB at a 0 threshold value as shown in Fig. 2.13 is −$3246, with 95% CI from −$2637 to −$3854. More generally, INMB depends on the threshold

value for incremental effects as well as incremental cost, with expected INMB changing with $\lambda$ at a rate of $\Delta E$ per unit of the threshold value. That is, the slope of the INB line as a function of $\lambda$ is $\Delta E$. In the case of LIPID, the expected INB line has a slope of 0.03013 (reflecting the absolute 3.013% mortality reduction), increasing at a rate of \$30.13 (=0.03013 × 1000) for every \$1000 increase in the value of a life saved. On the horizontal axis, expected INMB is 0 at the threshold value where $\lambda\Delta E - \Delta C = 0$, and hence at a threshold value where $\lambda = \Delta E/\Delta C$, the ICER estimate. In the case of LIPID, the expected INB curve crosses the horizontal axis at \$107,730 per life saved.

In general, the expected incremental net monetary benefit line $\text{INMB} = \lambda\Delta E - \Delta C$ passes through points of intersection on the vertical axis at $\text{INMB} = -\Delta C$ and horizontal axis ($\lambda$ value with INMB = 0) at the expected ICER ($\lambda = \Delta C/\Delta E$ when INB = 0) and have slope $\Delta E$. Hence, INMB lines will be upward sloping as a function of $\lambda$ where there is a positive expected treatment effect and downward sloping where there is a negative treatment effect. Expected INMB lines start with an implicit threshold value for valued of 0 on the vertical axis. Hence, INMB on the vertical axis simplifies to minus incremental cost, with negative INMB where the strategy has net additional costs, while starting with positive INMB if the strategy is cost saving. This in general leads to four types of expected INMB line (Willian and Briggs 2006) reflecting different potential combinations of cost and effects on the four quadrants of the CE plane:

(i) INMB lines which start negative and become positive corresponding to positive incremental cost and effects (NE quadrant on CE plane).
(ii) INMB lines which start positive and increase corresponding to negative incremental cost and positive effect (SE quadrant on CE plane), and indicate a new therapy dominates existing care.
(iii) INMB lines which start positive and become negative corresponding to negative incremental cost and negative effects (SW quadrant on CE plane).
(iv) INMB lines which start negative and decrease corresponding to positive cost and negative effect (NW quadrant on CE plane), and indicate the comparator (e.g. existing care) dominates then new therapy.

Similarly, the lower and upper 95% CI curves for INMB start on the vertical axis at minus the 95% CIs for incremental cost. The lower and upper 95% CI curves for INMB will cross the horizontal axis at the lower and upper 95% CI for the ICER unless they don't arise – the new strategy dominates or is dominated at these points. Where new strategies are expected to dominate, expected INMB is positive for all feasible positive threshold values for a positive effect, while where new strategies are expected to be dominated, expected INMB is negative for all plausible threshold values. These curves do not cross the horizontal axis (have INMB=0) over feasible ranges for threshold values.

## 2.10 How Should Economically Meaningful Threshold Values for Effects Be Estimated?

Historically, processes of health technology assessment in developed countries such as the PBAC in Australia or NICE in the UK have focused on the NE quadrant with requirements for new technology to demonstrate incremental effectiveness to justify

a price premium over comparators, which profit-motivated manufacturers' try to maximise. From a societal decision maker perspective on the NE quadrant if the estimated combination of incremental costs and effects lies to the SE of (below) the threshold line, then the value of incremental effects is greater than incremental costs, and the new strategy should be preferred. Conversely, if the estimated combination of incremental costs and effects on the NE quadrant lies above (to the NW of) such a threshold line, then the value of incremental effects is less than incremental cost of the new intervention and the comparator strategy should be preferred.

However, one should note that in presentations such as Fig. 2.5, the threshold line and its use in distinguishing what should be preferred assumes a decision-making threshold value which appropriately reflects opportunity cost of reimbursing the new technology. In general, appropriately determined threshold values should reflect opportunity costs in a jurisdiction of interest under local health system conditions, to enable optimisation given investment options and budget (resource) constraints. If there was a pool of new money available to spend on health care then the opportunity cost of spending that new budget on adopting a new technology is the best alternative adoption action, the most cost effective expansion of exisiting programs and technology. However, more generally budgets are fixed and hence reimbursing a new technology requires both adoption and financing actions (Pekarsky 2012, 2015). Given an underlying objective of maximising net benefit from a fixed budget, the decision maker threshold values per unit effect should reflect best alternative adoption and financing actions in relation to this goal (Eckermann and Pekarsky 2014). In reimbursing new technology with net incremental cost, the best alternative actions are most cost effective alternative expansion of current programs or technology funded by contraction of the least cost effective current program or technology, reflected in the health shadow price of Pekarsky (2012, 2015).

The health shadow price is derived by Pekarsky as

$$\beta_c = \left( \frac{1}{n} + \frac{1}{d} - \frac{1}{m} \right)^{-1}$$

where $n$ is the ICER of the most cost effective expansion of current programs, $m$ is the ICER of the least cost effective current program in contraction, and $d$ is the ICER of services displaced.

This derivation arises from finding the threshold cost per unit effect (threshold ICER = $\beta_c$) for a new strategy or technology with net cost of investment (I) at which returns from adopting the new strategy or technology financed with displacement of services (ICER = $d$) given a fixed budget equate with the opportunity cost, that of best alternative actions. The best alternative actions are the most cost effective expansion of current programs (ICER = $n$) and technology financed by contraction of the least cost effective current programs or technologies (ICER = $m$). Hence, $\beta_c$ is solved from equating investment returns as

$$\frac{I}{\beta_c} - \frac{I}{d} = \frac{I}{n} - \frac{I}{m}$$

which dividing through by I and rearranging simplifies to

$$\frac{1}{\beta_c} = \frac{1}{n} + \frac{1}{d} - \frac{1}{m}$$

and hence

$$\beta_c = \left(\frac{1}{n} + \frac{1}{d} - \frac{1}{m}\right)^{-1}$$

This represents the true opportunity cost (Pekarsky 2012, 2015) of reimbursing new strategies or technologies where they have a net incremental cost.

Importantly the health shadow price in comparing with best alternative adoption and financing actions encourages optimal displacement as well as optimal investment actions (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014). If displacement is optimal (the least cost effective current program or technology is displaced, $d = m$), then the health shadow price equates to the most cost effective expansion of current programs and technology, with ICER $n$. That is, if $d = m$, then

$$\beta_c = \left(\frac{1}{n} + \frac{1}{d} - \frac{1}{m}\right)^{-1} = \left(\frac{1}{n} + \frac{1}{m} - \frac{1}{m}\right)^{-1} = n$$

However, where displacement is not optimal ($d<m$), then

$$\frac{1}{d} > \frac{1}{m}$$

$$\frac{1}{n} + \frac{1}{d} - \frac{1}{m} > \frac{1}{n}$$

$$\left(\frac{1}{n} + \frac{1}{d} - \frac{1}{m}\right)^{-1} < n$$

$$\beta_c = \left(\frac{1}{n} + \frac{1}{d} - \frac{1}{m}\right)^{-1} < n$$

These results are key to appropriately interpreting the appropriate threshold value for effects in net benefit – that which reflects opportunity cost and allows budget-constrained optimisation. As Chap. 11 highlights, this should be the case whether net benefit for decision making relates to new technology reimbursement, research decisions or regulatory and policy making assessment in practice. Until Chap. 11, as with other health economic practitioners, we will condition analysis across potential decision making threshold values for effects in analysing and summarising cost effectiveness evidence.

Summary measures conditional on plausible threshold values for effects include:

(i) Cost effectiveness acceptability and net benefit curves for two strategies comparisons with one effect introduced in this chapter;
(ii) Expected net gain in optimising the expected value relative to cost of research designs and associated joint research and reimbursement decision locally (Chap. 5) and globally (Chaps. 6 and 7);
(iii) Expected net loss curves and frontiers for multiple strategies (Chap. 8);
(iv) Net benefit efficiency measures (Chap. 9); and
(v) Expected net loss planes and contours for multiple outcomes (Chap. 10).

The implications of the health shadow price in expansion for net benefit maximisation and budget-constrained optimisation are considered at length in Chap. 11. The health shadow price in contraction is also considered following Eckermann (2015), with empirical estimates of the health shadow price of expansion and contraction in the UK presented based on program budgeting marginal analysis (PBMA) evidence. Alternative threshold values for health which have previously been proposed for comparing and pricing new technology with higher net costs against – willingness to pay or the ICER of displaced programs (d) are also critiqued in Chap. 11 following Eckermann and Pekarsky (2014). Unlike the health shadow price, these alternatives are shown to not reflect opportunity cost of best alternative actions or allow a pathway to budget-constrained optimisation from current allocative ($n<m$) or displacement ($d<m$) inefficiency.

Critically, the health shadow price (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014) points to the need for research in relation to expansion and contraction of current technology and programs, that is, research to identify the most cost effective expansion of current technology and programs and where to contract or displace the least cost effective current programs and technology. Hence, the health shadow price is shown in Chap. 11 to establish the threshold values societal decision makers should be using in creating a pathway to optimisation across research, reimbursement and regulatory decisions (pricing and provider performance in practice).

## 2.11 Conclusion

Satisfying decision analytic principles of coverage, comparability and consistency are keys to obtaining unbiased estimates for relative comparison of absolute costs and effects and cost effectiveness analysis.

These represent the primary considerations to best inform decision making of bodies such as NICE and the PBAC about cost effectiveness in process of health technology assessment. To meaningfully model cost effectiveness uncertainty or summary measures such as CEA and NB curves requires costs, and effect uncertainty is also jointly considered (Briggs et al. 2002). However, first and foremost, these distributions need to be based around unbiased estimates of absolute incremental costs and effects, where coverage and comparability principles are

jointly satisfied. Satisfying comparability, coverage and consistency principles in estimating joint costs and effects along alternative pathways prior to consideration of decision uncertainty lays the foundation stone for robust, unbiased cost effectiveness analysis. These principles are central to allowing robust analysis for within-study RCT evidence comparing two strategies illustrated in this chapter, but also any more complex forms of analysis. Coverage, comparability and consistency principles are also central to robust analysis throughout the text in:

(i) Synthesizing, translating and extrapolating evidence (Chap. 3);
(ii) Evaluating health promotion and prevention strategies (Chap. 4);
(iii) Informing and optimizing joint research and reimbursement decisions locally and globally (Chaps. 5, 6 and 7);
(iv) Multiple strategy and multiple outcome comparisons (Chaps. 8 and 10);
(v) Evaluating efficiency in performance of health care providers, and health funding systems, in practice allowing for quality of care consistent with maximizing net benefit (Chap. 9 and Sect. 12.5), where the net benefit correspondence theorem underlying these methods makes explicit the need to satisfy coverage and comparability conditions to enable robust analysis and create appropriate incentives in practice (Eckermann 2004; Eckermann and Coelli 2013);
(vi) Establishing economically meaningful opportunity costs and threshold values for effects in jurisdictions of interest given relevant decision contexts for health system allocative and displacement inefficiency (Chap. 11), following Pekarsky (2012, 2015) and Eckermann and Pekarsky (2014); and
(vii) Policy analysis (Chap. 12).

## 2.12 Discussion – Satisfying Coverage, the Need for Robust Evidence Synthesis, Translation and Extrapolation

The LIPID study satisfies comparability and coverage principles in providing RCT evidence with adequate scope (mortality and quality of life) and duration of coverage (6-year median follow-up) for a trial-based analysis. This also doubled as an Australian analysis given study patients and their treatment were representative of practice at time of decision making. The inclusiveness of patient in the LIPID study supports trial analysis providing a robust estimate of baseline risk expected in practice in secondary prevention populations in Australian decisions related to adopting statin therapy, as well as relative treatment effect. However, synthesis, translation and potentially extrapolation of trial evidence are more generally needed to allow valid estimation in a jurisdiction of absolute incremental effects, cost and INB. Differences in population INB in practice in a jurisdiction of interest can differ from that in a trial where trial inclusion and exclusion criteria as well as geography and associated populations, practice, prices and preferences differ.

In this chapter, we have considered the simplest case of within-trial two-strategy comparisons, as often reported in CE literature alongside trials. Such estimates can provide meaningful analysis of expected effects, costs and cost effectiveness for decision making in the jurisdiction where the trial is undertaken, provided the comparator arm reflects usual practice and the trial population is the same as that expected in practice.

Robust, unbiased methods for trial evidence translation to reflect the baseline risk expected in practice in any given jurisdiction of interest are established in Chap. 3 (Eckermann et al. 2011).

Chapter 3 more generally makes clear that avoiding bias requires consistent methods for evidence synthesis, translation and extrapolation as well as coverage of the scope and duration of incremental cost and effects. In general, both coverage and comparability need to be satisfied to enable unbiased estimates of INB for any given jurisdiction, a precursor to any meaningful consideration of cost effectiveness (INB) decision uncertainty. Hence, principles of coverage and comparability form the basis for robust cost effectiveness analysis whether analysis is purely based on a RCT or is undertaken with model-based analysis. Bernie O'Brien's seminal paper 'Frankenstein's Monster or the Vampire of Trials' (O'Brien 1996) takes centre stage in Chap. 3 establishing the need to jointly satisfy coverage and comparability principles with model and trial-based analysis.

## References

Arrow KG, Lind RC. Uncertainty and the evaluation of public investment decision. Am Econ Rev. 1970;60:364–78.

Briggs AH, O'Brien BJ, Blackhouse G. Thinking outside the box: recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. Annu Rev Public Health. 2002;23(1):377–401.

Claxton K, Posnett J. An economic approach to trial design and research priority-setting. Health Econ. 1996;5(6):513–24.

Eckermann S. Hospital performance including quality: creating incentives consistent with evidence-based medicine PhD Dissertation. Sydney: UNSW; 2004. http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/.

Eckermann S. Measuring health system efficiency and funding for net benefit maximisation: the health economics of quality of care. SA Health Department working paper. 2009. ISBN 13 978 1 921402 050. Accessed 29 May 2011. http://clinicalchange.flinders.edu.au/files/workingpapers/wp_2009_08.pdf.

Eckermann S. Kinky thresholds revisited: opportunity costs differ in the NE and SW quadrants. App Health Econ Health Policy. 2015;13:7–13. doi:10.1007/s40258-014-0136-3.

Eckermann S, Briggs A, Willan A. Health technology assessment in the cost-disutility plane. Med Decis Making. 2008;28:172–81.

Eckermann S, Coelli T. Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice. Soc Sci Med. 2013;76:159–68. http://dx.doi.org/10.1016/j.socscimed.2012.10.020

Eckermann S, Coory M, Willan A. Consistently estimating absolute risk difference when translating evidence to jurisdictions of interest. Pharmacoeconomics. 2011;29(2):87–96.

Eckermann S, Coory M, Willan AR. Indirect comparison: relative risk fallacies and odds solution. J Clin Epidemiol. 2009;62:1031–6.

Eckermann S and Kirby A and on Behalf of the LIPID Study Investigators. Cost effectiveness analysis: uncertainty, predictive conditioning and extrapolation post study – results from LIPID. Economics and health: 2002 Proceedings of the Twenty Fourth Australian Conference of Health Economists. In: Butler JRG, Quinn C, editors. Sydney: AHES; 2003. p. 54–83.

Eckermann S, Pekarsky B. Can the real opportunity cost stand up: displaced services, the straw man outside the room. PharmacoEconomics. 2014;32(4):319–25.

Eckermann S, Willan A. Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies. Pharmacoeconomics. 2011;29(7):563–77.

Glasziou P, Eckermann S, Mulray S, Simes RJ, Martin A, Kirby A and on Behalf of the LIPID Study Investigators. Cholesterol-lowering therapy with pravastatin in patients with average cholesterol levels and established ischaemic heart disease: is it cost effective? Med J Aust. 2002;177:428–34.

Graham DA. Cost-benefit analysis under uncertainty. Am Econ Rev. 1981;71:715–25.

Graham D. Public expenditure under uncertainty: the net-benefit criteria. Am Econ Rev. 1992;82:822–46.

McCaffrey N. Modelling joint cost and outcomes uncertainty on the cost-disutility plane – case studies in palliative care. PhD, Flinders University; 2013.

McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Better informing decision making with multiple outcomes cost-effectiveness analysis under uncertainty in cost-disutility space. PLoS ONE. 2015;10(3):e0115544. Accepted 8th January 2015.

Nixon RM, Wonderling D, Grieve RD. Non-parametric methods for cost-effectiveness analysis: the central limit theorem and the bootstrap compared. Health Econ. 2010;19(3):316–33.

O'Brien B. Economic evaluation of pharmaceuticals. Frankenstein's monster or vampire of trials? Med Care. 1996;34(12 Suppl):DS99–108.

O'Brien BJ, Drummond MF, Labelle RJ, Willan A. In search of power and significance: issues in the design and analysis of stochastic cost effectiveness studies in health care. Med Care. 1994;32:150–63.

Pekarsky B. The new drug reimbursement game: a regulator's guide to playing and winning. London: Springer; 2015.

Pekarsky B. Trusts, constraints and the counterfactual: reframing the political economy of new drugs. PhD Thesis University of Adelaide, 2012.

Stinnett AA, Mullahy J. Net health benefits: a new framework for analysis of uncertainty in cost-effectiveness analysis. Med Decis Making. 1998;17(4):483–9.

Stinnett AA, Paltiel AD. Estimating CE ratios under second =−order uncertainty: the mean ratio versus the ratio of means. Med Decis Making. 1997;17(4):483–9.

Wakker P, Klaassen MP. Confidence intervals for cost-effectiveness ratios. Health Econ. 1995;4:373–81.

Willan A, Briggs A. (2006). The Statistical Analysis of Cost-effectiveness Data. Wiley and Sons, Chichester.

# Chapter 3
# Avoiding Frankenstein's Monster and Partial Analysis Problems: Robustly Synthesising, Translating and Extrapolating Evidence

## 3.1 Introduction

Chapter 2 highlighted how satisfying decision analytic principles of coverage and comparability with joint consideration of cost and effects are fundamental to allowing unbiased trial-based cost effectiveness analysis for two-strategy comparisons. In this chapter, we extend these principles to identify methods for robust cost effectiveness analysis relevant to decision making in any given jurisdiction of interest (e.g. Australia with a PBAC remit, the UK with a NICE remit), while highlighting common biases, pitfalls and fallacies that should be avoided. Satisfying coverage and comparability principles for unbiased analysis in a jurisdiction of interest is shown to usually require robust methods for evidence synthesis and translation and potentially extrapolation. In particular, methods are required which avoid common inferential fallacies and selection biases associated with partialisation of analysis (O'Brien 1996; Briggs and O'Brien 2001) and alternative framing of binary outcomes (Eckermann et al. 2009, 2011).

The seminal paper 'Frankenstein's Monster or the Vampire of Trials' (O'Brien 1996) provides the backdrop and sets the stage, identifying the usual need for modelled analysis to generalise the protocol-specific nature of RCT evidence of treatment effect, in informing analysis and decision making relevant to a jurisdiction of interest. Modelling is usually required to enable adequate coverage of the scope and duration of expected impacts in estimating incremental absolute effects, costs and net clinical and economic benefit relevant for that jurisdiction. However, while such modelling can aid coverage, robust unbiased methods for evidence synthesis, translation and extrapolation methods are required to satisfy comparability principles and avoid Frankenstein's monster. That is, enable unbiased estimation of incremental absolute costs, effects and net benefit expected for any given jurisdiction/s of interest.

Common examples of fallacies and biased methods leading to Frankenstein's monster are illustrated in this chapter for the cost minimisation method as proposed by Drummond et al. (1987), use of relative risk in evidence synthesis and translation



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_3

of binary events and parametric extrapolation of trial-based primary outcomes. Cost minimisation inferential pathways treat the absence of evidence as evidence of absence where effects are ignored in the absence of statistical significant based on a type I error. Ignoring effects based on type I error biases analysis and creates perverse incentives for vested interests to collect inadequate research on likely inferior interventions and strategies. This cost minimisation fallacy extends the dangers of partialising cost and effects in undertaking cost effectiveness analysis introduced in critiquing the 'box method' in Chap. 2. Hence, the need to jointly consider costs and effects in cost effectiveness analysis established in Chap. 2 is further reinforced in avoiding similar partial analysis fallacies arising with cost minimisation.

In evidence synthesis and translation of RCT evidence, common inferential fallacies and selection biases associated with alternate framing of binary outcomes (survival or mortality, progression or no progression, etc.) are shown to arise with the use of relative risk as a metric (Eckermann et al. 2009, 2011). The use of relative risk is shown to result in differences in direction as well as extent of estimated treatment effect with alternate framing of binary outcomes in indirect comparison (Eckermann et al. 2009). Similarly, in translating trial evidence of treatment effect to estimate absolute risk differences in a jurisdiction of interest, relative risk is similarly shown to be inconsistent with alternate framing of binary events. Indeed, inconsistent results are shown to arise with the use of relative risk whenever binary outcome evidence requires translation, that is, whenever base event risk in the jurisdiction of interest differs from that in relevant comparator trial arm/s (Eckermann et al. 2011) and there is evidence of a treatment effect.

The use of odds ratio metrics is shown to overcome these problems of relative risk, allowing consistent estimation of relative treatment effect in indirect comparison and absolute effect differences for binary events in evidence translation (Eckermann et al. 2009, 2011). These results also provide key building blocks for being able to robustly undertake more complex processes of evidence synthesis and translation, such as evidence translation across jurisdictions with optimal global trial design (Chaps. 6 and 7) and multiple strategy and multiple outcome comparisons (Chaps. 8 and 9).

Robust methods for extrapolation of cost effectiveness evidence beyond trial duration are similarly shown to require consistent unbiased estimation of absolute cost and effects relevant to the indicated use of interventions in any given jurisdiction of interest. Decision analytic methods are shown to be preferred over parametric methods in enabling extrapolated treatment effect and absolute incremental effect and cost differences consistent with continuation rules of strategies compared, levels of side effects, compliance and resistance expected over time, as well as competing risks in surviving populations.

## 3.2 Setting the Scene: Frankenstein's Monster or the Vampire of Trials

To estimate incremental net clinical benefit ($\Delta E$) and net benefit (INB = $\lambda\Delta E - \Delta C$) in any jurisdiction of interest requires estimating absolute incremental effects, resource use and costs of alternative strategies in the relevant jurisdiction. In Chap. 2

key decision analytic and health economic principles of coverage (scope and duration) and comparability of relative and absolute incremental effects were pointed to as needing to be satisfied to obtain unbiased absolute effectiveness or net benefit evaluation estimates in trial-based analysis. In this chapter, we extend the use of coverage and comparability principles to allow robust evidence synthesis, translation and extrapolation of trial-based evidence, necessary to inform estimates of unbiased absolute incremental cost, effect and net benefit estimates relevant to real decision making across jurisdictions of interest (e.g. PBAC in Australia, NICE in the UK). Consequently, this chapter advances along the optimal decision making cycle in Fig. 3.1 from the top left sphere to the second highlighted sphere on its right.

O'Brien (1996) considered the relative merits of two approaches to cost effectiveness evaluation: trial- and model-based analysis. The title of the paper 'Frankenstein's Monster or the Vampire of Trials' aptly summarises and highlights the inherent strengths and weakness of each approach in attempting to jointly satisfy coverage and comparability principles and robustly inform decisions. Randomised control tri-

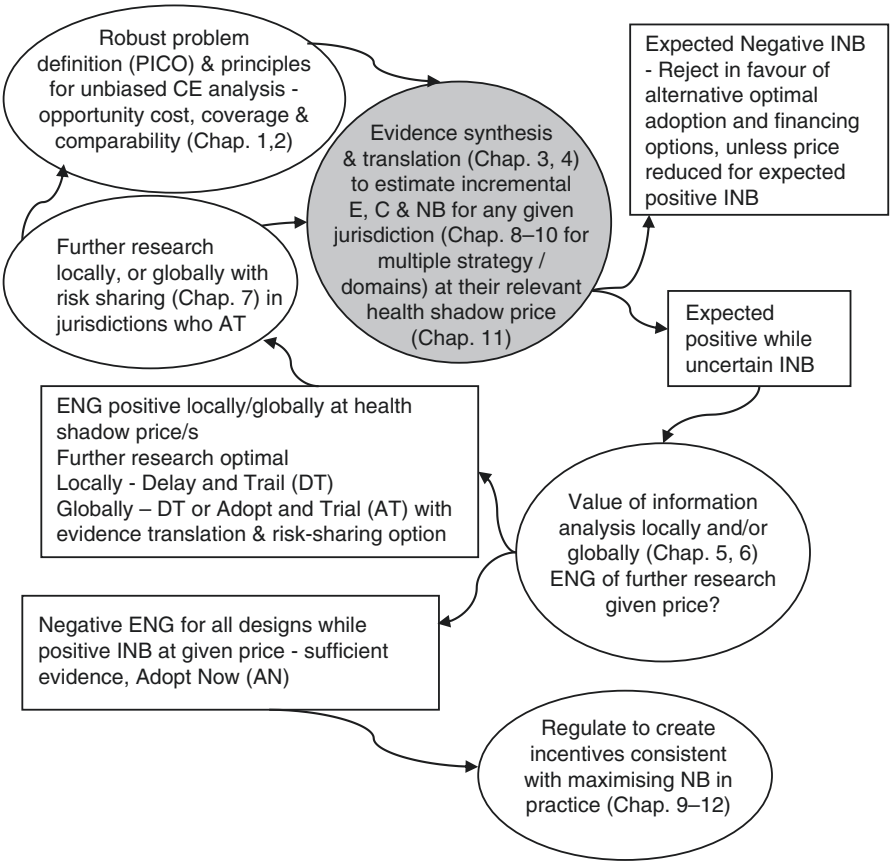


**Fig. 3.1** Optimal decision making cycles for joint research, reimbursement and regulatory processes locally and globally

als, in allowing unbiased estimation of relative treatment effect on health outcomes and associated health-care resource use and costs satisfy the principle of comparability. However, while the ability to avoid biases in estimating relative treatment effects within trial is a strength, trial estimates of absolute effects and costs are like the puncture marks of a vampire, highly specific to the settings and conditions of the trial. That is, absolute incremental effects, resource use and costs estimated from trial evidence are specific to the trial conditions with respect to:

(i) Trial arm interventions or strategies compared;
(ii) Trial populations recruited and associated baseline risk with trial inclusion and exclusion criteria, diagnostic build-up and background level of care;
(iii) Effects collected, where a primary effect alone will usually be inadequate in estimating net incremental effect or net benefit;
(iv) Length of follow-up; and
(v) Protocol conditions more generally.

Consequently, coverage is almost always an issue in informing cost effectiveness analysis for a jurisdiction (e.g. Australia in the case of PBAC) presented with RCT evidence. Such RCT evidence usually requires synthesis with other relevant evidence and evidence translation to that jurisdictions circumstances to reflect the absolute expected effects, resource use and associated costs relevant to that jurisdiction, given their population, practice, preference and prices.

Conversely, modelling can reflect local practice and pathways (structure of tree), synthesis of trial evidence in informing relative treatment effects (indirect comparisons, meta-analysis, etc.), translation of evidence to allow for different population risks (base risk modification), jurisdiction prices and preferences (change payoffs) and extrapolation beyond trial (markov chain monte carlo modelling methods). Hence, modelling provides the potential for coverage to be robustly satisfied. However, maintaining comparability with models to prevent becoming a Frankenstein's monster requires robust methods in structuring models, synthesising evidence from multiple sources and translation and extrapolation of evidence.

The vampire of trials highlights that trial-based analysis presented (alongside an introduction to coverage and comparability principles) in Chap. 2 rarely directly informs decision making in any given jurisdiction of interest. Nevertheless, Frankenstein's monster emphasises that in modelling to address this, processes of unbiased evidence synthesis, translation and extrapolation are usually required to robustly inform such decision making, as described in the second sphere of Fig. 3.1 and depicted in action in Fig. 3.2.

Estimating incremental absolute event rates, cost and effects, incremental net clinical benefit and incremental net benefit for a jurisdiction of interest such as Australia for the PBAC in Fig. 3.2 usually requires the following steps:

(i) Synthesising trial evidence of relative treatment effects for relevant effects (survival, morbidity, side effects, etc.) in a systematic review of trial evidence. Ideally, a meta-analysis of direct comparison but in practice also indirect

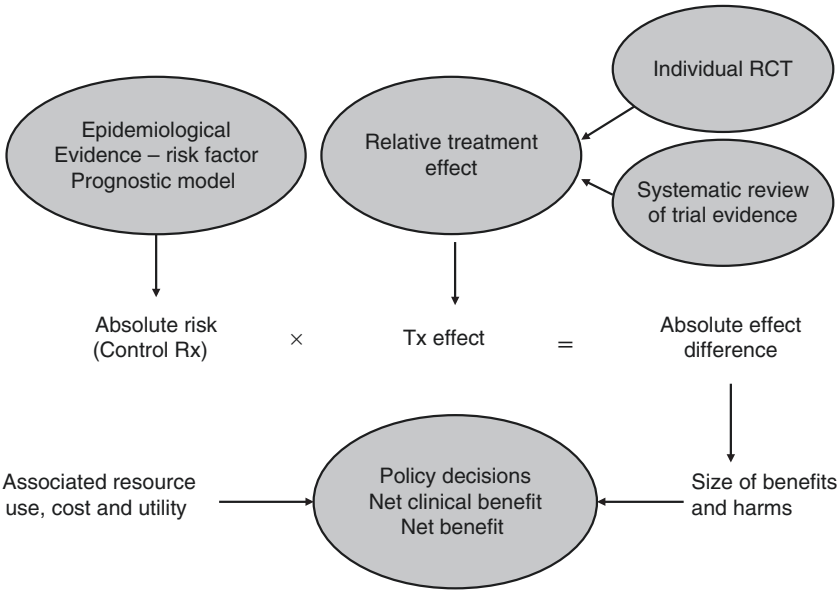


**Fig. 3.2** CE analysis and decision making for a jurisdiction of interest (Source: Health Economics from Theory to Practice course)

comparisons via common comparators (see Sect. 3.4), most commonly placebo-controlled trials.

(ii) Estimating the expected baseline risk of events in the population relevant to decision making. That is, evidence of baseline risk for the relevant patient population and indication (population risk factors – age, sex, comorbidities and diagnostic build-up for current interventions in practice) in the jurisdiction of interest (Australia for the PBAC). This can be estimated directly with epidemiological evidence or indirectly from population characteristics using approaches such as prognostic risk factor modelling.

(iii) Translating trial evidence to estimate absolute event rates (survival, morbidity, side effects, etc.) expected for the relevant population in the jurisdiction of interest (Australia for the PBAC). This requires consistent methods for applying relative treatment effect to baseline risk of the relevant population for decision making (Australia in the case of the PBAC) estimating absolute effect differences for typical binary outcome measures (survival, progression, etc.).

(iv) Valuing combined event rate evidence over time (for mortality, morbidity, side effects, etc. as relevant) to estimate net effects in each arm and incremental net clinical effect. Net effects are often able to be estimated with QALYs where utility measures can be integrated with survival over time in weighting morbidity effects net clinical benefit (trading off the size of harms and benefits). However, where there are multiple domains of importance that cannot be integrated with survival as in palliative care (see Chap. 4), evidence of incremental

net effect benefit requires multiple outcome comparisons and decision maker weighting of individual effects (see Chap. 10).

(v) Estimating incremental costs accounting for expected cost offsets with resource use expected from treatment needs associated with effects from (iv), multiplied by relevant prices over time to enable estimating follow-up costs to combine with expected costs of directly considered treatment alternatives.

(vi) Estimating incremental net benefit as the value of incremental effects less incremental costs, conditional on threshold values for effects.

Note that in steps (i) and (ii), the best evidence of treatment effect is from randomised controlled trials; however, the best evidence of baseline risk is usually from current practice in the jurisdiction of interest, epidemiological evidence rather than RCT evidence. Changes in effects calculated in step (iii) are incremental absolute effects, not efficacy. Combining the value of harms and benefits, net clinical benefit in the population and indication of interest for the jurisdiction of interest (Australia in the case of the PBAC) can be estimated. Often ideally this is in terms of incremental quality adjusted life years, integrating survival and quality of life impacts over time, but more generally allowing for adequate coverage of multiple effect domains of interest (see chapter 4 and 10 for palliative care examples particularly).

Steps (i) and (iii) require robust and consistent methods to undertake evidence synthesis for binary outcomes such as survival, disease progression or symptom response. Consistent estimation of the direction and extent of relative treatment effect with binary outcomes in evidence synthesis and absolute effect differences in evidence translation is shown in Sects. 3.4 and 3.5 to require use of odds ratio rather than relative risk, following Eckermann, Coory and Willan (2009, 2011). Odds ratios overcome selection biases associated with relative risk, where the extent and direction of relative and absolute effects change with framing of binary outcomes (such as survival, mortality; response, no, response; disease progression or no progression). Indeed, these issues with relative risk are shown to arise whenever the use of a relative treatment effect in indirect comparison or evidence translation might have been useful. In contrast, odds ratios provide an unbiased method, ensuring consistent estimation of the direction and extent of relative treatment effect in indirect comparison and absolute risk differences in evidence translation.

The combining of absolute effects and costs to estimate incremental net clinical benefit and net benefit in steps (iv) to (vi) enables analysis to directly inform policy decisions around whether therapies are effective and cost effective in any given jurisdiction. While providing the potential to address coverage issues of trials and associated biases that result, modelling needs to employ robust methods that avoid biases associated with comparability issues. Appropriate care needs to be taken to maintain comparability and avoid inferential fallacies and biases in modelled analysis. Otherwise, modelling can easily lead to Frankenstein's monster with inferential fallacies and biases arising from the way evidence is combined, translated and extrapolated to estimate absolute incremental effects, costs and cost effectiveness for jurisdictions of interest.

O'Brien (1996) makes clear that the weaknesses of each approach need to be addressed – coverage with trial-based analysis and comparability with modelling. Hence, there is usually a need to combine trial evidence and model-based analysis to

enable unbiased estimation of absolute incremental cost, effects and net benefit (cost effectiveness) in a jurisdiction of interest. While Chap. 2 focused on the need for coverage in avoiding the vampire of trials, we focus in this chapter on the problems of comparability to avoid Frankenstein's monster in modelling.

The following sections highlight common problems in maintaining comparability and show solutions to overcome common fallacies that arise in:

(i) Evidence synthesis with cost minimisation (Briggs and O'Brien 2001) and indirect comparison (Eckermann et al. 2009) in Sects. 3.3 and 3.4, respectively
(ii) Evidence translation (Eckermann et al. 2011) in Sect. 3.5
(iii) Extrapolation in Sect. 3.6

## 3.3 Cost Minimisation and the Absence of Evidence Fallacies

Drummond et al. (1987) suggested cost minimisation analysis as appropriate where there is 'no statistically significant difference in relative effect' between trial arms, proposing that only direct cost of interventions is considered in such cases. This approach effectively sets up a partial sequential hypothesis test where the hypothesis is tested that effects are statistically different at a pre-specified level of type I error (typically 5% with a two-sided test) and then:

(i) If effects are tested as different, consider cost effectiveness, cost benefit and cost utility analysis; or
(ii) If effects are not tested as different, then cost minimisation analysis is indicated.

For example, consider comparing interventions A and B, where the relative risk for a 12-month survival for A versus B is estimated as 0.50, with 95% CI (0.20,1.20) and the direct cost of A is $1000 less than B (say $19,000 vs. $20,000). Following Drummond (1987), the lack of statistical significance of the trial finding with a 5% two-sided type I error leads to the following inferential trial:

(i) Infer an equivalent 12-month survival between A and B from a lack of statistical significance against a 5% two-sided type I error.
(ii) Cost minimisation is justified by 1.
(iii) A is inferred to weakly dominate B with $1000 lower direct cost and 'equivalent' effect from (i).

Briggs and O'Brien (2001) highlight that problems with the inferential trial for cost minimisation analysis as proposed in Drummond (1987) arise by inferring equivalence of effect from a hypothesis test against a probability of type I error in:

(i) Not allowing for type II error (probability of false negative) or associated notions of powering in testing with difference trials, rather than equivalence trials; and
(ii) Ignoring expected cost implications that arise from differences in effects (i.e. ignoring cost and effect covariance).

Given (i) and/or (ii) a weak dominance inferred from the inferential trial followed by Drummond (1987), can easily represent a dominated or cost ineffective strategy. Further, (i) and (ii) also imply that incentives are created for manufacturers or more generally those with a vested interest in promoting an intervention where initial evidence points towards such interventions being inferior to collect inadequate evidence. That is, by treating absence of evidence as though it were evidence of absence (Altman and Bland 1995), scant evidence can be collected for likely inferior interventions and then be ignored in the absence of statistically significant differences.

For example, in the case of A versus B, while the direct cost of A may be $1000 cheaper, A is expected to have half the survival rate of B at 12 months, which for a late-stage cancer population might translate to a 10% versus 20% survival rate at one year.

If we directly present this evidence on the incremental cost effectiveness plane only comparing A versus B, then the point estimate of an ICER for A versus B would lie in the south-west quadrant, where a $1000 reduction in costs for A is at the expense of an expected 10% absolute reduction in survival rate. This clearly does not indicate dominance of A relative to B. Indeed if such interventions have expected lower effect while cost saving (in the SW quadrant), they should be compared with the best alternative actions in raising funds for investment, as discussed at length in Chap. 11 following Eckermann (2015). That is, interventions or strategies with incremental expected costs and effects in the SW quadrant should be compared with the best alternative disinvestment action that with the lowest reduction in health expected to arise in saving the health system costs, in order to fund alternative actions, or equivalently the greatest cost saving (funding raised) for any given health loss.

More generally, the absence of evidence being treated as evidence of absence (Altman and Bland 1995; Briggs and O'Brien 2001) is particularly problematic with cost effectiveness analysis where effect inference is based on a type I error as proposed by Drummond (1987), given biases that can result for cost as well as effect estimation and their interaction, in informing cost effectiveness estimates and decision making. That is, effects ignored are usually also expected to have cost implications for the health system where they have associated treatment, such as hospitalisation for various forms or morbidity or treatment of side effects. Hence, an intervention which on the face of a cost minimisation analysis has lower direct cost and 'equivalent effect' can easily represent worse effects, but also higher health system costs in appropriately including the cost of treating such effects. What cost minimisation following Drummond (1987) suggests as weak dominance ('equivalent effect' and lower direct strategy cost) can therefore easily represent a dominated strategy in reality.

To avoid inferential fallacies that arise with cost minimisation as proposed by Drummond (1987) and more generally clarify cost effectiveness analysis, the focus should be on jointly presenting estimates of incremental costs and effects relative to relevant comparator/s and ideally their joint density under uncertainty (Briggs and O'Brien 2001). Consequently, critiquing cost minimisation analysis leads to the same

conclusions for principles and methods required for robust cost effectiveness analysis as critiquing of the box method, following Briggs et al. (2002) in Chap. 2. Costs and effects should always be considered together to avoid partial analysis biases that otherwise arise in consideration of societal decision making in relation to cost effectiveness, or net benefit, analysis.

## 3.4 Indirect Comparison and Avoiding Framing Biases with Relative Risk in Evidence Synthesis of Binary Outcomes

Direct RCT evidence may not always be readily available to inform relative and absolute incremental effect, incremental cost and net benefit assessment between two strategies of interest, say A and B, for a variety of reasons. For example, due to parallel development of strategies, lack of co-operation between vested interests representing alternative interventions or lack of independent bodies to broker and robustly undertake relevant unbiased head-to-head research to directly inform societal decision making.

Nevertheless, while direct head-to-head RCT evidence of A versus B may not be available, RCT evidence will often be available for A and B relative to a common comparator (C), where typically C is placebo. That is, trials will often have been undertaken for A versus placebo and B versus placebo. In such cases, indirect comparison of A versus B via common comparator C can be employed in an attempt to estimate relative treatment effects and subsequently absolute incremental effects, net clinical benefit, costs and net benefit between A and B. Indirect comparison of A versus B via common comparator C graphically can be represented as A – C C – B.

The validity or otherwise of such simple indirect comparison or more complex, multiple forms of indirect comparison such as network and meta-analysis (Efthimiou et al. 2016) relies on the assumption of exchangeability of evidence between trials (Greenland and Robins 1986). That is, the assumption that the same result would have arisen with the indirect comparison method employed if the trial settings (protocol conditions for inclusion and exclusion criteria for populations, diagnostic build-up and treatment practice, etc.) had been exchanged. Hence, indirect comparisons become more robust and tenable the more exchangeable trial settings are, or, to the extent trial setting factors such as baseline level of population risk differ, unbiased and consistent methods can be found to adjust for such differences.

To illustrate relative merits of alternate methods for attempting to improve exchangeability in practice, consider a simple example in a multiple sclerosis population comparing MS progression where the comparison of interest is between natalizumab and interferon (Eckermann et al. 2009). Indirect comparison is based on RCT evidence of the risk of multiple sclerosis progression in trials of natalizumab versus placebo (Polman et al. 2006; Australian Government Department of Health and Ageing 2007) and interferon versus placebo (Douquette et al. 1995) presented in Table 3.1.

**Table 3.1** Inadequacy of indirect comparison using difference in differences methods for multiple sclerosis progression natalizumab versus placebo and interferon versus placebo

| No. (%) with no progression | No. (%) with no progression |
|---|---|
| Natalizumab | Placebo |
| 418/627 (66.7%) | 129/315 (41.0%) |
| Interferon | Placebo |
| 36/115 (31.3%) | 18/112 (16.1%) |
| ARD (% with no progression)<br>Natalizumab vs. placebo = 66.7 − 41.0 = 25.7%<br>Interferon vs. placebo = 31.3 − 16.1 = 15.2% | |

A difference in differences estimate for an indirect comparison of natalizumab versus interferon in Table 3.1 would lead to a 10.5% estimate of the absolute risk of no progression for natalizumab versus interferon (25.7% less 15.2%). However, where comparator base risks differ across arms, direct calculation of risk differences between A and B using a difference in differences method is biased where risk difference is modified by baseline risk (Greenland 1987; Engels et al. 2000; Deeks 2002; Furakawa et al. 2002). Hence, note that the baseline risk in the placebo arm substantially differs across the two studies (41.0% in the natalizumab RCT and 16.1% in the interferon RCT). This difference in placebo (baseline) risk is expected to be reflected in absolute risk difference for any given relative treatment effect of natalizumab and interferon relative to placebo. That is, with absolute risk difference modified by baseline risk and positive relative treatment effects for both natalizumab and interferon relative to placebo, the use of a difference in differences method biases against interferon, given a much lower placebo arm (baseline) rate of multiple sclerosis nonprogression in the interferon study.

If common comparator arm risks differ, and indeed strictly unless common comparator arm rates are exactly the same, then the use of relative measures such as relative risk and odds ratio have been suggested as a means of avoiding bias and helping maintain exchangeability between trials (Greenland 1987; Rothman 1996). Such relative effect measures show less heterogeneity across trials than arithmetic effect (absolute difference) measures in the vast majority of cases (Engels et al. 2000; Deeks 2002; Furakawa et al. 2002). Consequently, the need to use relative treatment effects to avoid bias in estimating absolute risk differences where baseline risks are not the same in trial arms has been well established. However, the natural question to consider is which relative treatment effect should be used, relative risk (RR) or odds ratio (OR)?

Assessing net clinical benefit and net benefit with indirect comparisons will usually require estimation of risk differences for binary events and outcomes such as survival (mortality), disease progression (no disease progression), response (no response), etc. Where absolute risk difference is modified by relative treatment effect and trial evidence comes from an indirect comparison then net clinical or economic benefit assessment for a jurisdiction of interest involves both:

(i) Estimation of relative treatment effect from trial indirect comparison, an important consideration in its own right which we consider in this Sect. (3.4); and
(ii) Translation of relative trial treatment effect to jurisdiction/s of interest, considered in the next Sect. (3.5).

While only the second step is required to estimate absolute risk difference for binary outcomes with direct randomised control trial evidence, in the case of indirect comparisons robust estimation of treatment effect is required first.

Hence, the first question we consider is whether OR or RR provides consistent estimation of relative treatment for binary outcomes with indirect comparisons.

The use of relative risk to estimate the direction and extent of treatment effect in indirect comparison is undertaken using a 'clinically intuitive' ratio of ratios method. Namely, where relative risk of A versus B is estimated as the ratio of relative risk for A versus placebo relative to the relative risk for B versus placebo. However, binary outcomes such as survival (mortality), disease progression (no disease progression) and response (no response) can each be framed one of two ways, form a positive (utility bearing) prespective or a negative (disutility bearing) perspective. Hence we need to consider whether the use of relative risk allows consistent estimation of relative treatment effect in indirect comparisons such as interferon versus natalizumab via placebo, given alternative framing of a relevant binary outcome of interest–e.g. multiple sclerosis disease progression or no progression (Table 3.2).

Consider this binary outcome framed first from a negative or disutility bearing prespective as multiple sclerosis (MS) progression on the left side of Table 3.2. RCT evidence for natalizumab versus placebo MS progression has a relative risk of 0.57 (0.333/0.590), which relative to a relative risk of progression of 0.82 (0.687/0.839) with interferon versus placebo leads to an indirect comparison relative risk of 0.70 (0.57/0.82) for natalizumab versus interferon in terms of MS progression. Hence, indirect comparison of relative risk for disease progression suggests natalizumab is favoured. That is, the indirect comparison estimate of RR progression for natalizumab versus interferon of less than 1 favours natalizumab, given multiple sclerosis progression is an outcome to be avoided.

Now, consider the same indirect comparison where the same binary outcome is framed from a positive or utility bearing perspective as no progression of multiple sclerosis (MS). RCT evidence for natalizumab versus placebo with no MS progression has a relative risk of 1.63 (0.667/0.410), which relative to a relative risk of no MS progression of 1.95 (0.313/0.161) with interferon versus placebo leads to an indirect comparison relative risk of 0.84 (1.63/1.95) for no MS progression with natalizumab versus interferon. Now, a relative risk less than 1 with no MS progression for

**Table 3.2** Framing bias with inconsistent estimation of direction and extent of treatment effect from indirect comparison with RR

| *Natalizumab no. (%) with progression* | *Placebo no. (%) with progression* | *Natalizumab no. (%) with no progression* | *Placebo no. (%) with no progression* |
|---|---|---|---|
| 209/627 (33.3%) | 186/315 (59.0%) | 418/627 (66.7%) | 129/315 (41.0%) |
| $RR_{progression} = 0.57$ | | $RR_{no\ progression} = 1.63$ | |
| *Interferon progression* | *Placebo progression* | *Interferon no progression* | *Placebo no progression* |
| 79/115 (68.7%) | 94/112 (83.9%) | 36/115 (31.3%) | 18/112 (16.1%) |
| $RR_{progression} = 0.82$ | | $RR_{no\ progression} = 1.95$ | |
| Indirect comparison N. vs. I.<br>Progression RR = 0.57/0.82 = 0.70 (favours natalizumab) | | Indirect comparison N. vs. I<br>No progression RR = 1.63/1.95 = 0.84 (favours interferon!) | |

natalizumab versus interferon is unfavourable for natalizumab; given no progression of MS is a positive (utility bearing) outcome over time. Hence, indirect comparison of RR for no MS progression favours interferon, a diametrically opposite finding to that with RR for MS progression!

Alternate framing of the binary outcome (no progression of multiple sclerosis) with relative risk leads to reversal of the direction of treatment effect in indirect comparison. Consequently, such use of relative risk in indirect comparison creates selection bias, where those with vested interests can select the result they prefer simply by choosing which way to frame the binary outcome. The lack of consistent estimation with relative risk in the direction let alone extent of relative or absolute treatment effect for binary outcomes with relative risk is critical, given the imperative for consistent methods in supporting the exchangeability of evidence assumption (Greenland 1987; Rothman 1996), underlying indirect comparison (Eckermann et al. 2009). To understand why the treatment reversal and lack of consistency arise with the use of relative risk in indirect comparison, and consider how to address this problem, we need to explore what is causing it.

### 3.4.1 *What Causes Reversal of Treatment Effect with Relative Risk?*

Relative risk is not a symmetric metric with respect to framing binary outcomes from a positive (utility bearing) or negative (disutility bearing) perspective. That is:

$$RR_{\text{event}} \neq {}^{1}\!\Big/\!_{RR_{\text{no event.}}}$$

In indirect comparisons, interaction between this lack of symmetry and differences in common comparator (e.g. placebo) risks across trials creates conditions for reversal of treatment effect with alternative framing of outcomes. For example, as observed with indirect comparison of natalizumab versus interferon, where relative risk was less than 1 with binary outcomes framed from both a utility and disutility bearing perspective ($RR_{UAB}$ and $RR_{DAB} < 1$). Equally, reversal of treatment effect can arise where relative risk is greater than 1 with alternative framing of outcomes for indirect comparison from a utility or disutility bearing perspective (i.e. where $RR_{UAB}$ and $RR_{DAB} > 1$; Eckermann et al. 2009).

Reversal of the direction, but more generally differences in the extent, of treatment effect estimated with the use of relative risk in indirect comparison is inconsistent and undermines exchangeability required to support indirect comparison and creates clear scope for selection bias in framing outcomes. To overcome these problems with relative risk, their cause needs to be addressed, the lack of symmetry of the relative risks metric with alternative framing of binary outcomes. That is, a symmetric metric is required to allow consistent estimation of the direction and extent of treatment effect in indirect comparisons.

Odds ratios (ORs), as the ratio of odds, where odds are the probability of an event happening divided by the probability of it not happening:

$$\left( \text{Odds}_{\text{event}} = \frac{P(\text{event})}{P(\text{no event})} = 1/\text{Odds}_{\text{no event}} \right)$$

provide a symmetric measure of treatment effect. That is for comparisons between treatments *a* and b for a binary outcome (which can either be framed as having an event or no event):

$$\text{OR}_{a,b\,\text{event}} = \text{Odds}_{a(\text{event})} / \text{Odds}_{b(\text{event})}$$

$$= \frac{P_a(\text{event})}{P_a(\text{no event})} \left( \frac{P_b(\text{event})}{P_b(\text{no event})} \right)^{-1}$$

$$= \left( \frac{P_b(\text{no event})}{P_b(\text{event})} \right) \left( \frac{P_a(\text{no event})}{P_a(\text{event})} \right)^{-1}$$

$$= \text{Odds}_{b(\text{no event})} / \text{Odds}_{a(\text{no event})}$$

$$\Rightarrow \text{OR}_{a,b\,\text{event}} = 1/\text{OR}_{a,b\,\text{no event}}$$

Eckermann, Coory and Willan (2009) show that the symmetric odds ratio ensures consistent estimation of the direction and relative extent of treatment effect in indirect comparisons. Table 3.3 illustrates this for the case of indirect comparison between natalizumab and interferon.

Considering progression first (left-hand side of Table 3.3), the odds ratio of MS progression for natalizumab versus placebo is 0.35 ( = (0.333/0.667)/(0.590/0.410)), while for interferon versus placebo is 0.42 ( = (0.687/0.313)/(0.839/0.161)). This leads to an indirect comparison odds ratio with MS prgression of 0.83 (0.35/0.42) for natalizumab versus interferon. Hence, indirect comparison of the odds ratio for progression favours natalizumab, given multiple sclerosis progression is an event

**Table 3.3** Consistent estimation of direction and extent of treatment effect in indirect comparisons with OR

| *Natalizumab progression* | *Placebo progression* | *Natalizumab no progression* | *Placebo no progression* |
|---|---|---|---|
| 33.3% | 59.0% | 66.7% | 41.0% |
| Natalizumab vs. placebo Progression OR = 0.35 | | Natalizumab vs. placebo No progression OR = 2.88 | |
| *Interferon progression* | *Placebo progression* | *Interferon no progression* | *Placebo no progression* |
| 68.7% | 83.9% | 31.3% | 16.1% |
| interferon vs. placebo Progression OR = 0.42 | | interferon vs. placebo No progression OR = 2.38 | |
| OR for progression *n* vs. I 0.83 (favours natalizumab) | | OR for no progression *n* vs. I 1.21 (favours natalizumab) | |

the target population want to avoid and hence an odds ratio of progression less than 1 is favourable for natalizumab.

Now, consider indirect comparison for no progression (right side of Table 3.3). For binary events the odds, and odds ratios, of no events are in general the reciprocal of those for events. The odds ratio for no progression with natalizumab versus placebo is 1/0.35 = 2.88 = (0.667/0.333)/(0.410/0.590), while for interferon versus placebo is 1/0.42 = 2.38 ( = (0.313/0.687)/(0.161/0.839)). This leads to an indirect comparison odds ratio for no progression of 1.21 (2.88/2.38) for natalizumab versus interferon, which is equal to the reciprocal of that for progression (1.21 = 1/(2.38/2.88) = 1/0.83).

Hence, whether framed as disease progression or no disease progression, indirect comparison of odds ratios favours natalizumab with the same direction and extent of treatment effect. In general, Eckermann, Coory and Willan (2009) show that the problems of lack of consistent estimation with alternate framing with relative risk in indirect comparisons are overcome with odds ratios, which ensure consistency in the direction and extent of treatment effect with alternative framing of binary outcomes. The reciprocal nature of symmetric odds ratios ensures that for binary outcomes the indirect comparison estimates of an event (mortality, morbidity, progression, etc.) are the reciprocal of that without the event (survival, no morbidity, no progression).

The bottom line then in choosing a robust metric for simple or complex forms of indirect comparison to improve exchangeability of evidence with binary outcomes is that:

(i) A relative treatment effect is required to inform trial-based or translated jurisdictional evidence of absolute incremental effects, where absolute treatment effect is expected to be modified by baseline risk.
(ii) Of candidate relative treatment effects (relative risk or odds ratio), only the odds ratio provides a consistent estimate of the direction and extent of treatment effect and prevents selection bias with alternative framing of binary outcomes.

These advantages of odds ratio over relative risk with indirect comparisons also extend to network analysis (i.e. more complex forms of indirect comparison; Welton et al. 2012). Further, they also extend to translating relative treatment effects to estimation of absolute risk difference in any given jurisdiction of interest with direct or indirect comparisons, as we now show in Sect. 3.5, following Eckermann, Coory and Willan (2011).

## 3.5 Preventing Framing Biases in Evidence Translation

Evidence from a trial of relative treatment effect usually requires translation to any given jurisdiction of interest in estimating absolute effect and cost differences allowing for differences in baseline risk between the trial setting and those expected in the jurisdiction of interest. To allow unbiased estimation of absolute incremental effects and costs relevant to any given jurisdiction and their consequent relevant decision making (net clinical benefit, INB), such estimates have been shown to usually require trial relative treatment effects modifying (Greenland 1987; Engels et al. 2000; Deeks 2002; Furukawa et al. 2002) epidemiological estimates of that jurisdictions baseline risk. That is, the use of a relative treatment effect should modify the baseline risk in

the relevant jurisdiction of interest to allow unbiased estimation of absolute risk differences, net clinical benefit and net benefit in that jurisdiction, as depicted in Fig. 3.2.

For binary events such as survival, response or disease progression, the moot question in translating trial evidence to a jurisdiction of interest, as it was in indirect comparison, is which treatment metric allows consistent estimation of such absolute risk differences? Under the usual assumption that relative treatment effects are homogenous across baseline risk (Greenland 1987; Engels et al. 2000; Deeks 2002; Furukawa et al. 2002), relative risk or odds ratio could be applied to estimate risk differences for binary events, noting binary outcomes can be framed one of two ways (survival vs. mortality; response vs. no response; progression vs. no progression, etc.).

Historically, several epidemiologists (Sinclair and Brackn 1994; Davies et al. 1998; Sackett et al. 1996) have suggested that as relative risk estimates are easy to calculate and apply directly to baseline risk in translating evidence, clinicians might think on a relative risk scale and misinterpret odds ratios. On this basis, relative risk has been suggested as a preferred metric for characterising relative treatment effect in translating evidence. However, they did not consider whether relative risk provides a consistent estimate of risk differences for binary outcomes in translating trial evidence to jurisdictions of interest.

### *3.5.1 Does Relative Risk Consistently Estimate Absolute Risk Difference in Translating Evidence with Alternate Framing of Binary Events?*

In translating trial evidence to a jurisdiction of interest, the convention for estimating the jurisdiction risk for the active therapy with relative risk is to simply multiply the relevant trial relative risk estimate by epidemiologically estimated baseline risk in the jurisdiction of interest for the control arm therapy. Hence, for example, if baseline risk of survival for a condition were estimated as 70% with current practice in a jurisdiction of interest and relative risk of survival with a new treatment relative to the current practice therapy in that jurisdiction were 1.3, then the estimated survival for the new treatment option in that jurisdiction with relative risk would be 0.91 (1.3 multiplied by 0.7). Consequently, the estimated increase in absolute survival rate with the new relative to current therapy would then be 91% less 70% or 21%.

However, with binary events such risks can be framed and estimated from a utility bearing or a disutility bearing perspective (survival or mortality, progression or no progression, etc.). Relative risk, as discussed in Sect. 3.4, is not a symmetric metric, a property which:

(i) Walter (2000) noted as very troubling;
(ii) Fleiss (1994) argued should effectively rule relative risk out as a useful metric for meta-analysis; and
(iii) Eckermann, Coory and Willan (2009) show prevents consistent estimation of the direction let alone extent of relative treatment effect in indirect comparisons.

Hence, we need to consider whether in translating trial evidence of binary effects to estimate absolute effect differences in any given jurisdiction relative risk provides

a consistent estimate with alternate framing. Is absolute risk difference estimated with relative risk from a utility bearing perspective (e.g. survival) consistent with that from a disutility bearing perspective (e.g. mortality) when translating trial evidence of a treatment effect to estimate expected absolute risk differences in a jurisdiction?

Following Eckermann, Coory and Willan (2011), we consider direct trial evidence for therapy A versus B with risk of a binary event (say survival or mortality) in each arm. Let the absolute rates in trial arms ($a$, $b$) framed from a utility bearing perspective (e.g. survival) be $RU_a$ and $RU_b$, while from a disutility bearing perspective (mortality) be $RD_a = 1 - RU_a$ and $RD_b = 1 - RU_b$. Relative risk framed as survival (utility bearing perspective) is $RU_a/RU_b$ and framed as mortality (disutility bearing perspective) is $RD_a/RD_b$.

Now, consider translating the trial evidence of relative risk to a jurisdiction of interest whose rates we denoted with *. Hence, for the jurisdiction of interest, their baseline risk is $RU_b$* framed from a utility bearing perspective (e.g. survival) and $RD_b$* from a disutility bearing perspective (e.g. mortality).

Estimating absolute risk difference ARD from a utility ($ARD_u$) and disutility ($ARD_d$) bearing perspective with relative risk following Eckermann, Coory and Willan (2011) leads to estimates of:

$$\mathrm{ARD}_u = \mathrm{RU}_b^* \left( \mathrm{RU}_a / \left( \mathrm{RU}_b - 1 \right) \right)$$

and

$$\mathrm{ARD}_d = \left(1 - \mathrm{RU}_b^*\right)\left(1 - \left(\left(1 - \mathrm{RU}_a\right) / \left(1 - \mathrm{RU}_b\right)\right)\right)$$

Now, if these estimates were consistent, then we would have $ARD_u = ARD_d$ or equivalently $ARD_u - ARD_d = 0$.

In Eckermann, Coory and Willan (2011), the expression for differences between absolute risk differences framed from a utility and disutility bearing perspective is shown in appendix 1 to simplify to:

$$\mathrm{ARD}_u - \mathrm{ARD}_d = \frac{\left(\mathrm{RU}_a - \mathrm{RU}_b\right)}{\left(1 - \mathrm{RU}_b\right)} \frac{\left(\mathrm{RU}_b - \mathrm{RU}_b^*\right)}{\left(\mathrm{RU}_b^*\right)}$$

Thus, absolute risk differences estimated with relative risk from a utility and disutility bearing perspective are the same, if and only if the first and/or second term of this equation is 0. That is, translating with relative risk is consistent only if there is no treatment effect and hence $\frac{\left(\mathrm{RU}_a - \mathrm{RU}_b\right)}{\left(1 - \mathrm{RU}_b\right)} = 0$ and/or the baseline risk in the jurisdiction of interest (RU*b) matches that from the trial (RUb) and hence:

$$\frac{\left(\mathrm{RU}_b - \mathrm{RU}_b^*\right)}{\left(\mathrm{RU}_b^*\right)} = 0$$

However, these conditions under which relative risk provides a consistent estimate of absolute risk difference in translating evidence reflect the only cases where evidence does not require translation. That is, where there is no treatment effect from the trial and/or the baseline risk is the same in the jurisdiction of interest as that in the trial. For cases where translation of evidence is of interest, there is a treatment effect ($RU_a \neq RU_b$) and baseline risk in the jurisdiction of interest differs from that in the trial, relative risk does not provide a consistent estimate of risk difference.

To illustrate this result and explain why relative risk should not be used for evidence translation of binary outcomes, consider a simple translation case. Consider a trial with 20% mortality in the active arm and 40% in the control arm being translated to a jurisdiction where the mortality rate was 30% for the control arm therapy.

Framed in terms of mortality, the relative risk from this trial is 0.5 (0.2/0.4), and hence the mortality risk in the jurisdiction if the new therapy was adopted would be estimated with relative risk as 15% ($0.5 \times 0.30 = 0.15$) and represents an absolute mortality reduction of 15% (30% less 15%).

Now, framed in terms of survival, the relative risk from this trial is 1.333 (0.8/0.6). Hence, the survival rate expected in the jurisdiction if the new therapy was adopted would be estimated with relative risk as 93.3% ($1.333 \times 0.70 = 0.933$) and represents an absolute survival increase of 23.3% ($0.933 - 0.70 = 0.233$).

Hence, alternative framing of the same binary event evidence translated to the same jurisdiction does not provide consistent estimates of effect differences with relative risk. Different estimates of absolute risk difference arise depending on which way the binary event is framed.

Further, note that an additional bounding problem can also arise with relative risk. For example, consider if the trial evidence had suggested a larger treatment effect for survival, say 1.6. Translating that trial evidence to the jurisdiction of interest with a base rate of survival of 70% would lead to a survival rate estimate in the active arm of 112% ($1.6 \times 0.7 = 1.12$)!

Hence, the use of relative risk to translate trial evidence to a jurisdiction of interest can easily lead to estimates of binary event probabilities which are not appropriately bound between 0 and 1. Greenland (1987) suggested such bounding problems with relative risk establish logical reasons with binary outcomes for not believing constancy in the ratios of risk that relative risk represent. That is, bounding problems with relative risk provides purely logical reasons for not employing relative risk in translating trial evidence of binary outcomes.

In summary, relative risk should not be used to translate trial evidence of binary outcomes to jurisdictions of interest both because (i) it is not consistent with respect to framing of binary outcomes whenever translation is required and (ii) it does not appropriately bound risk of binary outcomes between 0 and 1.

Having highlighted problems of relative risk, we now turn our attention to see whether the other candidate, the odds ratio, performs any better in translating evidence of binary outcomes to a jurisdiction of interest.

### 3.5.2 *Does the Odds Ratio Allow Consistent Estimation of Absolute Risk Difference in Translating Trial Evidence to Jurisdictions of Interest?*

Let us consider the same evidence translation case we explored with relative risk in Table 3.2 but now using odds ratios in Table 3.4.

In Table 3.4, we first consider absolute risk difference calculated with odds ratios where the binary event is framed as mortality (left-hand side of Table 3.4). The odds ($O = P/(1 - P)$) of mortality in arm A is 0.2/0.8 = 1/4, arm B is 0.4/0.6 = 2/3 and hence the odds ratio for A relative to B is 1/4/(2/3) = 1/4 × 3/2 = 3/8. In the jurisdiction of interest, base risk of mortality with the control arm therapy in current practice is 30% reflecting a prior base odds of mortality of 0.3/0.7 = 3/7. Hence, applying the treatment odds ratio to the prior odds, the odds of treatment arm therapy expected in the jurisdiction of interest is 3/8 × 3/7 = 9/56. Converting the treatment odds back to a probability ($P = O/(1 + O)$) = (9/56)/(65/56) = 9/65. Consequently, the absolute risk difference is estimated as the base rate of mortality of 3/10 less the treatment mortality of 9/65. Converting to a common denominator, the reduction in mortality equates to 39/130 − 18/130 = 21/130.

Now we consider the right-hand side of Table 3.4, estimating absolute risk difference calculated in translating trial evidence with odds ratios for survival. The odds $O = P/(1 - P)$ of survival in arm A is 0.8/0.2 = 4, while in arm B is 0.6/0.4 = 3/2, and hence the odds ratio for A relative to B is 4/(3/2) = 4 × 2/3 = 8/3. Note that in each case, the calculations with odds ratios for survival were the exact reciprocals of that for mortality. In the jurisdiction of interest base survival rate with the control arm therapy in current practice is 70%, and hence base odds of survival is 0.7/0.3 = 7/3, again the reciprocal of that for mortality. Hence applying the treatment odds ratio for survival to the base (jurisdiction of interest current control arm) odds of survival, the odds of treatment arm survival expected in the jurisdiction of interest is 8/3 × 7/3 = 56/9, again the reciprocal of that with mortality. Converting the treatment odds back to a probability ($P = O/(1 + O)$) = (56/9)/(65/56) = 56/65. Consequently, the absolute risk difference is estimated as treatment survival of

**Table 3.4** Translating trial evidence to a jurisdiction using odds ratios

| A mortality | B mortality | A survival | B survival |
|---|---|---|---|
| 20% | 40% | 80% | 60% |
| Trial OR mortality A vs. B | | Trial OR survival A vs. B | |
| (0.2/0.8)/(0.4/0.6) = 3/8 | | (0.8/0.2)/(0.4/0.6) = 8/3 | |
| Jurisdiction Y: mortality (arm B) in practice 30% | | Jurisdiction Y: survival (arm B) in practice 70% | |
| B odds mortality = 0.3/0.7 = 3/7 | | B odds survival = 0.7/0.3 = 7/3 | |
| A odds mortality = 3/8 × 3/7 = 9/56 | | A odds survival = 8/3 × 7/3 = 56/9 | |
| A risk mortality = 9/65 | | A risk survival = 56/65 | |
| ARD mortality = 3/10 − 9/65 = 21/130 | | ARD survival = 56/65 − 7/10 = 21/130 | |

56/65 less base risk of survival of 7/10. Converting to a common denominator, the increase in survival equates to 112/130–91/130 = 21/130, the same as the reduction in mortality. The symmetric nature of odds and odds ratio creates a mirror image with alternate survival or mortality framing and the same estimate of absolute risk difference.

More generally, Eckermann, Coory and Willan (2011) show algebraically that the use of odds ratios generates the same expression for absolute risk difference in translating evidence to a jurisdiction of interest regardless of binary outcome (event or no event) framing. That is, regardless of whether binary effects are framed from a utility (survival, no disease, no progression, etc.) or disutility bearing perspective (mortality, disease, progression, etc.), the absolute risk difference is the same, consistent, estimate. Further, this result applies regardless of whether evidence of treatment effect being translated comes from a direct trial, meta-analysis of trial data or indirect comparison, since OR and ARD are consistently estimated in each of these settings (Eckermann et al. 2009, 2011).

Consequently, the odds ratio addresses inconsistency and selection bias problems arising with using relative risk for binary outcomes in translating evidence to allow consistent estimation of absolute risk differences. Importantly, such consistent estimation is necessary for unbiased clinical and cost effectiveness analysis relevant to any given jurisdiction of interest. To consistently translate evidence with odds ratios in estimating risk differences in a jurisdiction of interest involves three simple steps:

(i) Convert baseline probability estimate for usual care (control arm therapy) in the jurisdiction of interest (indicated by * superscript) to odds ($O_b^* = P_b^*/(1 - P_b^*)$).
(ii) Apply odds ratio from trial (treatment relative to usual care control) to current odds in the jurisdiction of interest (from i) to estimate odds in treatment arm in that jurisdiction of interest ($O_a^* = \mathrm{OR}_{ab} \times O_b^*$).
(iii) Convert the treatment arm odds estimate for the jurisdiction of interest back to a probability ($P_a^* = O_a^*/(1 + O_a^*)$).

Note that applying odds ratios to baseline control arm odds in the jurisdiction of interest to estimate treatment arm odds in that jurisdiction and then converting back to a probability also ensures probability is appropriately bounded between 0 and 1, unlike relative risk. Appropriate bounding is ensured given odds are 0 or positive, and hence in converting back to probability, odds divided by 1 plus odds is appropriately bounded between 0 (where odds are 0) and 1 (as odds asymptotically head towards infinity).

Combining the steps (i) to (iii) in the odds ratio method of evidence translation above leads to an expression for treatment risk of $P_a^* = \mathrm{OR}_{ab} \times P_b^*/(1 - P_b^* + \mathrm{OR}_{ab} \times P_b^*)$, where $\mathrm{OR}_{ab}$ is the odds ratio from trial evidence of a relative to comparator b and $P_b^*$ is the base risk of the binary event in the jurisdiction of interest. In a decision analytic model, this direct expression for $P_a^*$ can be placed directly into the treatment arm for the relevant jurisdiction of interest to inform decision making in that jurisdiction along with $P_b^*$, the epidemiologically determined risk in that jurisdiction for the control arm.

The bottom line is that in evidence synthesis and translation of binary outcomes odds ratios, unlike relative risk, allow consistent estimation of the direction and

extent of treatment effects and risk differences and appropriately bound treatment risk of binary outcomes. This also provides a key stepping stone for undertaking more complex analysis without bias. The rejection of relative risk in favour of odds ratios in estimating absolute risk differences for binary outcomes also carries over to comparison of multiple strategies (see Eckermann and Willan 2011; Eckermann, Briggs and Willan 2008 and Chap. 8) and multiple outcomes (McCaffrey 2013; McCaffrey et al. 2015 and Chap. 9); meta-analysis of multiple studies and network analysis (Efthimiou et al. 2016; van Valkenhoef and Ades 2013); and standardisation of binary event rates (for example see Eckermann and Coelli 2013 and Chap. 9).

In relation to indirect comparison and network meta-analysis, van Valkenhoef and Ades (2013) discussion of rank reversal being unlikely where transitivity (exchangeability) is met does not address the more general issue in maintaining exchangeability of RR inconsistency and selection bias with alternate framing of binary outcomes in any form of indirect comparison or evidence translation with RR (Eckermann et al. 2009, 2011). This is particularly key for public policy and economic analysis where robust unbiased evidence translation (Eckermann et al. 2011) as well as indirect comparison methods (Eckermann et al. 2009) are required, including those underlying network meta-analysis.

Inconsistency and selection bias with RR in framing of binary outcomes arise wherever baseline risk differs in common comparator arms in indirect comparison (Eckermann et al. 2009) or between trial control arm and that in the jurisdiction of interest in translating evidence (Eckermann et al. 2011), that is whenever it might have been useful. This is simply overcome with the use of odds ratios which are consistent and hence prevent selection bias with alternative framing of binary outcomes, better maintaining evidence exchangeability.

Consequently, regardless of whether there is reversal of the direction of treatment effect or not, economic analysis will be inconsistent and biased with the use of relative risk whenever required in evidence synthesis or translation of binary outcomes, while consistent and unbiased with odds ratios. Indeed, inconsistency and resulting selection bias with alternate framing of binary outcomes in using relative risk estimates for indirect comparison or their translation to estimating effect differences should rule out the use of RR in all such economic analysis.

## 3.6 Extrapolating Cost Effectiveness Evidence Beyond Trial Duration for a Jurisdiction of Interest

Economic and decision analytic principles for unbiased cost effectiveness analysis and related decision making support randomised control trial evidence as the most comparable evidence but also require robust methods for evidence coverage. That is, avoiding biases with adequate coverage of scope and duration of relevant effects and costs using robust, consistent methods for evidence synthesis, translation and extrapolation as necessary to inform cost effectiveness decision making by relevant jurisdictions, such as Australia with the PBAC in Fig. 3.2.

To enable robust extrapolation of effects, costs and their joint distribution for cost effectiveness analysis, we need to extend principles and methods for unbiased analysis to beyond study impacts. To achieve unbiased analysis when extrapolating cost effectiveness evidence in practice, the following key factors should be considered in jointly satisfy principles of coverage, comparability and consistency:

(i) Explicitly state and justify assumptions with respect to treatment effect and baseline risk beyond study follow-up, conditional on indication, and in particular, whether treatment is continued or not beyond study given stopping rule conditions, resistance, intolerance and side effects or compliance over time.
(ii) Consistently translate trial evidence to the jurisdiction/s of interest.
(iii) Extrapolate incremental costs and outcomes consistently.
(iv) Consistently allow for within and beyond trial follow-up decision uncertainty.

What is often presented to decision makers as extrapolated cost effectiveness analysis are parametric extrapolation of a primary outcome in both arms, usually survival or some other binary event (disease progression, response). Typically, this is undertaken using parametric functional forms for extrapolation of within-study effects, such as Gompertz, Weibull, lognormal or log-logistic functions (Fig. 3.3).

Such analysis might also include discussion about why one function has been preferred over others in terms of goodness of fit – fitting within-study data.

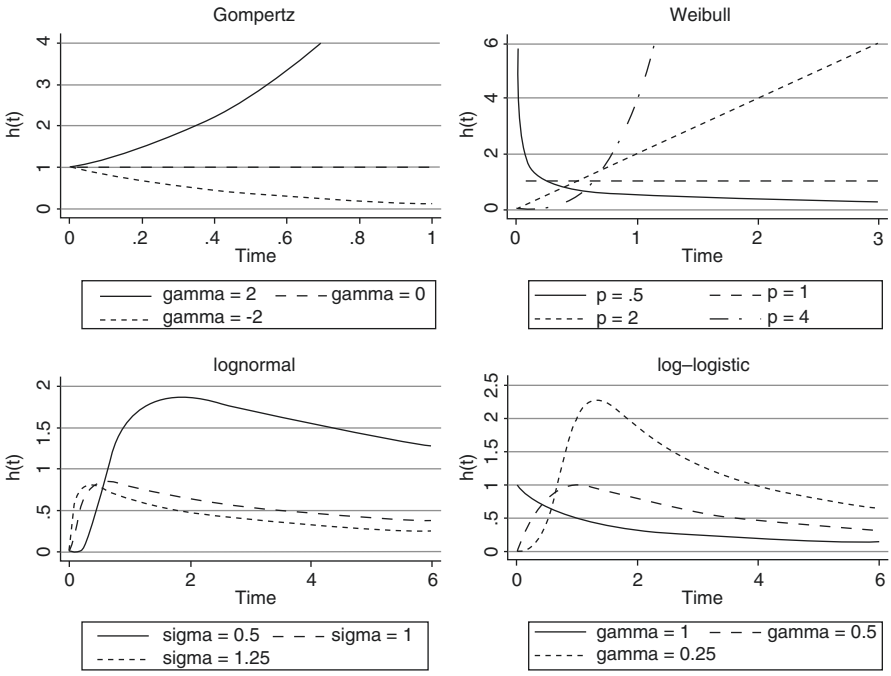


**Fig. 3.3** Alternative parametric extrapolation functions

However, in relation to satisfying decision analytic and health economic principles and in particular the four key factors for unbiased extrapolation, parametric extrapolation of trial survival curves in both arms:

(i) Makes an implicit assumption of within-study treatment effect being continued beyond study – while often not justified or justifiable.
(ii) Does not translate evidence to jurisdiction of interest – while baseline risk should reflect those of population, practice, etc. in the jurisdiction of interest and relative treatment effect appropriately translated (with odds ratio rather than relative risk with binary outcome – see Sect. 3.5).
(iii) Does not extrapolate effects other than the primary outcome (side effects, competing risks, etc.) nor extrapolate associated costs.
(iv) Does not consistently allow for combined uncertainty of cost and all relevant effects within study, let alone within and beyond trial.

Hence, while such parametric extrapolation approaches may often have considerable appeal to authority in relation to goodness of fit, they do not satisfy criteria for an unbiased extrapolation of cost effectiveness evidence relevant to a jurisdiction of interest.

In contrast, extrapolation with decision analytic models can appropriately be used to address these four key factors for unbiased analysis by:

(i) Modelling of treatment effect consistent with indication – no further, reversal or continued treatment effect beyond trial – as well as consistent evidence synthesis metrics and methods used for binary outcomes (see Sect. 3.4 and Eckermann et al. 2009).
(ii) Translating evidence to the jurisdiction/s of interest – baseline risks of local population, practice, etc. and prices relevant to the decision being made within study and beyond study allowing for competing risks, as well as consistent evidence translation metrics and methods with binary outcomes (see Sect. 3.5 and Eckermann et al. 2011).
(iii) Consistently modelling costs and all effects (Briggs, Blackhouse and O'Brien 2002) within and beyond study.
(iv) Modelling joint cost and effect uncertainty within trial combined with sensitivity analysis with respect to scenario uncertainty beyond trial.

In relation to (i), extrapolation of treatment effect is conditional on indication. Treatment effect beyond study follow-up should be consistent with whether compared treatments are continued or not beyond trial periods and more generally continuation/stopping rules, time profile of resistance, intolerance, side effects, non-compliance etc. Ceteris paribus (all other things being equal), continued treatment effect is less likely and reversal of treatment effect more likely where active treatment is not continued or there are expected levels of resistance, intolerance, side effects and/or non-compliance.

To illustrate, consider extrapolating within-study trial evidence of cost per life saved presented in Chap. 2 for the LIPID study of pravastatin use as a secondary preventive for coronary heart disease (Eckermann and Kirby 2003). Within study

there was an absolute mortality reduction of 3.01% at 6 years, corresponding to a 0.78 relative risk. In extrapolation beyond 6 years for such an individual-based pharmacotherapy, one needs to consider the relative treatment effect beyond 6 years and in broadest terms whether there is no further treatment effect, a continuation, or reversals, of treatment effect (as depicted in Fig. 3.4).

In considering treatment effects and costs beyond 6 years, the primary consideration is whether statin therapy is continued or not. That is, expected treatment effects and direct and follow-up costs beyond study are conditional on whether the indication is to continue statin therapy or stop. If the indication is stopped, then no further treatment effect beyond study or reversal of treatment effect may be considered most likely, while if treatment is continued, then further treatment effect becomes more likely.

In general, the indication beyond study evidence is a primary consideration in extrapolating treatment effect before considering other evidence such as:

(i) Risk factors amongst study survivors by arm given events within-study and prior risk factors (which may differ by arm with selection effects for survival) as risk factors for events post study;
(ii) Evidence from beyond study follow-up of study cohorts; and
(iii) Evidence of biological mechanisms for therapies actions (how long are mechanisms expected to last).

With respect to risk factors amongst survivors (i), in the case of the LIPID study, analysis of prognostic risk factors amongst survivors at end of study follow-up (the populations to extrapolate beyond study impacts from) suggests that survivors in the pravastatin arm were at lower risk of future CHD events compared to survivors in the placebo arm. Overall, pravastatin arm survivors had a significantly lower multifactor CHD prognostic risk score (1.34 vs. 1.42, $p < 0.05$; Marschner et al. 2001).

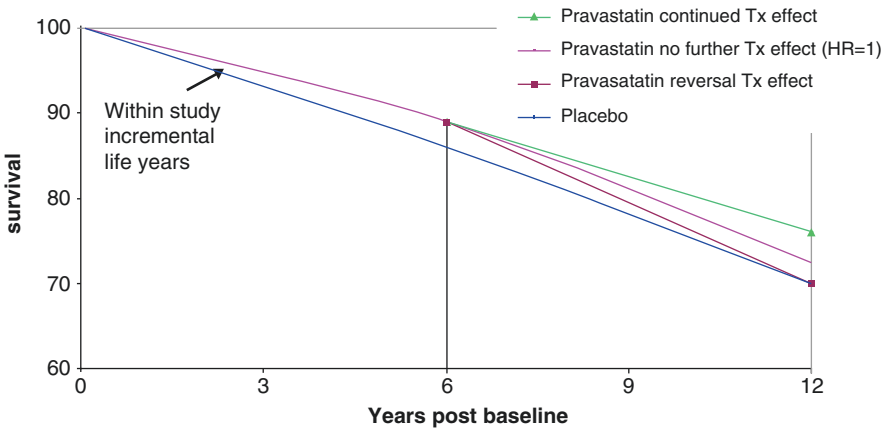


**Fig. 3.4** Contingencies for beyond study treatment effect, LIPID study

**Table 3.5** Differences in event rates amongst LIPID study survivors

| Risk factor | Reduction in events amongst pravastatin vs. placebo survivors | *P* value |
|---|---|---|
| MI | 1.48% | 0.0008 |
| Stroke | 0.34% | 0.063 |
| UAP | 0.84% | 0.215 |
| Any of MI, stroke, UAP | 2.25% | 0.0005 |
| Revascularisation | 1.99% | 0.0005 |

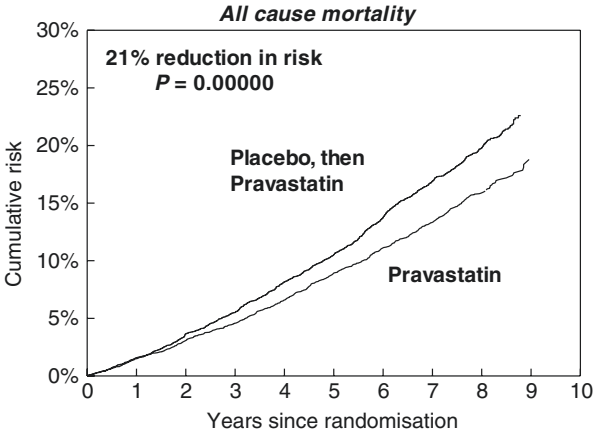


**Fig. 3.5** Mortality in LIPID cohort over average 2 years further follow-up with both arms on pravastatin

This is attributable to lower absolute rates of MI, stroke and unstable angina pectoris (UAP) amongst survivors and in spite of protective effects of higher rates of revascularisation (Table 3.5).

This evidence provides some support for continued treatment effect on CHD events beyond study where further treatment of this population is indicated and where statin therapy is not continued a conservative assumption of no further treatment or potentially even some form of limited continued treatment effect. A continued treatment effect beyond 6 years where treatment is continued is further supported by the observed mortality rate up to 8 years in the pravastatin arm of a follow-up study (see Fig. 3.5).

It should be noted that the relative treatment effect observed beyond study follow-up of 6 years in this figure is likely somewhat conservative given the placebo arm surviving cohort beyond trial end was provided with study medication. Never the less in that respect note the lack of separation of survival curves between pravastatin and placebo over the first year or indeed 15 months of the study. Hence, the placebo arm (then pravastatin) rates and treatment effect could less conservatively be inferred as still reflecting that with placebo alone somewhere beyond 7 years.

However, the actual study follow-up for survival beyond 6 years with open-label pravastatin therapy shown for 2 additional years in both arms in Fig. 3.5 given high levels of compliance with open-label therapy effectively shows the impact of delaying therapy by 6 years. The open-label study follow-up has con-

tinued from 6 years within-study follow-up to now 16 years follow-up (Hague et al. 2016), where the absolute survival benefit observed at 6 years has been maintained. That is, the absolute 3.0% mortality rate improvement to 6 years has been maintained over the additional 10 years of follow-up with an average 85% and 84% statin use in the post-study open-label follow-up for study patients assigned pravastatin and placebo, respectively.

Consequently, the extended follow-up evidence generally supports a clinically plausible impact of statin therapy within study being maintained or continuing beyond 6 years with further treatment. Nevertheless, with stopping of therapy, a clinically plausible reversal of treatment effect (converging of survival curves) could also be expected. It should be clear that such clinically plausible mechanisms while supportive generally require relevant evidence to move beyond conjecture and prior beliefs.

Evidence of prognostic event risks by arm for survivors at study end and continued study follow-up interpreted conditional on the use of intervention/s beyond study is key for decision makers to consider in informing the likelihood of alternative beyond study relative treatment effect scenarios in extrapolation. Nevertheless, given unknown decision maker priors, and uncertainty in relation to beyond study relative treatment effects from alternative scenarios, presenting decision makers with conditional analysis to cover cases of no further, continued or reversal of treatment effect is suggested alongside evidence where available to inform the likelihood of such eventualities. For example, in the case of the LIPID study, considering costs, effects and cost effectiveness conditional on alternative continued, no further and reversal of treatment effect (respectively, further separated, proportionally constant and converging survival curves in Fig. 3.4).

In the case of LIPID, life expectancy of survivors at study end was able to be estimated at a median age of survivors of 69 from within-study life tables up to age 82 as 8.9 years. This estimate corresponds to the additional life expectancy of survivors up to age 82 with no further treatment effect. For the other contingencies of continued or reversal of treatment effect for the same period as the study (6 years) followed by no further treatment effect, areas under survival curves calibrate with a life expectancy up to age 82 for additional survivors of 14 and 4 years, respectively. The range from 4.0 to 8.9 and 14.0 years of additional life years per survivor at study end with reversal, no further or continued treatment effect, reflects beyond study uncertainty about effects. However, we need to combine within and beyond study uncertainty to reflect extrapolated uncertainty.

Applying these contingent life expectancies to survival rates in each bootstrap replicate within study from estimates in Sect. 2.7; life years by arm conditional on post-study relative treatment effect can be estimated for each replicate and considered relative to incremental cost in constructing conditional CEA curves (Fig. 3.6).

The resulting contingent CEA curves allow for an interaction between within and beyond study uncertainty in extrapolation as a function of potential threshold values for effects. In the case of pravastatin for secondary CHD prevention, for thresholds at or above $50,000 per life year saved, pravastatin is expected to be cost effective with 100% or close to 100% certainty for continuing, no further or even reversal of

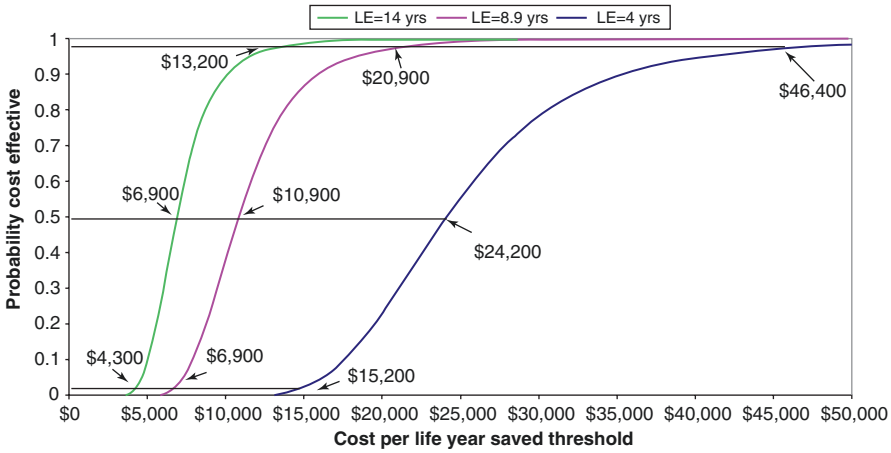


**Fig. 3.6** Lipid extrapolated CEA curves conditional on continued, no further and reversal of within-study treatment effect (Source: Eckermann and Kirby 2003)

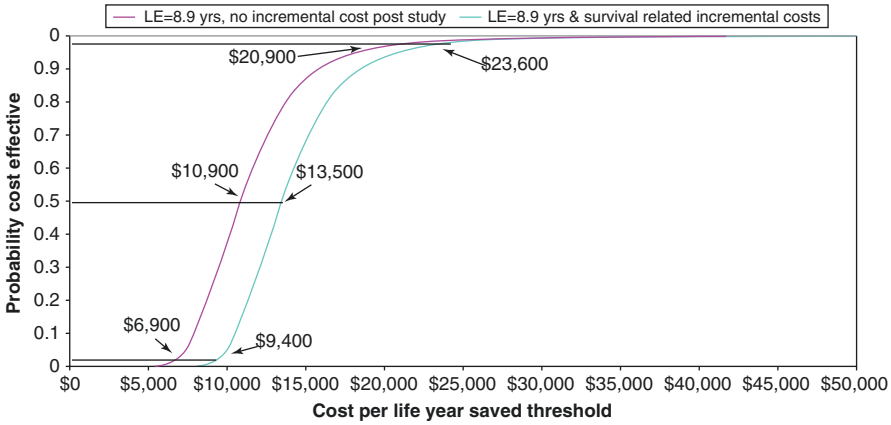


**Fig. 3.7** Conditional LIPID cost effectiveness acceptability curves including incremental costs of treating additional survivors (Source: Eckermann and Kirby 2003)

treatment effect. However, at a threshold of say $20,000 per life year saved, this is only the case for the continuing or no further treatment effect contingencies.

Note that the extrapolated analysis for costs informing such decisions may not be conservative to the extent it models incremental costs as not changing from that within study under the assumption that increased costs associated with additional survivors are 'offset' by lower event rates from healthier survivors (Glasziou et al. 2002).

To be more conservative in relation to incremental costs, one can consider the cost of additional survivors, that is, model incremental cost allowing for expected cost of treating additional survivors. This is presented in Fig. 3.7 for the case of no further treatment effect, where applying absolute survival difference at study end (3.0%) to lifetime costs (up to age 82 as per life years) leads to an estimated

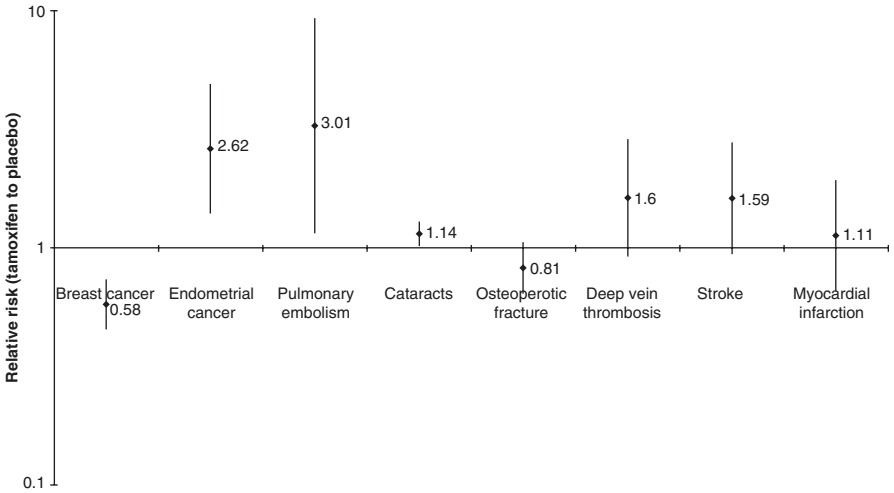


**Fig. 3.8** Benefits and harms of tamoxifen as a preventative (Source: Eckermann, Martin, Sockler and Simes 2003)

additional incremental cost per patient of $667. This increases total incremental costs from $3246 to $3913 per patient and results in an expected incremental cost per life year increasing from $10,900 to $13,500 (Eckermann and Kirby 2003).

It should be clear that the extrapolation approach employed for the LIPID study where extrapolated survival analysis relies on within-study life table evidence and the marrying of within and beyond study analysis are extrapolated from one major event – survival – faces distinct limitations in being applied more generally to other settings, and particularly in extrapolating costs.

More generally robust decision analytic modelling is required to allow consistent estimation of incremental cost and effects and unbiased cost effectiveness analysis extrapolation. To illustrate such decision analytic approaches, we consider the use of Monte Carlo Markov Chain (MCMC) methods in modelling extrapolated incremental cost effectiveness. This is illustrated for a modelled analysis of the extrapolated incremental cost effectiveness of tamoxifen used as a preventative for breast cancer in Australia given evidence from trials internationally (Eckermann et al. 2003).

The largest and first reporting RCT, the US National Surgical Adjuvant Breast and Bowel Project (NSABP) trial (P-1), suggested daily oral tamoxifen significantly reduced risk of breast cancer at 5 years by 49% in women at risk and recommended adopting for all women over 50. However, they also found significantly increased risks of endometrial cancer, pulmonary embolism, cataracts as well as increased prevalence of vasomotor and gynaecological dysfunction (Fisher et al. 1998). Later international trials found lower relative risk reduction, while in extrapolating evidence beyond 5 years tamoxifen use could plausibly either delay or truly prevent breast cancer and in undertaking cost effectiveness analysis, costs need to be considered. A meta-analysis combining evidence from available trials for relative risk of all harms and benefits is shown in Fig. 3.8.

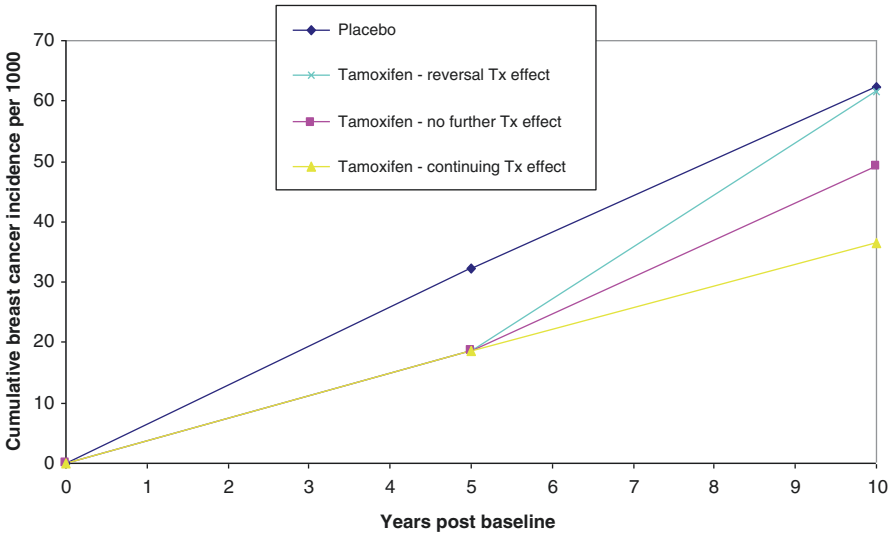


**Fig. 3.9** No further, reversal and continued treatment effect contingencies for tamoxifen study extrapolation (Source: Eckermann, Martin, Sockler and Simes 2003)

The relative risk for breast cancer at 5 years with combined RCT evidence was 0.58 or a 0.42 relative risk reduction.

In considering the extrapolation of 5-year evidence available to 10-year evidence, there could plausibly have been no further treatment effect, reversal (converging of curves at 10 years) or continued treatment effect (continued diverging of event curves to 10 years), as Fig. 3.9 depicts.

In modelling treatment effects on breast cancer from 5 to 10 years and breast cancer and competing risks mortality beyond that to lifetime impacts MCMC methods were used with decision models of relevant health states and treatment costs and outcomes. These methods enable consistent modelling of survival and quality of life (disutility) impacts of harms (PE, cataract, endometrial cancer, side effect) and benefits (breast cancer reduction) and age-dependant competing risks (other deaths) under uncertainty for the Australian population. Importantly they also enabled consistent estimation of the associated expected Australian health system cost of acute (PE, cataract, endometrial cancer) and terminal (breast cancer) conditions and side effects (gynaecological and vasomotor dysfunction). These decision models also enabled for contingent extrapolation scenarios with true prevention of breast cancer (no further or continued treatment effect from 5 to 10 years) and delay of breast cancer (with reversal of treatment effect from 5 to 10 years such that breast cancer incidence converged at 10 years).

From these alternative extrapolation model scenarios (no further, reversal and continued treatment effect) contingent estimates of the expected incremental cost, effects and cost effectiveness ICER (Table 3.6), as well as their bivariate distribution on the incremental cost effectiveness plane (Fig. 3.10) and CEA curves (Fig. 3.11) and tornado diagrams (Fig. 3.12) were constructed.

**Table 3.6** Tamoxifen discounted incremental costs, QALYs and cost/QALY for alternate extrapolation contingencies from 5 to 10 years

| Incremental | Prevent 5 years | Prevent 10 years | Delay (5 years) |
|---|---|---|---|
| Total costs/patient | $2193 | $1818 | $2698 |
| Life years saved/1000 | 45 | 75 | 9 |
| QALY saved /1000 | 57 | 94 | 13 |
| Cost/life year saved | $48,000 | $24,000 | $288,000 |
| Cost/QALY saved | $38,000 | $19,000 | $199,000 |

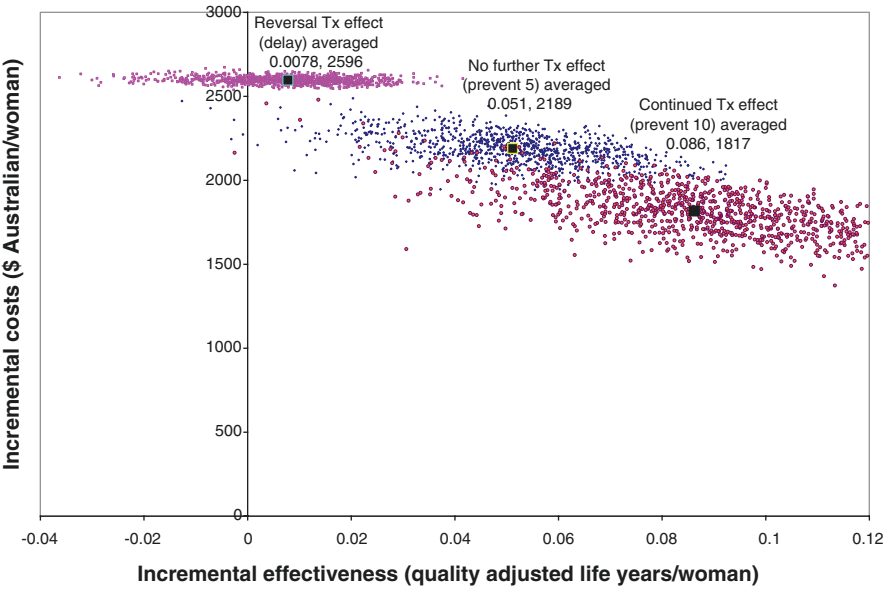


**Fig. 3.10** CE distributions for contingent tamoxifen extrapolations (reversal, no further, continued treatment effect)

These analyses showed decision makers that key uncertainty with extrapolation for decision making was in relation to whether tamoxifen truly prevents or only delays breast cancer. Informed by this analysis, the International Breast Cancer Intervention Study (IBIS) trial was extended from a 5- to 10-year follow-up, rather than further trials of treatment effect up to 5 years to consider the cost effectiveness of tamoxifen as a preventative therapy in these settings.

Such analysis points towards the importance of considering value of information in decision making when considering trial design, which we formally and systematically consider for optimal local and global decision making in relation to joint reimbursement and research decisions and trial design in Chaps. 5, 6 and 7.

Importantly, in leading up to such VOI considerations, unbiased extrapolation of evidence with decision analytic methods for strategies and interventions, such as medications, targeted at individuals can be expected to be considerably less rose tinted than parametric approaches with implicit continued treatment effects suggest. For such

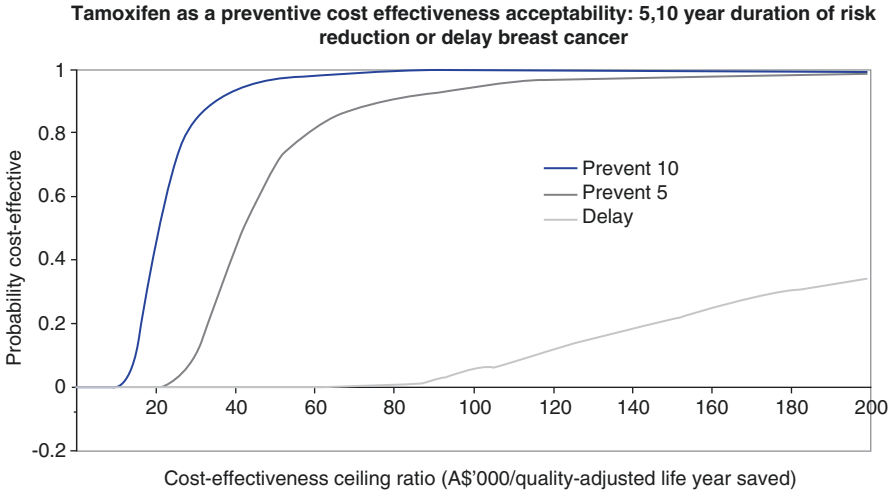


**Fig. 3.11** Contingent CEA curves for tamoxifen reversal, no further or continued treatment effect in extrapolating from 5 to 10 years (Source: Eckermann, Martin, Sockler and Simes 2003)

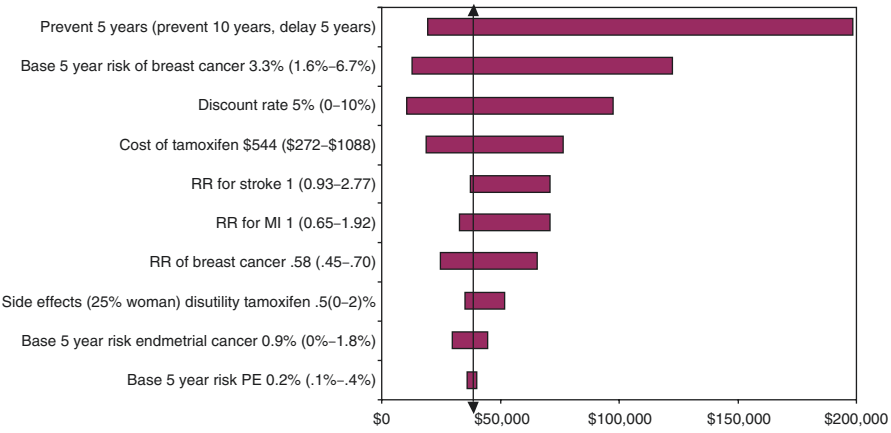


**Fig. 3.12** Tornado diagram one-way sensitivity analyses for tamoxifen incremental cost per QALY (Source: Eckermann, Martin, Sockler and Simes 2003)

individual-targeted therapies, where short-term evidence might suggest a promising treatment effect, such effects should be appropriately tempered when extrapolating with consideration of side effects, tolerance and compliance issues, interactions with other medications, generalisability of results to wider populations and unintended consequences at population levels. For example, consider the societal impact of tamoxifen used as a preventive with an expected additional 25% of the over 50 female population treated having minor side effects such as gynaecological and vasomotor dysfunction. The unintended societal impact of such side effects could be considerably greater at a population level than the sum of individual disutility might suggest, as well as reduced compliance in practice from that in trials.

However, such appropriate tempering in extrapolation of short-term promising results for individual-targeted therapies such as medications is not to say that extrapolation for all programs and strategies is always overestimated for all types of strategies and programs. As Chap. 4, shows for health promotion and prevention programs in community settings, where such programs are effective and have growing multiplier effects over time, then they provide the ability to costlessly expand benefits and significantly improve cost-effectiveness (net benefit) in extrapolation.

That is, effective community-level health promotion and prevention strategies that are successful in building social capital and have community ownership of the strategies can have large multiplier impacts on community networks. The scope of these multiplier effects can be in terms of the combined impacts both of increased population and duration of impacts and with minimal or no costs beyond that observed within-study periods. Hence, when extrapolating impacts of prevention and promotion programs in community settings, community acceptance and network considerations are highlighted in Chap. 4 as key to the success and cost effectiveness of such programs.

## References

Altman DG, Bland JM. Absence of evidence is not evidence of absence. BMJ. 1995;311:485.

Australian Government Department of Health & Ageing. Public sum-mary documents by product: natalizumab, concentrated solution for IV infusion, 300 mg per 15 mL, TysabriÒ November 2006. 2007. Accessed 14 Oct 2008; Available at http://www.health.gov.au/internet/main/ publishing.nsf/Content/pbac-psd-natalizumab-nov06.

Briggs A, O'Brien B. The death of cost minimization analysis. Health Econ. 2001;10:179–84.

Briggs A, O'Brien B, Blackhouse G. Thinking outside the box: recent advances in the analysis and presentation of uncertainty in cost effectiveness studies. Annu Rev Public Health. 2002; 23:377–401.

Davies H, Crombie I, Tvarakoli M. When can odds ratios mislead? BMJ. 1998;316:989–91.

Deeks J. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. Stat Med. 2002;21:1575–600.

Drummond MF, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 1987.

Duquette P et al. Interferon beta-1b in the treatment of multiple sclerosis: final outcome of the randomized controlled trial. Neurology. 1995;45(7):1277–85.

Eckermann S. Kinky thresholds revisited: opportunity costs differ in the NE and SW quadrants. Appl Health Econ Health Policy. 2015;13:7–13. doi: 10.1007/s40258-014-0136-3.

Eckermann S, Willan AR. Globally optimal trial design for local decision making. Health Econ. 2009;18:203–16.

Eckermann S, Willan A. Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies. Pharmacoeconomics. 2011;29(7):563–77.

Eckermann S, Willan A. Optimal global VOI trials: better aligning manufacturer and decision maker interest and enabling feasible risk sharing. PharmacoEconomics. 2013;31:393–401.

Eckermann S, Briggs A, Willan A. Health technology assessment in the cost-disutility plane. Med Decis Mak. 2008;28:172–81.

Eckermann S. Hospital performance including quality: creating incentives consistent with evidence-based medicine. PhD Dissertation. Sydney: UNSW; 2004. http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/.

Eckermann S, Coelli T. Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice. Soc Sci Med. 2013;76:159–68. http://dx.doi.org/10.1016/j.socscimed.2012.10.020.

Eckermann S, Coory M, Willan A. Consistently estimating absolute risk difference when translating evidence to jurisdictions of interest. Pharmacoeconomics. 2011;29(2):87–96.

Eckermann S, Coory M, Willan AR. Indirect comparison: relative risk fallacies and odds solution. J Clin Epidemiol. 2009;62:1031–6.

Eckermann S, Kirby A. Cost effectiveness analysis: uncertainty, predictive conditioning and extrapolation post study – results from LIPID. In: Butler JRG, Quinn C, editors. Economics and health: 2002 Proceedings of the Twenty Fourth Australian Conference of Health Economists, AHES, Sydney. 2003. p. 54–83.

Eckermann S, Martin A, Stockler M and Simes J. The benefits and costs of tamoxifen for breast cancer prevention. ANZJPH 2003;27(1):34–40.

Efthimiou O, Debray TPA, van Valkenhoef G, Trelle S, Panayidou K, Moons KGM, et al. GetReal in network meta-analysis: a review of the methodology. Res Synth Methods. 2016. Available from: https://www.researchgate.net/publication/290219003_GetReal_in_network_meta-analysis_a_review_of_the_methodology. Accessed 22 Jan 2016.

Engels E, Schmid C, Terrin N, Olkin I, Lau J. Heterogeneity and statistical significance in meta-analysis: an empirical study of 125 meta-analyses. Stat Med. 2000;19(13):1707–28.

Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, Vogel V, Robidoux A, Dimitrov N, Atkins J, Daly M, Wieand S, Tan-Chiu E, Ford L, Wolmark N. Tamoxifen for Prevention of Breast Cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study J Natl Cancer Inst. 1998 Sep 16;90(18):1371–88.

Fleiss J. Measures of effect size for categorical data. In: Cooper H, Hedges L, editors. The handbook of research synthesis. New York: Russell Sage; 1994.

Furukawa T, Guyatt G, Griffith L. Can we individualize the 'number needed to treat'? An empirical study of summary effect measures in meta-analyses. Int J Epidemiol. 2002;31:72–6.

Greenland S, Robins JM. Identifiability, exchangeability, and epidemiological confounding. Int J Epidemiol. 1986;15(3):413–9.

Greenland S. Interpretation and choice of effect measures in epidemiologic analyses. Am J Epidemiol. 1987;125(5):761–8.

Hague WE, Simes J, Kirby A, Keech AC, White HD, Hunt D, Nestel PJ, Colquhoun DM, Pater H, Stewart RA, Sullivan DR, Thompson PL, West M, Glasziou PP, Tonkin AM. Long-term effectiveness and safety of pravastatin in patients with coronary heart disease: sixteen years of follow-up of the LIPID study. Circulation. 2016;133(19):1851–60. Epub 2016 Mar 25

Marschner IC, Colquhoun D, Simes JR, Glasziou P, Harris P, Singh BB, Friedlander D, White H, Thompson P, Tonkin A. Long-term risk stratification for survivors of acute coronary syndromes: results from the long-term intervention with pravastatin in ischemic disease (LIPID) study. J Am Coll Cardiol. 2001;38(1):56–63. doi:10.1016/S0735-1097(01)01360-2.

McCaffery N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Is home-based palliative care cost effective? An economic evaluation of the Palliative Care Extended Packages at Home (PEACH) pilot. BMJ Support Palliat Care. 2013;3(4):431–5.

McCaffrey N, Skuza P, Breaden K, Eckermann S, Hardy J, Oaten S, Briffa M, Currow D. Preliminary development and validation of a new end-of-life patient-reported outcome measure assessing the ability of patients to finalise their affairs at the end of life. PLoS One. 2014;9(4):e94316.

McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Better informing decision making with multiple outcomes cost-effectiveness analysis under uncertainty in cost-disutility space. PLoS One. 2015;10(3):e0115544.

O'Brien B. Economic evaluation of pharmaceuticals. Frankenstein's monster or vampire of trials? Med Care. 1996;34(12 Suppl):DS99–108.

Polman CH, O'Connor PW, Havrdova E, Hutchinson M, Kappos L. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. N Engl J Med. 2006;354(9):899–910.

Rothman K. Placebo mania. BMJ. 1996;313:3e4.
Sackett DL, Deeks JJ, Altmen DG. Down with odds ratios! Evid Based Med. 1996;1:164–6.
Sinclair J, Brackn M. Clinically useful measures of effect in binary analyses of randomized trials. J Clin Epidemiol. 1994;47:881–9.
van Valkenhoef G, Ades AE. Evidence synthesis assumes additivity on the scale of measurement: response to 'Rank reversal in indirect comparisons' by Norton et al. Value Health. 2013;16:449–51. doi:10.1016/j.jval.2012.11.012.
Walter S. Choice of effect measure for epidemiological data. J Clin Epidemiol. 2000; 53(9):931–9.
Welton NJ, Sutton AJ, Cooper N, Abrams KR, Ades AE. Mixed and indirect treatment comparisons. In: In evidence synthesis for decision making in healthcare. New York: John Wiley & Sons; 2012.

# Chapter 4
# Beyond the Individual: Evaluating Community-Based Health Promotion and Prevention Strategies and Palliative Care Domains of Effect

**Simon Eckermann and Nicola McCaffrey**

## 4.1 Introduction

Chapter 4 considers challenges faced undertaking health economic analysis to evaluate health promotion and palliative care programs and strategies in complex community settings and more importantly highlight principles and some promising methods and approaches to address these challenges.

When evaluating community-based health promotion programs, the principles and evaluation approach to health system decision making need to take a community perspective. In other words, assessing community acceptance of strategies and expected community net benefit given behavioural, lifestyle and health effects (intended and unintended) across community populations reached are central to determining the success and long-term effectiveness and cost effectiveness of these interventions.

Conventional within-study and individual focussed cost-effectiveness methods, with typical short-term evaluation time frames, and extrapolation modelling methods (such as those considered in Chaps. 2 and 3) have struggled to appropriately assess health promotion and primary prevention strategies. These methods fail to tractably gauge community acceptance or capture the diffusion of long-term outcomes across populations in complex community settings such as schools, aged communities, community gardens and walking groups, etc. Alternative evaluation methods are needed to navigate the coverage (population scope and duration) and comparability of the long-term and diffuse community and population impacts of health promotion strategies.

The research of Shiell and Hawe, pointing to the value of assessing community impacts with network and multiplier methods, is shown in this chapter as a more robust and appropriate approach to satisfying coverage and comparability principles in such community-based health promotion programs. In modelling terms, such multipliers and their trajectory over time are key to assessing long-term community acceptance, network impacts and success of community-based health promotion and prevention programs, in triangulation with qualitative assessment of community acceptance and program impacts on individuals' attitudes, behaviours, etc.



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_4

Use of multiplier methods to enable robust evaluation in such settings is illustrated in evaluating the Stephanie Alexander Kitchen Garden National Program (SAKGNP), a health promotion and primary prevention program undertaken in primary schools (Eckermann et al. 2014; Yeatman et al. 2014).

The need for new methods to address difficulties with the use of conventional individual patient-based effectiveness (e.g. QALY) and cost-effectiveness approaches is also shown to arise at the other end of the health-care spectrum – palliative care. Limitations of conventional cost-effectiveness methods in the palliative setting are shown to arise with lack of robust consideration of the multiple key domains reflecting palliative patients' preferences, including finalising personal and financial affairs, process of death, place of care and distress of family, friends and carers (McCaffrey et al. 2016b). Such domains are not amenable to integration with survival time and hence cannot be incorporated into single-effect comparisons in conventional cost effectiveness analysis – even where patient QALYs are estimated. In addressing these problems, the research of McCaffrey (McCaffrey 2013; McCaffrey et al. 2015a) is highlighted as enabling robust comparison of multiple outcome domains under uncertainty, methods which are considered in detail in Chap. 10, and naturally extends multiple strategy methods from Chap. 8. Such multiple outcome domain comparisons are shown to be necessary to consider diverse outcomes beyond those able to be integrated with survival into individual patient QALYs, informing wider community utility functions and aspects of utility within health. Further, even within a QALY framework, significant value to decision makers in many circumstances is suggested from being able to robustly present multiple events or effects, given the potential for these rates, as well as utility weights, to differ between individuals, communities and across jurisdictions and over time.

## 4.2 Evaluating Health Promotion and Prevention: Moving Beyond Individual Measures Within Study

While the Framingham study (Dawber et al. 1957; Truett et al. 1967) showed major causes of death that are preventable, individual patient rather than community-level programs have had marginal if any impact (Zaza et al. 2005). An increasing body of evidence shows that health promotion and disease prevention programs that do not engage with community networks fail to have an impact on population health over time, particularly beyond program obligation and evaluation periods (Hawe et al. 2009). Health promotion and prevention strategies in community settings are in general most effective where they engage with social networks and build social capital to enable community ownership and embedding of strategies (Hawe and Shiell 2000; Moore et al. 2006).

Without community involvement and ownership, impacts are short lived (Schensul 2005) and fail to be effective in impacting on disadvantaged populations targeted (Hill et al. 2005). Where behaviour models have been successful at population levels, such as telephone messaging in improving physical activity and diet (Eakin et al.

2007), they generally coincide with approaches which actively engage with and enable dynamic impacts over time across community networks. Success of health promotion programs in community settings generally requires engaging with complex systems of networks in those communities (Shiell et al. 2008) – in addition to having impacts on attitudes, behaviour and lifestyle of individuals targeted.

Consequently, conventional cost effectiveness analysis models which estimate effects and costs in individual patients, without considering community network impacts, struggle to estimate long-term effects across populations in community settings (Shiell and Hawe 1995). Conventional assessment of within-study costs and effects typically does not enable assessment of whether programs:

(i) Have community acceptance;
(ii) Will continue beyond evaluation periods; or
(iii) Can be expected to have community impacts or adequate return on investment beyond study.

Beyond observed short-term within-study effects, long-term effects are either not considered or postulated around assumption-based sensitivity and scenario analysis – ignoring community interaction and impacts.

Multiplier effects from program investment flowing across networks into community activities provide a robust quantitative indicator of community ownership, engagement with, and building of, social networks and capital and sustainability of programs over time (Hawe et al. 2009; Shiell et al. 2008). Triangulated with qualitative evidence and typically short-term program effects on individuals during evaluation periods, this enables informed assessment of whether findings are expected to translate into sustainable programs with long-term outcomes across communities (Fig. 4.1). Triangulated and combined assessment of individual and community

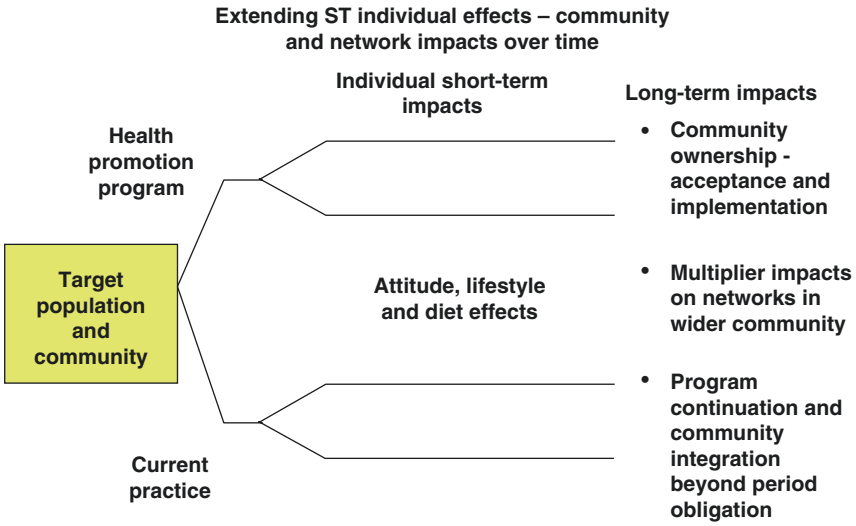


**Fig. 4.1** Extrapolating community impacts of health promotion and prevention programs over time

impacts is key in health promotion, as the WHO guide to economic evaluation in health promotion highlights:

> To make a true economic assessment of health promotion options, one must be forward thinking and consider many different avenues to arrive at a given result (de Salazar et al. 2007: p. 1).

Within-study analysis of health promotion and prevention can be undertaken as per HTA of chronic and acute interventions looking at within-study incremental costs and effects for individuals. However, given the community focused and long-term nature of health promotion and disease prevention strategies, extrapolated analysis is required to evaluate the long-term impacts and potential of such strategies. Otherwise, analysis systematically biases comparison between community health promotion and individual focused treatment strategies.

For interventions focused on the individual (e.g. medications), community ownership and acceptance are often shift factors on compliance rates and implementation in practice. Where compliance is lower without community acceptance, impacts in practice are expected to have lower effects and higher cost than in trial. Hence, modelling which extrapolates within-study analysis without allowing for compliance in practice will often represent a best-case scenario for the long-term effectiveness and cost-effectiveness of individual-based interventions. Modelling cost effectiveness with trial-based compliance is unbiased where practice compliance matches trial compliance, but for trial compliance to be replicated requires greater engagement than that usually observed in practice.

Conversely, within-study analysis will usually represent a worst-case scenario for the expected long-term costs, effectiveness and cost-effectiveness of effective community-based health promotion strategies where they have community ownership. Provided there is community ownership of programs, widening and continuing program effects are expected from network impacts in spreading the use of strategies over time and across widening populations, while not necessarily incurring any further implementation costs.

Hence, unlike patient-centred interventions, community interventions have the potential to costlessly widen impact of effective strategies where effective. This can be both in terms of the number of people benefiting across communities and the time over which these populations benefit. Long-term effects with no or minimal further implementation cost are expected if social capital has been built with community acceptance and ownership of the strategies and their use and impacts are sustained beyond the study period.

Consequently, multiplier effects from program investment flowing across networks into community activities provide a robust quantitative indicator of community ownership, engagement with, and building of, social networks and capital and sustainability of programs over time (Hawe et al. 2009; Shiell et al. 2008). Alongside (triangulated with) qualitative evidence, this enables informed assessment of whether typically short-term program effects on individuals during evaluation periods can be expected to translate into sustainable programs with long-term effects across communities. This in turn provides the key evidence to assess whether health promotion or prevention programs are expected to be successful and cost-effective.

### 4.2.1 The Stephanie Alexander Evaluation Case Study

The Stephanie Alexander Kitchen Garden National Program (SAKGNP) has been provided as an Australian primary school-based health promotion and disease prevention strategy since 2008, nationally extending a state-piloted Victorian program which started in 2001. The program funds garden and kitchen capital (up to A$60,000), with an obligation of successful schools to weekly run Stephanie Alexander garden (45 min) and kitchen (90 min) classes for year 3–6 and students for 2 years. The SAKGNP was evaluated between 2011 and 2012 applying mixed qualitative and quantitative methods (Yeatman et al. 2013, 2014) including multiplier methods for economic evaluation (Eckermann et al. 2014).

The overall aim was to triangulate evidence of attributable short-term impacts on student attitudes, lifestyle and behaviour with the community (government, school and wider community) multiplier impacts of initial investment on the value of total activity up to and beyond 2 years along with qualitative evidence of community acceptance to enable informed assessment of expected long-term community impacts and returns from SAKGNP investment.

In assessing short-term effectiveness on individuals in the immediate target population of schoolchildren, the evaluation triangulated multiple evidence sources for signals of four policy-relevant domains:

(i) Kitchen lifestyle and attitudes;
(ii) Garden lifestyle and attitudes;
(iii) Eating habits and social behaviours; and
(iv) Food Choices.

Comparative student and parent evidence of domain impacts in students from 28 initiative and 14 matched (state curriculum, school size, rurality, index of community socio-educational advantage) schools were compared adjusted for potential confounders (student grade, time at school, gender; parent level of education, country of birth) and triangulated with pre-post analysis in initiative schools.

Long-term effectiveness and cost effectiveness of the health promotion and prevention program the SAKGNP represents were assessed as critically depending on:

(i) Acceptance, implementation and local ownership by school and community networks – quantitatively measured by the multiplier of grant investment on school and community activity over time; and
(ii) Continuation and program adaptation and integration – garden and kitchen classes continuing beyond a 2-year agreement period, changing in program scale and/or scope and integration with school curriculum.

In terms of short-term individual impacts, students in SAKGNP schools ($n = 491$)showed improved student food choices ($p = 0.024$) and kitchen lifestyle behaviour ($p = 0.019$) domains compared to controls ($n = 260$). This was triangulated with intervention pre-post analysis where 20.0% (58/290) reported eating fruit and vegetables more often and 18.6% (54/290) preparing food at home more often.

No statistically significant differences were found in case-control analysis for eating habits or garden lifestyle behaviour domains, although 32.3% of children helped more in the garden (91/278) and 15.6% (45/289) ate meals together more often in pre-post analysis. Findings of the SAKGNP multiplier impact up to 2 years and decomposition (Eckermann et al. 2014; Yeatman et al. 2013, 2014) are summarised in Table 4.1.

An average grant contribution of $44,758 generated a value of observable community activity (Salamon et al. 2011) at 2 years totalling $226,737 and hence a multiplier of 5.07. In considering impacts beyond 2 years, all eight schools observed beyond the 2-year obligation period continued garden and kitchen classes with an average 17% upscaling of classes (371–436 class hours per year) and local adaptation, including full curriculum integration.

Importantly, in considering full curriculum integration while kitchen and garden classes were found to be easily integrated into broader primary school curriculum, with the average cost of running kitchen and garden classes and preparation to teaching time of classes was found to be equivalent to or less than that of the curriculum more generally.

The ratio for specialist kitchen and garden staff total hours to contact hours in the garden and kitchen classes and class costs (Table 4.2) was comparable to the mini-

**Table 4.1** Multiplier impacts of investment in the SAKGNP

| Activity (hours) and value (AU$) | Total | Garden | Kitchen |
|---|---|---|---|
| Average hours of community donated time per school up to 2 years | 2641.1 | 1351.7 | 1289.4 |
| Value of community volunteer time[a] | $99,238.49 | $48,029.24 | $51,209.24 |
| Donated capital | $11,252.44 | $6991.70 | $4260.75 |
| Total community contribution up to 2 years | $110,490.93 | $55,020.94 | $55,469.99 |
| Total school contributions | $71,488.41 | $27,553.17 | $43,935.24 |
| Total grant contributions | $44,757.67 | $15,147.40 | $29,610.26 |
| Total | $226,737.01 | $97,721.51 | $129,015.50 |
| Multiplier | 5.07 (= $226,737 / $44,758) | | |

[a]Volunteer time for multiplier valued at same rate as classes kitchen or garden specialist

**Table 4.2** SAKGNP kitchen and garden classes and staff specialist and co-ordinator class costs per hour

| Comparing specialist staff costs (AU$) | Total | Garden | Kitchen |
|---|---|---|---|
| Average class hours per school to 2 years | 787.4 | 307.0 | 480.4 |
| Per year average class hours | 409.40 | 159.22 | 250.18 |
| Specialist staff hours per hour of classes | | 1.778 | 1.542 |
| Staff specialist cost per class | | $63.18 | $61.26 |
| SAKGNP staff costs per class including program co-ordinator costs | | $72.19 | $70.00 |

mum of those typically found in conventional contemporary Australian primary school classroom settings (ABS 2006).

Overall, triangulated policy-relevant assessment of short-term individual impacts and 2-year and longer-term multiplier and acceptance impact of SAKGNP capital investment found:

(i) Improved food choices ($p = 0.02$) and kitchen lifestyle behaviour ($p = 0.02$) of individuals in case-control and pre-post analysis.
(ii) Multiplier on Commonwealth investment of 5.07-fold ($226,737/$44,758) at 2 years, 1.60 attributable to school and 2.47 to wider community activity.
(iii) Beyond 2 years SAKGNP classes were upscaled (average 17%), and local adaptation in schools included full curriculum integration as well as greater emphasis on kitchen classes and local ingredients and cultural dishes, indicating strong long-term community ownership.

Combined, these results provided strong triangulated evidence of the effectiveness of SAKGNP as a health promotion and prevention program in changing student attitudes and behaviour, engaging with school and community networks and building social capital to enable cultural embedding, community ownership and local adaptation of the SAKGNP strategies (Eckermann et al. 2014). These findings were also supported by prior studies showing positive impacts on attitudes, skills, lifestyle and eating behaviours of community gardens particularly in young, as well as older, community populations (Libman 2007; Litt et al. 2011; Robinson-O'Brien et al. 2009; Morris et al. 2001; Morgan et al. 2010; Somerset and Markwell 2008). The Australian Government committed further funding of $5.4 million – providing opportunities for 400 new schools, bringing total number of schools to 650.

Measuring multiplier impacts over time and particularly beyond obligation periods provides a good first step in assessing whether community health promotion programs have community ownership and network impacts. Never the less the use of more sophisticated network analysis methods, such as pre- and post-network density (number of linkages per individual) or the extent of champion and gatekeeper involvement, Hawe et al. (Hawe et al. 2004a, b, 2009; Hawe and Ghali 2008) could further strengthen methods and findings. For example, in the case of the SAKGNP, such network analysis could have been undertaken across specialist garden and kitchen staff, curriculum integrated and wider school staff and volunteers, students and parents. Qualitatively, the continuity of kitchen and garden staff and particularly their relationship to volunteers were highlighted as important for maintaining volunteer involvement, while more generally curriculum integration was seen as key to program continuation.

Research methods could also link to impacts on wider community systems, networks and events, e.g. building of other community gardens and kitchens that SAKGNP may have enabled. Nevertheless, investment multipliers monitored over time combined with qualitative evidence provide a strong platform to both triangulate with and direct the best use of such more sophisticated methods.

### 4.2.2 *Conclusion: Health Promotion and Prevention in Complex Community Settings*

Evaluating multiplier effects over time (triangulated with evaluation of community ownership of interventions) is key to assessing long-term effectiveness and cost effectiveness of health promotion and primary prevention programs in complex community settings. Such extrapolated assessment is also key to allow fair comparison with individual-focused treatment interventions.

Effective health promotion programs with community ownership and network multiplier impacts provide the potential to costlessly widen benefits over time and across community networks.

This compared with individual-targeted interventions where compliance issues, side effects, time profile of resistance and intolerance, etc. reduce LT effects, increase cost and reduce cost effectiveness in practice from that in trial setting.

## 4.3 Palliative Care

Economic evaluations in palliative settings are extremely rare (Smith et al. 2014). Many authors have highlighted the challenges of conducting economic evaluations in this context, particularly capturing and valuing benefits (Normand 1996; Bruera and Suarez-Almazor 1998; Higginson and Edmonds 1999; Hughes et al. 1999; McCaffrey and Currow 2015). Palliative care is a complex disease area; aetiology, symptoms, treatment and needs are multi-faceted requiring a broad range of services provided by diverse disciplines across all health-care sectors. According to the World Health Organisation, palliative care:

> …improves the quality of life of patients and families facing the problem associated with life-threatening illness, through the prevention and relief of suffering by means of early identification and impeccable assessment and treatment of pain and other problems, physical, psychosocial and spiritual (Australian Institute of Health and Welfare 2013).

The demand for palliative care is expected to continue rising rapidly, particularly with baby boomer ageing. Although the incidences of the major causes of death have been declining in Australia in the 33- to 77-year age groups particularly (ABS 2013), cancer is bucking the trend while older age group mortality rates have moved little in comparison (Productivity Commission 2006). Also, the changing pattern of death means, as individuals, we are much more likely to survive into older age with advanced notification of our impending death and therefore to require palliative care (Najman 2000; Davies and Higginson 2004). Consequently, there is a growing imperative for systematic comparison of costs and benefits of palliative care interventions to inform health-care decision making. It is crucial that the cost and consequences of palliative care interventions are robustly evaluated to usefully inform health policy decision making.

Palliative care is multidimensional, evident from the WHO definition phrases 'pain and other problems, physical, psychosocial and spiritual' and 'quality of life' (QOL) (Australian Institute of Health and Welfare 2013). There is utility from the dying process: preferences for place of care and place of death between inpatient- or home-based services, preparation for death and discussions about death and treatment modalities, e.g. avoiding mechanic ventilation (Higginson and Sen-Gupta 2000, Patrick et al. 2001; Steinhauser et al. 2001; Agar et al. 2008). Not only the patients' but also the families' QOL is an important dimension of palliative care. Evidence in the scientific literature suggests that palliative care interventions affect patients' and informal carers' health-related QOL (Hughes et al. 1999; Christakis and Allison 2006; Dixon et al. 2006; Carretero et al. 2009; Davidson et al. 2008; McNamee and Seymour 2008), and family members commonly act as informal carers for loved ones at the end of life. In Australia, approximately one in ten people in the community has provided hands-on care for someone at the end of life in the last 5 years, with estimated annual contributions by informal caregivers of AUD$40 billion (McCaffrey et al. 2016a). However, one study estimated that a quarter of people that have cared for a loved one at the end of life could not commit to taking on the role again (Currow et al. 2011). Given societies increasing and implicit reliance on carers' willingness to take on, maintain and repeat this role and the economic consequences should such care diminish, it is crucial that carer effects are considered in economic evaluations (McCaffrey and Currow 2010; McCaffrey et al. 2015b).

In summary, there are multiple important outcomes in palliative care service delivery and economic methods that incorporate such multiple effects that are required to robustly evaluate the relative costs and benefits of palliative care interventions.

### *4.3.1 Economic Evaluation in Palliative Care*

Conventional cost-effectiveness analysis (CEA) is limited to consideration of one measure of effect such as life years gained (Drummond et al. 2005). However, when multiple outcome domains are important, such single outcome comparison can lead to conflicting conclusions concerning preferred strategies (Sindelar et al. 2004; Al-Janabi et al. 2013; McCaffrey et al. 2013). Consequently, decisions about the costs and benefits of funding allocations can be misinformed and lead to inefficient distribution of finite health-care resources (Sindelar et al. 2004; Hoch and Dewa 2007). The widely applied QALY measure overcomes this limitation to the extent that impacts on multiple domains of health can be integrated with survival (Drummond et al. 2005). However, QALYs may lack sensitivity for people receiving palliation because important aspects of QOL such as preparation for death (finances, funeral arrangements, wills, handing over tasks to other family members, advance directives, saying goodbye) (McCaffrey et al. 2009, 2014, 2016b) and existential issues like hope and dignity are not amenable to integration with individual survival. These domains generally remain unconsidered by the popular generic

multi-attribute utility instruments (MAUIs) measuring the QOL component of the QALY (see Table 4.1). The one notable exception is research led by Coast (Coast et al. 2012) devising a supportive care measure (ICECAP-SCM) for use in economic evaluations of palliative and end-of-life care. Attributes in the ICECAP-SCM developed from interviews with samples of older people, the general population, aged care residents and patients receiving palliative care include autonomy, love, physical suffering, emotional suffering, dignity, support and preparation, in contrast to the domains/dimensions listed in Table 4.3.

Quality-adjusted life years focus on health alone as the sole indicator of value. Utility from the dying process is not captured, and the way care is delivered may be just as or more important to patients' QOL than the health outcomes achieved in the palliative setting (Ratcliffe and Buxton 1999). For example, many patients obtain value from receiving home based rather than hospital care (Agar et al. 2008). Such aspects of care are not captured by existing generic multi-attribute utility instruments (MAUIs) (see Table 4.3). Hence, evaluations of palliative care alternatives using QALYs calculated by integrating utility weights generated from generic MAUIs such as the EQ-5D, with survival time, will not cover these key patient-valued domains. Indeed, patient health-related quality of life (HRQOL) utility weights generated by such measures and their integration with survival have been shown to be at best third- and fourth-order considerations of palliative patients trumped by process of death in relation to ability to finalise personal and financial affairs and family and carer distress (Steinhauser et al. 2000; Finkelstein et al. 2015; Malhotra et al. 2015). Relying solely on patient QALYs generated from generic MAUIs which fail to reflect the key domains important to palliative care patients risks inefficient and inequitable allocation of scarce public funds. The risks are both in supporting interventions and strategies that do not satisfy palliative patient and societal primary needs and preferences and failure to support those that do.

Further, recent research suggests that general public preferences in palliative care settings, for those close to death, are for such populations preferences to be respected and reflected in resource allocations (Pinto-Prades et al. 2014; Finkelstein et al. 2015; McHugh et al. 2015; van Exel et al. 2015). However, these public preferences should not be misinterpreted as a call for higher threshold values for QALYs in palliative populations. Rather, efforts should be made to improve the coverage of domains important to palliative populations in health economic analysis to better inform societal decision making (McCaffrey et al. 2016b). Given societal perspectives, there is an urgent need to develop methods which reflect palliative

**Table 4.3** Comparison of domains for the pre-scored generic multi-attribute utility-based questionnaires EQ-5D, SF-6D, HUI3 and AQOL (Marra et al. 2005; Richardson et al. 2011)

| Questionnaire | Domains/dimensions |
|---|---|
| EQ-5D | Anxiety; pain; mobility; self-care; usual activities |
| SF-6D | Mental health; pain; physical function; role limitation; social function; vitality |
| HUI3 | Ambulation; cognition; dexterity; emotion; hearing; pain; speech; vision |
| AQoL | Coping; independent living; life satisfaction; mental health; pain; relationship; self-worth; senses |

patient preferences for multiple domains in resource allocation decisions. Chapter 10 illustrates some promising methods for how this might be achieved to increase utility of palliative care cost effectiveness evaluation, decision making and investment.

Furthermore, the options to increase palliative patient utility across multiple domains have the distinct potential (see Chap. 10 and Sect. 12.4) to be achieved without increasing health system costs. Options that allow patients to finalise their affairs and have a good process of death with loved ones at home are often not the most expensive. For example, in a typical palliative population with cancer, while very expensive options such as radiotherapy and chemotherapy and a plethora of associated lines of therapy of medication and regimens have been developed which arguably may marginally increase survival and potentially QALYs (Cardona-Morrell et al. 2016; Kamal et al. 2016), they usually require extensive treatment in hospital settings with significant side effects. Hence such therapies often deny satisfying palliative care patients the ability to finalise their affairs at home with loved ones – combining key domains of importance to palliative patients. Conversely, for example, the use of pain relief such as medicinal cannabis (Johnson et al. 2010; Carter et al. 2011) can often be managed within a home setting by patients and loved ones with palliative physician oversight and intervention as necessary. Such strategies better target and create the ability to address the primary domains of importance to palliative patients, while also having significant potential to be cheaper than alternative current therapies.

Most importantly for this chapter, what the above discussion highlights is the importance of having robust methods to allow for joint outcome domains of interest and importance in palliative care settings and ideally consideration alongside costs to enable palliative patient and societal optimisation of resource (budget) allocations across alternatives.

One possible option is to develop a condition-specific MAUI from an existing palliative care QOL instrument to generate quality weights for economic evaluations (Brazier et al. 2012). However, there is no gold standard for measuring QOL in the palliative setting and no single instrument includes all relevant aspects (Albers et al. 2010). For example, topics such as existential issues are missing from the EORTC QLQ-C15-PAL (Groenvold et al. 2006), one of the most popular palliative care QOL questionnaires, and a key aspect of QOL, the ability to finalise personal and financial affairs in preparation for death, is not explicitly captured by this nor other commonly used instruments (McCaffrey et al. 2009, 2014, 2016b). In a recent systematic review of palliative care QOL instruments, the McGill Quality of Life Questionnaire (MQOL) (Albers et al. 2010) was ranked highest with the best measurement properties, but aspects such as quality of care, place of care and financial issues are missing from this tool (Shahidi et al. 2010). Two potentially related cancer-specific MAUIs being developed for use in economic evaluations are the QLQ-U (King et al. 2016) and FACT-U (Costa et al. 2014). Once again though, these instruments are hampered by the limited coverage of QOL dimensions important in the palliative setting such as spiritual issues, and thus these instruments may miss the issues of most concern to this population (Shahidi et al. 2010).

Alternatively, cost-consequences analysis (CCA) (Drummond et al. 2005) with disaggregated mean costs and multiple outcomes explicitly presented has been advocated as a preferred method in these situations (Sculpher and Price 2003; Coast 2004). However, in CCA, cost and multiple outcomes are treated separately without consideration of their interaction or joint uncertainty. As Briggs et al. argued in their seminal papers on the death of cost minimisation (Briggs and O'Brien 2001) and cost effectiveness under uncertainty (Briggs et al. 2002), separate and sequential hypothesis tests on differences in outcomes and costs lead to fallacious inferences. It is important that CEA represents the joint uncertainty associated with cost and key effect domains so that funders and policy makers can assess the degree of confidence in CEA and associated decision making. Evaluation should enable robust trade-offs between expected relative costs and effects under uncertainty to aid consideration of the consequences of the funding decision and contribute to decisions concerning the value of obtaining additional information from further research (Eckermann and Willan 2007; Claxton 2008; Eckermann et al. 2010), as considered in Chaps. 5 to 7.

Ultimately, economic evaluations will misrepresent the relative net benefit of palliative and end-of-life care alternatives without simultaneous consideration of costs and multiple outcome domains under uncertainty. Accessible, robust and generalisable methods for jointly comparing cost and multiple outcomes under uncertainty are needed to better inform funding decisions in such settings, which we explicitly address following McCaffrey et al. (2015a) in Chap. 10.

## References

Agar M, Currow DC, Shelby-James TM, Plummer J, Sanderson C, Abernethy AP. Preference for place of care and place of death in palliative care: are these different questions? Palliat Med. 2008;22(7):787–95.

Albers G, Echteld MA, de Vet HC, Onwuteaka-Philipsen BD, van der Linden MH, Deliens L. Evaluation of quality-of-life measures for use in palliative care: a systematic review. Palliat Med. 2010;24(1):17–37.

Al-Janabi H, McCaffrey N, Ratcliffe J. Carer preferences in economic evaluation and healthcare decision making. Patient Patient Centered Outcomes Res. 2013;6(4):235–9.

Australian Bureau of Statistics. Education & work: school teachers. 2006. Retrieved 1 Apr 2013, from http://www.abs.gov.au/ausstats/abs@.nsf/2f762f95845417aeca25706c00834efa/557e80a2b6d29244ca2570ec001b193d!OpenDocument.

Australian Institute of Health and Welfare. Palliative care services in Australia 2013. HWI 123 Canberra. AIHW. 2013.

Brazier JE, Rowen D, Mavranezouli I, Tsuchiya A, Young T, Yang Y, Barkham M, Ibbotson R. Developing and testing methods for deriving preference-based measures of health from condition-specific measures (and other patient-based measures of outcome). Health Technol Assess. 2012;16(32):1–114.

Briggs A, O'Brien B. The death of cost-minimization analysis? Health Econ. 2001;10(2):179–84.

Briggs AH, O'Brien BJ, Blackhouse G. Thinking outside the box: recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. Annu Rev Public Health. 2002;23(1):377–401.

Bruera E, Suarez-Almazor M. Cost effectiveness in palliative care. Palliat Med. 1998;12:315–6.

Cardona-Morrell M, Kim J, Turner R, Anstey M, Mitchell I, Hillman K. Non-beneficial treatments in hospital at the end of life: a systematic review on extent of the problem. Int J Qual Health Care. 2016;28(4):456–69.
Carretero S, Garces J, Rodenas F, Sanjose V. The informal caregiver's burden of dependent people: theory and empirical review. Arch Gerontol Geriatr. 2009;49(1):74–9.
Carter GT, Flanagan AM, Earleywine M, Abrams DI, Aggarwal SK, Grinspoon L. Cannabis in palliative medicine: improving care and reducing opioid-related morbidity. Am J Hosp Palliat Care. 2011;28(5):297–303.
Christakis N, Allison P. Mortality after the hospitalization of a spouse. N Engl J Med. 2006;354(7):719.
Claxton K. Exploring uncertainty in cost-effectiveness analysis. PharmacoEconomics. 2008;26(9):781–98.
Coast J. Is economic evaluation in touch with society's health values? BMJ. 2004;329(7476):1233–6.
Coast J, Al-Janabi H, Sutton EJ, Horrocks SA, Vosper AJ, Swancutt DR, Flynn TN. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. Health Econ. 2012;21(6):730–41.
Costa DSJ, King MT, Aaronson NK, Brazier JE, Cella D, Grimison P, O'Brien C, Janda M, McTaggart-Cowan H, Pickard AS, Arbuckle R. The development of cancer-specific multiattribute utility instruments from the EORTC QLQ-C30 and FACT-G. Qual Life Res. 2014;23:28–9.
Currow DC, Burns C, Agar M, Phillips J, McCaffrey N, Abernethy AP. Palliative caregivers who would not take on the caring role again. J Pain Symptom Manag. 2011;41(4):661–72.
Davidson T, Krevers B, Levin L-Å. In pursuit of QALY weights for relatives: empirical estimates in relatives caring for older people. Eur J Health Econ. 2008;9(3):285–92.
Davies E, Higginson IJ. Palliative care: the solid facts. Copenhagan: World Health Organisation; 2004.
Dawber TR, Moore FEJ, Mann GV. Coronary heart disease in the Framingham Study. Am J Public Health. 1957;47(2):4–25.
de Salazar L, Jackson S, Shiell A, Rice M. Guide to economic evaluation in health promotion. Washington: Pan American Health Organization, Regional Office of the World Health Organization; 2007.
Dixon S, Walker M, Salek S. Incorporating carer effects into economic evaluation. PharmacoEconomics. 2006;24(1):43–53.
Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. Methods for the economic evaluation of health care programmes. Oxford, UK: Oxford University Press; 2005.
Eakin E, Lawler S, Vandelanotte C, Owen N. Telephone interventions for physical activity and dietary behaviour change: A Systematic Review. Am J Prev Med. 2007;32(5):419–34.
Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16(2):195–209.
Eckermann S, Karnon J, Willan AR. The value of value of information: best informing research design and prioritization using current methods. PharmacoEconomics. 2010;28(9):699–709.
Eckermann S, Dawber J, Yateman H, Quinsey K, Morris D. Evaluating return on investment in a school based health promotion and prevention program: the investment multiplier for the Stephanie Alexander Kitchen Garden National Program. Soc Sci Med. 2014;114:103–12. doi:10.1016/j.socscimed.2014.05.056.
Finkelstein EA, Bilger M, Flynn TN, Malhotra C. Preferences for end-of-life care among community-dwelling older adults and patients with advanced cancer: a discrete choice experiment. Health Policy. 2015;119(11):1482–9.
Groenvold M, Petersen MA, Aaronson NK, Arraras JI, Blazeby JM, Bottomley A, Fayers PM, de Graeff A, Hammerlid E, Kaasa S, Sprangers MA, Bjorner JB. The development of the EORTC QLQ-C15-PAL: a shortened questionnaire for cancer patients in palliative care. Eur J Cancer. 2006;42(1):55–64.

Hawe P, Ghali L. Use of social network analysis to map the social relationships of staff and teachers at school. Health Educ Res. 2008;23(1):62–9.
Hawe P, Shiell A. Social capital and health promotion: a review. Soc Sci Med. 2000;51(6):871–85.
Hawe P, Shiell A, Riley T. Complex interventions: how far 'out of control' should a randomised controlled trial be? Br Med J. 2004a;328:1561–3.
Hawe P, Shiell A, Riley T, Gold L. Methods for exploring implementation variation and local context within a cluster randomised trial. J Epidemiol Community Health. 2004b;58:788–93.
Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. Am J Community Psychol. 2009;43(3–4):267–76.
Higginson IJ, Edmonds P. Services, costs and appropriate outcomes in end of life care. Ann Oncol. 1999;10:135–6.
Higginson IJ, Sen-Gupta GJA. Place of care in advanced cancer: a qualitative systematic literature review of patient preferences. J Palliat Med. 2000;3(3):287–300.
Hill SE, Blakely TA, Fawcett JM, Howden-Chaman P. Could mainstream anti-smoking programs increase inequalities in tobacco use? New Zealand data from 1981-96. Aust N Z J Public Health. 2005;29(3):279–88.
Hoch JS, Dewa CS. Lessons from trial-based cost-effectiveness analyses of mental health interventions: why uncertainty about the outcome, estimate and willingness to pay matters. PharmacoEconomics. 2007;25(10):807–16.
Hughes SL, Giobbie-Hurder A, Weaver FM, Kubal JD, Henderson W. Relationship between caregiver burden and health-related quality of life. Gerontologist. 1999;39(5):534–45.
Johnson JR, Burnell-Nugent M, Lossignol D, Ganae-Motan ED, Potts R, Fallon MT. Multicenter, double-blind, randomized, placebo-controlled, parallel-group study of the efficacy, safety, and tolerability of THC:CBD extract and THC extract in patients with intractable cancer-related pain. J Pain Symptom Manag. 2010;39(2):167–79.
Kamal AH, LeBlanc TW, Meier DE. Better palliative care for all: improving the lived experience with cancer. JAMA. 2016;316(1):29–30.
King, M. T., D. S. J. Costa, N. K. Aaronson, J. E. Brazier, D. F. Cella, P. M. Fayers, P. Grimison, M. Janda, G. Kemmler, R. Norman, A. S. Pickard, D. Rowen, G. Velikova, T. A. Young and R. Viney (2016). QLU-C10D: a health state classification system for a multi-attribute utility measure based on the EORTC QLQ-C30. Qual Life Res. 2016;25(3):625–36.
Libman K. Growing youth growing food: how vegetable gardening influences young people's food consciousness and eating habits. Appl Environ Educ Commun. 2007;6(1):87–95.
Litt JS, Soobader M-J, Turbin MS, Hale JW, Buchenau M, Marshall JA. The influence of social involvement, neighborhood aesthetics, and community garden participation on fruit and vegetable consumption. Am J Public Health. 2011;101(8):1466–73.
Malhotra C, Farooqui MA, Kanesvaran R, Bilger M, Finkelstein E. Comparison of preferences for end-of-life care among patients with advanced cancer and their caregivers: a discrete choice experiment. Palliat Med. 2015;29(9):842–50.
Marra CA, Woolcott JC, Kopec JA, Shojania K, Offer R, Brazier JE, Esdaile JM, Anis AH. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. Soc Sci Med. 2005;60(7):1571–82.
McCaffrey N. Modelling joint cost and outcomes uncertainty on the cost-disutility plane – case studies in palliative care. PhD, Flinders University, South Australia; 2013.
McCaffrey N, Currow D. Including carer effects in economic evaluations: willingness to care, a missing link? Palliat Med. 2010;24(4):s166–7.
McCaffrey N, Currow D. Separated at birth? BMJ Support Palliat Care. 2015;5(1):2–3.
McCaffrey N, Currow DC, Eckermann S. Measuring impacts of value to patients is crucial when evaluating palliative care. J Pain Symptom Manag. 2009;37(6):e7–9.
McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Is home-based palliative care cost-effective? An economic evaluation of the Palliative Care Extended Packages at Home (PEACH) pilot. BMJ Support Palliat Care. 2013;3(4):431–5.

McCaffrey N, Skuza P, Breaden K, Eckermann S, Hardy J, Oaten S, Briffa M, Currow D. Preliminary development and validation of a new end-of-life patient-reported outcome measure assessing the ability of patients to finalise their affairs at the end of life. PLoS One. 2014;9(4):e94316.

McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Better informing decision making with multiple outcomes cost-effectiveness analysis under uncertainty in cost-disutility space. PLoS One. 2015a;10(3):e0115544.

McCaffrey N, Cassel JB, Coast J. Bringing the economic cost of informal caregiving into focus. Palliat Med. 2015b;29(10):866–7.

McCaffrey, N., H. Al-Janabi, D. Currow, R. Hoefman and J. Ratcliffe. Protocol for a systematic review of preference-based instruments for measuring care-related outcomes and their suitability for the palliative care setting. BMJ. 2016a;6(9):e012181.

McCaffrey N, Bradley S, Ratcliffe J, Currow DC. What aspects of quality of life are important from palliative care patients' perspectives? A systematic review of qualitative research. J Pain Symptom Manage. 2016b. pii S0885–39324(39316)30078–39321.

McHugh N, Baker RM, Mason H, Williamson L, van Exel J, Deogaonkar R, Collins M, Donaldson C. Extending life for people with a terminal illness: a moral right and an expensive death? Exploring societal perspectives. BMC Medical Ethics. 2015;16(1):1–15.

McNamee P, Seymour J. Incorporation of process preferences within the QALY framework: a study of alternative methods. Med Decis Mak. 2008;28(3):443–52.

Moore S, Haines V, Hawe P, Shiell A. Lost in translation: a genealogy of the "social capital" concept in public health. J Epidemiol Community Health. 2006;60(8):729–34.

Morgan PJ, Warren JM, Lubans DR, Saunders KL, Quick GI, Collins CE. The impact of nutrition education with and without a school garden on knowledge, vegetable intake and preferences and quality of school life among primary-school students. Public Health Nutr. 2010;13(11):1931–40.

Morris JL, Neustadter A, Zidenberg-Cherr S. First-grade gardeners more likely to taste vegetables. Calif Agric. 2001;55(1):43–6.

Najman J. The demography of death: patterns of Australian mortality. In: Kelliher A, editor. Death and dying in Australia. Melbourne: Oxford University Press; 2000.

Normand C. Economics and evaluation of palliative care. Palliat Med. 1996;10:3–4.

Patrick DL, Engelberg RA, Curtis JR. Evaluating the quality of dying and death. J Pain Symptom Manag. 2001;22(3):717–26.

Pinto-Prades J-L, Sánchez-Martínez F-I, Corbacho B, Baker R. Valuing QALYs at the end of life. Soc Sci Med. 2014;113:5–14.

Productivity Commission. Implications of an ageing Australia: cost of death and health expenditure. Productivity Commission Technical paper 13. Canberra; 2006.

Ratcliffe J, Buxton M. Patients' preferences regarding the process and outcomes of life-saving technology. An application of conjoint analysis to liver transplantation. Int J Technol Assess Health Care. 1999;15(2):340–51.

Richardson J, McKie J, Bariola E. Review and critique of health related multi attribute instruments. Clayton: Monash University; 2011.

Robinson-O'Brien R, Story M, Heim S. Impact of garden-based youth nutrition intervention programs: a review. J Am Diet Assoc. 2009;109(2):273–80.

Salamon LM, Sokolowski SW, Haddock MA. Measuring the economic value of volunteer work globally: concepts, estimates, and a roadmap to the future. Ann Public Cooperative Econ. 2011;82(3):217–52.

Schensul JJ. Sustainability in HIV prevention research. In: Trickett EJ, Pequegnat W, editors. Community interventions and AIDS. Oxford: Oxford University Press; 2005.

Sculpher MJ, Price M. Measuring costs and consequences in economic evaluation in asthma. Respir Med. 2003;97(5):508–20.

Shahidi J, Bernier N, Cohen SR. Quality of life in terminally ill cancer patients: contributors and content validity of instruments. J Palliat Care. 2010;26(2):88–93.

Shiell A, Hawe P. Health promotion community development and the tyranny of individualism. Health Econ. 1995;5(3):241–7.

Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. Br Med J. 2008;336(7656):1281–3.

Sindelar JL, Jofre-Bonet M, French MT, McLellan AT. Cost-effectiveness analysis of addiction treatment: paradoxes of multiple outcomes. Drug Alcohol Depend. 2004;73(1):41–50.

Smith S, Brick A, O'Hara S, Normand C. Evidence on the cost and cost-effectiveness of palliative care: a literature review. Palliat Med. 2014;28(2):130–50.

Somerset S, Markwell K. Impact of a school-based food garden on attitudes and identification skills regarding vegetables and fruit: a 12 month intervention trial. Public Health Nutr. 2008;12(2):214–22.

Steinhauser KE, Christakis NA, Clipp EC, McNeilly M, McIntyre L, Tulsky JA. Factors considered important at the end of life by patients, family, physicians, and other care providers. JAMA. 2000;284(19):2476–82.

Truett J, Cornfield J, Kannel WB. A multivariate analysis of the risk of coronary heart disease in Framingham. J Chronic Dis. 1967;20:511e524.

van Exel J, Baker R, Mason H, Donaldson C, Brouwer W. Public views on principles for health care priority setting: findings of a European cross-country study using Q methodology. Soc Sci Med. 2015;126(0):128–37.

Yeatman H, Quinsey K, Dawber J, Nielsen W, Condon-Paoloni D, Eckermann S, Morris D, Grootemaat P and Fildes D. Stephanie Alexander Kitchen Garden National Program Evaluation: Final Report. Centre for Health Service Development, Australian Health Services Research Institute, University of Wollongong, Wollongong, Australia. 2013. Available at: https://www.kitchengardenfoundation.org.au/sites/default/files/food%20education/sakgnp_evaluation_uow_finalreport_2012.pdf

Yeatman H, Quinsey K, Dawber J, Nielsen W, Condon-Paoloni D, Eckermann S, Morris D, Grootemaat P, Fildes D. Combining realism with rigour: evaluation of a National Kitchen Garden program in Australia primary schools. Eval J Australas. 2014;14(2):17–24.

Zaza S, Briss PA, Harris KW. The guide to community preventive service. In:What works to promote health? New York: Oxford University Press; 2005.

# Part II
# Joint Research and Reimbursement Questions, Optimising Local and Global Trial Design and Decision Making Under Uncertainty Within and Across Jurisdictions with Value of Information Methods

In Part II (Chaps. 5, 6 and 7) we highlight principles and methods that enable optimising joint research and reimbursement decisions under uncertainty with value of information methods for cases of interest where promising interventions or strategies have expected positive while uncertain INB. This naturally builds on robust coverage and comparability principles and methods for INB estimation, cost-effectiveness and return on investment assessment in societal decision-making evaluation, considering the expected performance of patient and community level interventions, strategies and programmes highlighted and illustrated in Part I (Chaps. 2, 3 and 4).

Values of information principles and methods are introduced that enable maximising expected value relative to cost or return on research design in making joint research and reimbursement decisions under uncertainty. That is, for cases of interest where promising strategies have positive while uncertain INB, whether it is optimal for decision makers to adopt with no trial (AN), delay and trial (DT) or adopt and trial (AT) where feasible; in optimising the expected value relative to cost of research (trial design) and reimbursement decisions. These principles and methods are considered in Chap. 5 for optimal joint reimbursement and research decision making in local jurisdictions (e.g. Australia, UK or Canada), noting that for interventions with prior positive while uncertain incremental net clinical benefit, adopting and trialling is usually infeasible locally, while feasible for new interventions or strategies where expected incremental net benefit is the result of health system cost savings. Chapter 6 extends these principles and methods to consider globally optimal societal decision maker's trials across jurisdictions, where the option for jurisdiction to adopt and translate trial evidence is shown to become generally feasible and be a valuable option as part of optimising the expected value relative to cost of global trial design and evidence translation. Chapter 7 further extends methods for optimal local and global trial design and joint research and reimbursement decisions in

Chaps. 5 and 6 to consider optimal pricing under uncertainty locally and globally, where potential arises for robust risk-sharing arrangements to mitigate against costs of reversal under uncertainty in jurisdictions who adopt and trial as part of a global trial.

While Chaps. 5 and 6 allow optimisation of societal decision maker funded trials, Chap. 7 considers decision making for manufacturer trials allowing for the interaction between societal decision maker and manufacturer perspectives in pricing and risk sharing under uncertainty (Willan and Eckermann 2012; Eckermann and Willan 2013). Overall Chaps. 5, 6 and 7 move us in the optimal decision-making cycle of Fig. p2.1 from assessment with positive while uncertain INB in identifying promising therapies or strategies to informing optimal joint research and reimbursement decision making in assessing:

(i) The expected value relative to expected cost, or expected net gain, of alternative joint research and reimbursement decisions to adopt now, delay and trial or where feasible adopt and trial, and identify associated optimal trial design (Chap. 5).
(ii) Globally optimal trial designs with explicit consideration of evidence translation and associated optimal local decision making to adopt and trial or delay and trial (Chap. 6).
(iii) Optimal pricing with the potential for robust risk sharing in jurisdiction who adopt and trial as part of a globally optimal trial, mitigating against potential impacts costs of reversal (Chap. 7).

Critically, in each case applying VOI principles and methods is shown to require allowing for key relevant decision contexts and their practical implications for optimisation in practice. In Chap. 5 key decision contexts highlighted include expected trial recruitment rate, follow-up and analysis time, option value and opportunity cost of delay, costs of reversal faced where adoption is feasible and expected degree of implementation conditional on strength of evidence (Eckermann et al. 2010; Eckermann and Willan 2007, 2008a, b; Willan and Eckermann 2010). In Chaps. 6 and 7, decision contexts are extended to global trial designs across jurisdictions and risk-sharing arrangements with pricing under uncertainty for jurisdiction where adopting and trialling becomes generally feasible and optimal with explicit evidence translation coverage across jurisdiction (Eckermann and Willan 2009, 2013; Willan and Eckermann 2012).

Chapters 6 and 7 highlight that globally optimal value of information trials across jurisdictions should be considered given such designs always improve on locally optimal value of information trials for cases of interest with positive while uncertain INB for societal decision making and manufacturer interests alike and better align these interests (Eckermann and Willan 2009, 2013). In particular, the option value of delay to societal decision makers and manufacturers can be avoided in jurisdictions with the feasible option to adopt and trial, while cost of reversal can be mitigated to the extent reversal can be avoided with the ability to undertake robust risk-sharing arrangements.

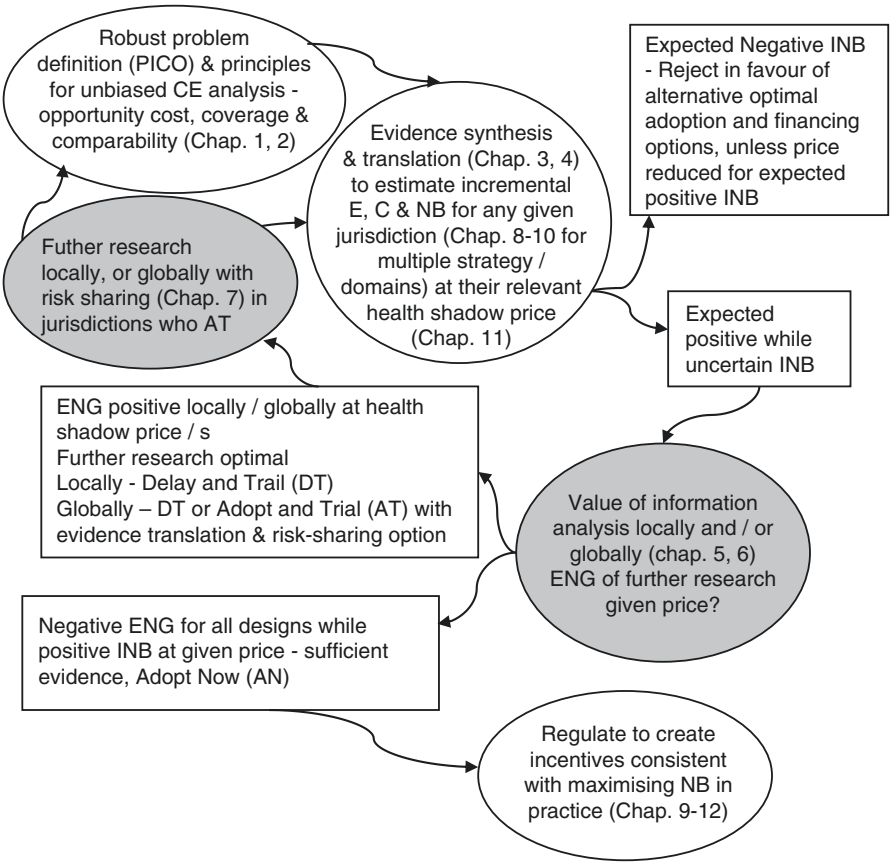


**Fig. p2.1** Optimal decision-making cycles for joint research, reimbursement and regulatory processes locally and globally

## References

Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16:195–209.

Eckermann S, Willan AR. Time and EVSI wait for no Patient. Value Health. 2008a;11:522–6.

Eckermann S, Willan AR. The option value of delay in health technology assessment. Med Decis Mak. 2008b;28:300–5.

Eckermann S, Willan AR. Globally optimal trial design for local decision making. Health Econ. 2009;18:203–16.

Eckermann S, Willan A. Optimal global VOI trials: better aligning manufacturer and decision maker interest and enabling feasible risk sharing. PharmacoEconomics. 2013;31:393–401.

Eckermann S, Karnon J, Willan A. The value of Value of Information: best informing research design and prioritization using current methods. PharmacoEconomics. 2010;28(9):699–709.

Willan AR, Eckermann S. Optimal clinical trial design using value of information methods with imperfect implementation. Health Econ. 2010;19(5):549–61.

Willan AR, Eckermann S. Expected value of information and pricing new health care interventions. PharmacoEconomics. 2012;30(6):447–59.

# Chapter 5
# The Value of Value of Information Methods to Decision-Making: What VOI Measures Enable Optimising Joint Research and Reimbursement Decisions Within a Jurisdiction?

## 5.1 Expected Value of Information Principles and Methods

Faced with making a decision under uncertainty given current evidence or of expected positive while uncertain INB, such as the distribution for incremental net benefit in Fig. 5.1, what is the expected value of additional information? Given costs of obtaining additional evidence or information, is it worth undertaking further research or not? If it is optimal to undertake further research, how much research is optimal?

In addressing the first of these questions, value of information methods clarify that the expected value of additional information is the expectation with further evidence of a reduction in decision uncertainty, and hence that expected opportunity losses arising from making decision without certainty (perfect information) reduce. That is, the expected value of further research or the expected value of sample information (EVSI) is the expectation with additional evidence of a reduction in the expected value of perfect information (EVPI) arising with the expected future rather than current INB uncertainty (Schlaiffer 1958, Raffia and Schlaiffer 1967, Claxton 1999).

To introduce and illustrate such value of information (VOI) concepts and measures in practice, consider a simple example with a game of pick a box. Suppose, as per Table 5.1, that there are five boxes and uncertainty about which box contains a $100,000 prize and which boxes are empty, other than box A, which with current information (analogous to data from a pilot trial) is known to be empty.

What is the expected value of having one or two additional boxes revealed? Without loss of generalisability, assume the prize is actually in box C, a fact the contestant is unaware of. Now, with current evidence, the contestant knows box A is empty while having no prior information in relation to the other 4 boxes. Hence, they have prior probabilities of 0.25 that the $100,000 is in each of the remaining 4 boxes (B to E), 3 of which are empty. Now, if they pick the box with the prize (box C) this has no opportunity loss. Conversely, if the contestant picks one of the remaining empty boxes (B, D or E), they have an opportunity loss of $100,000.



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_5

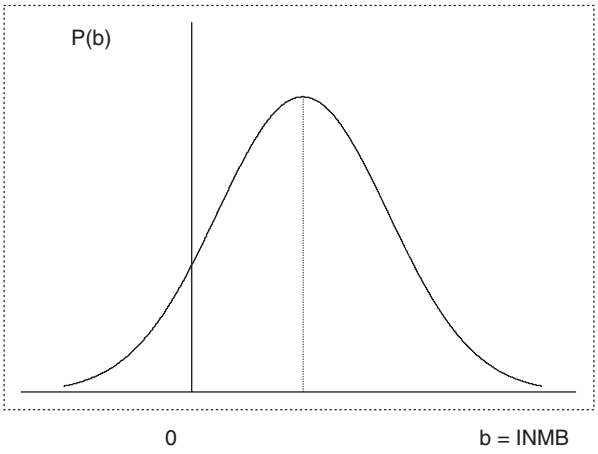


**Fig. 5.1** Cases of interest with positive while uncertain INB

**Table 5.1** Pick a box current EVPI with one empty box and EVSI of having 2 further empty boxes revealed

| Box | Content | Prior Pr | Opp. loss | Prior EOL | Post Pr | Post EOL |
|---|---|---|---|---|---|---|
| A | 0 | 0 | 0 | 0 | 0 | 0 |
| B | 0 | 0.25 | 100,000 | 25,000 | 0 | 0 |
| C | 100,000 | 0.25 | 0 | 0 | 0.5 | 0 |
| D | 0 | 0.25 | 100,000 | 25,000 | 0.5 | 50,000 |
| E | 0 | 0.25 | 100,000 | 25,000 | 0 | 0 |
| *EOL*[a] | | | | *75,000* | | *50,000* |

[a]*EOL* expected opportunity loss

Consequently, given their prior knowledge, the expected opportunity loss (EOL) avoided from eliminating uncertainty with perfect rather than current information (i.e. having 3 more boxes revealed) is $0.25 \times 0 + 0.75 \times 100{,}000 = \$75{,}000$. This expected opportunity loss of $75,000 represents the expected value of perfect information (EVPI) rather than current information. That is, the expected value of knowing all boxes contents with perfect information, rather than currently just the one box that doesn't contain the prize, is $75,000.

Now in reality, while uncertainty is expected to be reduced with additional evidence, it is not expected to be eliminated (analogously a trial does not provide perfect information). The expected value of additional (sample) information is the EVPI with current evidence less the expectation of EVPI with additional evidence.

Given current information of one empty box, if a second box were revealed as empty, then with two empty boxes revealed we would be picking between three boxes and hence the chance of picking the right box without any further prior knowledge is 1/3 and the wrong box is 2/3. Hence, the EOL or expectation of EVPI with a further box revealed (two boxes revealed rather than one box) reduces to

1/3 × 0 + 2/3 × 100,000 = $66,667, and the expected value of sample information (EVSI) is $75,000 – $66,667 = $8333.

Now, if two additional empty boxes were revealed, then the chance of picking the wrong or right box becomes 0.5, and the expected EOL or EVPI falls to 0.5 × $100,000 = $50,0000, as in Table 5.1. Consequently, the expected value of sample information (EVSI) from having 2 further boxes revealed given current knowledge of one empty box is $75,000 less $50,000 or $25,000:

$$\text{EVSI} = \text{EOL prior} - \text{E}\left(\text{EOL post}\right) = \text{EVPI prior} - \text{E}\left(\text{EVPI post}\right)$$

$$= 75,000 - 50,000 = 25,000$$

To address questions of whether it is worthwhile having further information (undertaking research) one needs to consider the expected value relative to expected cost of information. For example, consider a cost to revealing one additional box (2 empty boxes revealed in total) of $10,000, while $20,000 to reveal 2 additional boxes (3 empty boxes revealed in total). In that case, the expected value of sample information (EVSI) of having one additional empty box revealed of $8667 is $1333 less than the cost of $10,000. Hence the expected net gain (ENG) is $8667 – $10,000 = −$1333. This negative value of ENG indicates that at a cost of $10,000 it is not worth having 1 further box revealed. However, with two additional empty boxes revealed, the EVSI of $25,000 is $5000 greater than the expected total cost, and hence with poisitve ENG of $25,000 − $20,000 = $5000 is potentially worthwhile.

Nevertheless, note that if the decision-maker has other investment options, the opportunity cost of having two empty boxes revealed would be the expected value of the next best alternative action in investing $20,000, something we illustrate across potential research options in Sect. 5.3 and more generally in Chap. 11. An expected net gain (ENG) of $5000 from investing $20,000 represents an expected 25% return which could be compared with other rates of return in identifying optimal research and reimbursement decisions as Sect. 5.3 clarifies following Eckermann et al. (2010). Further, and more generally if, with a fixed or constrained budget, the $20,000 had to be raised from elsewhere, then both financing and adopting actions have opportunity costs which the health shadow price (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014) jointly considers, as highlighted in Chap. 11.

When facing current evidence of positive while uncertain incremental net benefit (INB) value of information methods can be applied to health economics analysis analogous to consideration of uncertainty about a prize in the ‘pick a box’ example. That is, in considering expected value, ENG (expected value relative to cost) and return (ENG relative to expected cost) from obtaining further information.

EVPI with uncertain INB is the expected value of losses avoided with perfect information, which reflects integration of the probability and associated losses where INB is less than 0 (Fig. 5.2). EVSI of any given research design represents the expected value of additional information from that research in reducing expected opportunity losses arising where INB is less than 0 (Fig. 5.3). That is, the EVSI of a given trial design (size) is EVPI with current information less the expectation of EVPI when information is updated with trial evidence.

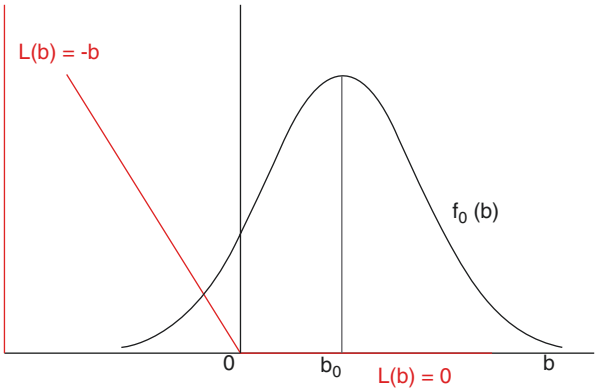


**Fig. 5.2** EVPI per patient integrating expected opportunity loss per patient with current rather than perfect information of the strategy maximising INB (Source: Eckermann and Willan (2007))

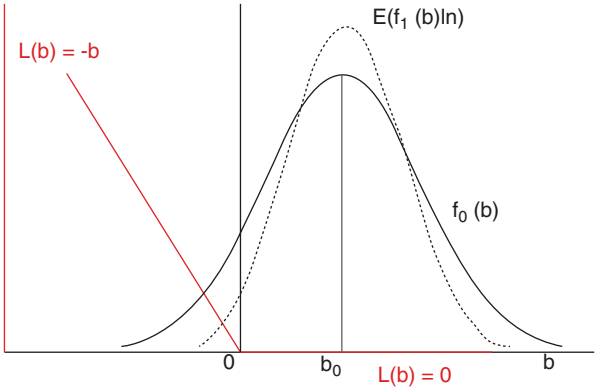


**Fig. 5.3** EVSI per patient with a trial of size *n* comparing two strategies as the expectation of an expected opportunity loss reduction (Source: Eckermann and Willan (2007))

The EVSI associated with a given research design can then be compared with the expected costs to decision-making of that design to estimate its expected net gain (ENG), the EVSI less expected costs. ENG (or return on research following Sect. 5.3) can in turn be optimised across potential trial sizes or research designs to identify optimal decision-making and trial design given current uncertainty.

In health-care evaluation, value of information (VOI) principles and methods have been proposed as a systematic decision-analytic approach for aiding optimal research and reimbursement decisions (Claxton 1999; Eckermann and Willan 2007; Eckermann et al. 2010). That is, in assessing whether there is sufficient evidence to adopt therapies, optimally designing research studies, and in efficiently allocating limited research funding across research proposals. The potential of VOI methods to

improve such decision-making is clear, given that, unlike conventional frequentist methods, their grounding in decision theory provides the ability to:

(i) Assess decision-maker uncertainty in relation to incremental costs, effects and cost effectiveness represented by the distribution of incremental net benefit (INB), rather than efficacy of a single clinical outcome (as per Chaps. 2–4).
(ii) Estimate expected value of research from integrating expected reductions in the probability of and negative payoffs from expected opportunity losses (negative regions of the INB distribution) under uncertainty from (i).
(iii) Efficiently design and prioritise research in maximising expected value less expected costs of, or return on, research rather than a hypothesis test with no necessary connection to expected value or cost of research.

To fulfil this potential, the use of the VOI toolkit needs to be both useful in addressing real decisions and simple enough to be applied by analysts and understood by decision-makers in practice. Considering the usefulness of VOI methods to inform decision-making, four related questions arise that these methods could be applied to:

(i) Is further research for a specific HTA potentially worthwhile?
(ii) Is the expected cost of a given research design less than its expected value?
(iii) What is the optimal research design for a specific HTA?
(iv) How is limited research funding best prioritised across alternative research proposal options?

## 5.2 Taking Occam's Razor to VOI Methods – What Is Necessary and Sufficient to Address Research Questions?

Following the principal of Occam's razor, the usefulness and value of alternative VOI measures in informing policy and decision-making in practice can be compared by considering their simplicity relative to their ability to inform decision-making. A method that is simpler to apply and equally or better informs decisions is clearly preferred under Occam's razor, while trade-offs between simplicity and informative value require further consideration. 'Better informed decision-making' may be interpreted in health economic terms as necessary and sufficient conditions to inform questions (i)–(iv) or, more generally, as the extent to which questions (i)–(iv) are answered or, conversely, losses arise from them not being fully addressed. Given limited resources for application of VOI methods, such assessment of simplicity and informative value can also be seen as akin to assessing the VOI equivalent of 'cost effectiveness of cost effectiveness'.

Following Eckermann et al. (2010) we first consider the ability of VOI measures (EVPI, EVSI, ENG) undertaken with different levels of sophistication, to inform decision-makers addressing research questions (i) to (iv).

When attempting to inform decision-making, VOI methods can be applied at different levels of sophistication to consider VOI measures including:

Expected value of perfect information (EVPI) given current decision uncertainty per patient and across the potential population expected to benefit.
Expected value of sample information (EVSI) of a given trial design (e.g. proposed frequentist trial design) per patient and across the population to benefit from trial information post trial analysis, population EVSI minus expected costs and, hence, expected net gain (ENG) of a given study design.
Optimal ENG or return on research investment across trial designs.

Considering which of these VOI measures are necessary and sufficient to answer questions (i) to (iv) is the primary issue in establishing which are useful to decision-making and point to how they can be best used in practice. In applying Occam's razor, we also later turn our attention to the simplicity and ability of alternative VOI methods to generate the necessary and sufficient VOI measures to inform questions ((i) to (iv)), allowing for key decision contexts.

### *5.2.1 Candidate Set 1: Per Patient and Population EVPI*

Current EVPI per patient, as the expected opportunity loss per patient avoided with perfect rather than current information, can be simply estimated at any given threshold value for a unit of effectiveness, given the current probability distribution of net benefit. With comparison of two strategies at a given threshold value, EVPI per patient is most easily graphically interpreted by considering the distribution of INB (noting that more than two strategies are shown as best considered with expected net loss frontiers in multiple strategy comparisons in Chap. 8). Figure 5.1 illustrated the case of interest to decision-makers, where current evidence suggests the new technology has expected positive but uncertain INB, with part of the distribution for b falling below 0.

EVPI per patient with current information is the expected value of the opportunity loss function integrated across the negative part of the distribution for INB per patient (b) as in Fig. 5.2. The value of the opportunity loss function is equal to zero where b is positive and the new therapy is optimal (there is no loss to choosing the net benefit maximising strategy), while where INB per patient (b) is negative, choosing the new therapy has an opportunity loss equal to the extent INB is negative (−b) per patient.

Hence, per-patient EVPI with two strategy comparisons reflects that with current information that there is both a:

(i) Chance that a decision to support such a new strategy with higher INB per patient will be wrong (P(INB < 0)); and an
(ii) Opportunity loss faced from being wrong to the extent INB is negative.

The expected opportunity loss (EOL) from adopting a strategy with positive expected while uncertainty INB based on current evidence is found integrating

opportunity losses across the probability density function (PDF) where INB is less than 0. This expected opportunity loss or expected value of perfect information (EVPI) per patient given current information represents a theoretical upper bound to the potential value of further information, that with perfect as opposed to current information.

Multiplying current EVPI per patient by the expected patient population who could potentially benefit from additional evidence over the relevant time horizon yields the current population EVPI. For example, if the current EVPI is Can$10 per patient and there were an expected one million future patients in a jurisdiction such as Canada across which information on the INB maximising strategy is valuable, then population EVPI equates to Can$10 multiplied by one million or Can$10 million. Such population EVPI has been suggested by several authors (Sculpher and Claxton 2005; Claxton and Sculpher 2006) as providing:

(i) An upper bound for the value of prospective research; and
(ii) A 'necessary condition' for further research where EVPI is 'large enough' to justify potential future research.

However, whether population EVPI is 'large enough' or not requires considering the expected cost and expected value of actual research options, both of which depend on the size and nature of research (Eckermann et al. 2010). Indeed, a jurisdictions local population EVPI is very much theoretical, as research to provide perfect information (maximise EVPI per patient) would require research on the whole population to benefit, but the population expected to benefit from further evidence is reduced by the time research takes to update decision-making. The more extensive the research undertaken to increase EVSI per patient towards that of EVPI per patient, the longer the research takes and the more the patient population who can benefit from the research is eaten up (Eckermann and Willan 2008b). Conversely, shorter trials while limiting the reduction in population who benefit also limit the extent to which EVSI per patient approaches EVPI per patient. Consequently, the expected value of actual research over time, the realisable expected value of sample information (EVSI) for the population who can benefit is consequently usually significantly smaller than population EVPI. Hence, optimal decision-making and trial design needs to recognise that the expected value of information from any actual trial will frequently be orders of magnitude less than population EVPI in depending on both the size and time taken to undertake the trial and their opposing impacts on population EVSI factors, EVSI per patient and the population to which they apply.

### *5.2.2 Candidate Set 2: Per Patient and Population EVSI*

The expected value of sample information (EVSI) per patient from an actual trial or research design is calculated as the *a priori* expectation of a reduction in EVPI per patient with additional information. That is, per-patient EVSI is EVPI per patient with current information minus the expectation of EVPI per patient with updated information from the trial. This is shown in Fig. 5.3, where a representative

posterior distribution of INB has been added, noting that, in practice, this is but one of potentially many posterior distributions.

The expectation is that the distribution for INB will become tighter with additional information due to a reduction in variance. This does not imply that the information from a new trial will reduce EVPI (have a tighter distribution of INB); however, the *a priori* expectation is that it will. Further, the amount by which EVPI per patient is expected to reduce is conditional on the size of the trial; hence per-patient EVSI is an increasing function of the proposed trial size.

Multiplying EVSI per patient by the patient population over a time horizon ($T$) expected to benefit from trial evidence represents population EVSI. Note that this means that in spite of per-patient EVSI increasing with trial size, population EVSI at some trial size will start to reduce as a function of trial size given the extent of the future population who benefit from a trial evidence is eaten up over the time patients recruited to trial take to accrue as well as follow up, analyse and report on, as considered in detail in Sect. 5.7. It should be clear that if any optimal sized trial eventuates, it will be well before a trial size where the future population who benefit or the time horizon are eaten up, given at that point population EVSI is 0, while costs of trialling would be maximised.

More generally for any research proposal population, EVSI is only half the equation to considering whether a trial design is potentially worthwhile, let alone optimising trial design in maximising expected value relative to cost or return on research funding. The expected cost as well as expected population EVSI need to be considered in identifying whether a trial is potentially worthwhile, represents an optimal trial design for that research area or should be competitively funded across proposals considered.

### 5.2.3 *Candidate 3: The Expected Value Less Cost or Expected Net Gain (ENG) of a Given Trial*

Whether population expected value of sample information is 'large enough' at any size of trial needs to consider the expected value of the trial to decision-makers relative to the expected costs and hence the expected net gain (ENG). Direct costs of research can vary from negligible with routinely collected evidence to that of a large RCT. In addition, the total cost of research to decision-making in a jurisdiction for cases of interest with positive while uncertain INB also includes an opportunity cost of delay for patients who don't receive the NB maximising therapy.

Where two strategies are compared for cases of interest with positive expected while uncertain INB, the per-patient opportunity cost of patients on standard therapy with delay and trialling is equal to the extent of positive expected INB. At a population level, such opportunity costs of delay can be substantial and indeed dwarf direct research costs. This is particularly the case where prior INB per patient is large and a large population do not receive the therapy outside the trial setting until decision-making is updated with trial evidence (Eckermann and Willan 2007, 2008b).

These opportunity costs are considered at length in estimating costs with delay and trialling in Sect. 5.5.

Questions in relation to whether further research is potentially worthwhile and the ENG of a given trial design (questions (i) and (ii)) can be addressed to the extent that a given trial design has positive ENG or not. However, where ENG of a trial design is positive, this still does not identify efficient or optimal trial design and decision-making, while where ENG of such a trial design is negative this does preclude other trial designs having positive ENG.

## 5.3 What Is Required to Inform Decision-Making Questions: Optimising ENG or Return on Research

The expected value, and expected cost and hence ENG (value less cost) of research are conditional on the extent of proposed research. Consideration of expected value, costs and ENG of actual trial designs are necessary to inform:

(i) Whether any further research is potentially worthwhile;
(ii) Whether a specific research design is worthwhile;
(iii) Optimal research design; and
(iv) Optimal prioritisation of research across HTAs

To illustrate why, we consider the population EVPI, EVSI, cost, expected net gain (ENG) of six hypothetical studies of research proposal areas A to F, in Table 5.2.

Table 5.2 makes clear that there is no necessary relationship between population EVPI and EVSI, let alone ENG or return on investment from an 'optimal' VOI trial design for any given jurisdiction. Using population EVPI would suggest an ordering of research funds supporting first D and then B, A, E, F and C. However, the relationship between EVPI, EVSI and optimal ENG aprt from prior evidence underlying the INB distribution depends on the expected cost and value of actual research which are in turn conditional on relevant key decision contexts (Eckermann and Willan 2007; Eckermann et al. 2010), such as those in relation to time, direct and

**Table 5.2** Population EVPI and EVSI and ENG of optimal[a] trial designs for six research proposals in a jurisdiction of interest

| HTA | Current EVPI | EVSI | Direct research costs | Opportunity cost of delay | Total cost US$ | ENG US$ | Return on direct investment |
|---|---|---|---|---|---|---|---|
| A | 50M | 10M | 1M | 4M | 5M | 5M | 500% |
| B | 100M | 50M | 10M | 15M | 25M | 25M | 250% |
| C | 5M | 2M | 1M | 0 | 1M | 1M | 100% |
| D | 101M | 10.1M | 6M | 4M | 10M | 0.1M | 2% |
| E | 25M | 9.8M | 2M | 8M | 10M | −0.2M | −10% |
| F | 6M | 3M | 3M | 0.5M | 3.5M | −0.5M | −17% |

[a]Optimal trial design given a trial is proposed (trial fixed costs faced)

opportunity cost of delay whether trials are feasible with adoption, global versus local trials and implementation (Eckermann and Willan 2008a, b, 2009, 2013; Willan and Eckermann 2010). Hence, while research option D has the highest theoretical population EVPI, it has a relatively low ENG and return on investment from actual research designs, reflecting orders of magnitude lower EVSI at an optimal trial design in practice.

It should be clear then that the size of population EVPI with current evidence does not come close to providing a necessary condition to inform the decision of whether further research is worthwhile or not. The size of population EVPI for a jurisdiction is only a theoretical upper bound, and one that often will be orders of magnitude greater than the highest realisable population EVSI let alone EVSI at an optimal sized trial if it exists. More generally in failing to consider the expected value relative to cost of actual trial designs can neither exclude nor recommend further research (Eckermann et al. 2010). To inform optimal joint research and reimbursement decisions and associated research designs requires comparing the expected value to expected cost of actual research, costs which can be trivial or substantial depending on direct and opportunity cost arising with type and scale of research.

Hence, population EVPI with current evidence is neither necessary nor sufficient to inform research decisions about whether further research is potentially worthwhile. Never the less, current EVPI per patient is useful to the extent that it provides the first technical step towards estimating the expected value of sample information (EVSI) from actual research designs. The technical necessity of estimating per-patient EVPI as a step towards estimating per-patient and population EVSI should not however be confused with population EVPI providing a necessary condition to inform decision-making, and particularly in relation to research questions (i)–(iv).

Establishing whether EVSI is greater than expected cost provides necessary and sufficient conditions to address decision-maker questions in relation to ENG being positive and further research being potentially worthwhile (questions (i) and (ii)), and is necessary while not sufficient to address questions in relation to optimal trial design and efficient allocation of constrained research funding (questions (iii) and (iv)). Fully adressing questions in relation to optimal trial design and research funding prioritisation require optimising ENG or investment return across designs and research options.

Table 5.2 also demonstrates the potentially important distinction between maximising ENG and return on investment. The optimal research design in addressing optimal trial design for a given HTA (question 3) historically was suggested as the point at which ENG is maximised. However, in considering research prioritisation (question (iv)) and, in particular, the opportunity cost of investing limited research funding across a set of research proposals, modification of the objective to maximising the return to research (i.e. ENG per additional dollar of direct research expenditure) is suggested (Eckermann et al. 2010). This is the case as the highest ENG may not necessarily represent the highest return to investing in research as illustrated in addressing research prioritisation across multiple research study options in Table 5.2.

For example, optimal trial design for B has the highest ENG of $US25 million from undertaking research, five times that of study A. However, the return from research option A is twice that of B (500% vs. 250%) as the direct cost research investment with 'A' ($US1 million) is one-tenth that of option B. In prioritising research based on return to investment (ENG per $ invested), option A should therefore be prioritised higher than B. Eckermann et al. (2010) show that to address questions (i) to (iii) within a jurisdiction requires identifying the maximum ENG. Repeating such analysis across different HTAs allows questions relating to prioritisation of research (question iv) to be addressed, where with limited research funding getting the best expected research value is clarified as requiring optimising ENG return per research dollar invested.

In contrast, prioritising research funding across studies based on population EVPI does not even address the risk that research with negative ENG is funded, let alone ensure that the research with highest expected return is funded. For example, if research funding were restricted to $US12 million and prioritising based on EVPI, then trials B, A, E and F would be chosen and have a combined ENG of at most $US4.4 million. It is at most $4.4 million given we are being favourble to EVPI in assuming each trial (B, A, E and F) were by chance proposed at the 'optimal' study size, where ENG was highest, given a research study has been proposed. However, note that even with this favourable assumption trials for options E and F have negative ENG and would never be considered for prioritisation where decision-making is based on maximising ENG or return to research.

In contrast to population EVPI, prioritising $US12 million funding based on maximising return to research (and ENG of total research funding invested), leads to ENG of $US31.0 million, a $US26.6 million (or 605%) higher return than that the best case with prioritisation based on EVPI, with funding of trials for A, B and C. If one was kinder still to the cause of EVPI in this example and allowed a research budget of $US16 million to enable support of the two strategies with highest EVPI (D and B), then the ENG from optimal studies for *D* and B of $US25.1 million would still be significantly less than the ENG of $31.0 million expected with optimal allocation of $US12 million based on return to research.

Indeed, with $US16 million and research options A to F, the optimal solution based on return to research would be to use $US12 million to support A, B and C with ENG of $US31 million and use the remaining $US4 million to support highest return across other potential actions (Eckermann et al. 2010). These actions could include the most cost effective expansion of current programs or interventions or future research for new interventions (Eckermann and Pekarsky 2014). Naturally this approach could also be applied more broadly to maximise outcomes across all potential actions – research, investment or implementation strategies – following programme budgeting and marginal analysis (PBMA) principles (Ruta et al. 2005), something we return to consider at length in Chap. 11.

In summary then, Table 5.2 has highlighted the clear need for, advantages and value of VOI measures moving beyond EVPI in taking the additional steps in estimating EVSI and expected costs of research and optimising ENG (expected value less cost) or ideally expected return (ENG per dollar of research invested) to research investment.

## 5.4 Broader Dangers of Population EVPI in Allocating Research Funding

Current population EVPI is the value of information equivalent to burden of disease in measuring size of problems rather than addressing what can be done about them. To address what can be done in identifying optimal design and research prioritisation with return on investment requires estimating EVSI of actual trial or research designs and optimising EVSI relative to expected costs and hence ENG or return on research of these designs. Prioritising with population EVPI alone is actively dangerous as it results in losses both from supporting research with unknown (and potentially negative) ENG as well as not identifying, encouraging and supporting research with high return as in Table 5.2.

Further, in considering the notion of a minimum sufficient level of population EVPI, it should be noted that measuring population EVPI within a jurisdiction assuming perfect implementation or only estimating the value of research within jurisdiction can significantly underestimate the expected value of undertaking further research.

Willan and Eckermann (2010) show that where a positive relationship is expected between strength of evidence and degree of implementation, relaxing the assumption of perfect implementation to allow for the more realistic expectation of imperfect implementation increases the expected value of research (EVSI). That is, for cases of interest with current positive while uncertain INB, where there is a positive relationship expected between the extent of implementation and strength of evidence, additional information has expected value in increasing degree of implementation, in addition to reducing decision uncertainty (Eckermann and Willan 2016). Consequently, population EVPI measured within a jurisdiction assuming perfect implementation can easily underestimate the upper bound for EVSI within any jurisdiction where implementation is expected to improve with additional evidence. Furthermore, it should also be noted that the EVSI of research is greater globally than that locally, given publicly available trial evidence has the non-rival and non-excludable characteristics of pure public goods (Eckermann and Willan 2009). Hence, EVSI can be vertically summed across jurisdictions, with positive value in each jurisdiction to the extent evidence is translatable, as we consider in detail in Chaps. 6 and 7.

Having established the dangers of population EVPI and importance of optimising population ENG and return of research of actual trial design in informing questions (i) to (iv), we now consider the ability and relative simplicity of current VOI methods in allowing this.

## 5.5 What VOI Method(s) Enable ENG Optimisation

Simple estimation of EVSI and ENG at any trial size for optimal overall trial design is enabled by the use of VOI methods with a closed-form solution applying the central limit theorem (CLT) (Willan and Pinto 2005a, b). In contrast to such CLT-based

VOI methods, non-parametric bootstrapping VOI methods are computationally expensive in estimating expected posterior EVPI for EVSI let alone expected cost and ENG allowing for relevant decision contexts (Eckermann et al. 2010) and prohibitive in optimising ENG across designs (Ades et al. 2004). Further, CLT methods assuming a bivariate normal distribution for INB have been shown to outperform alternative non-parametric bootstrapping methods with small samples and skewed data in particular with asymptotic properties of the CLT shown to arise at smaller samples than bootstrapping (Nixon 2010). Briggs et al. (1999) similarly found the performed better.

Most importantly for VOI methods to better inform real decision-making (Spiegelhalter 2004), a closed-form solution applying the CLT (Willan and Pinto 2005a, b) has been illustrated to provide an appropriate framework for optimising joint research and reimbursement decisions in allowing for key decision contexts including:

(i) Opportunity costs, time and option value of delay (Eckermann and Willan 2007, 2008a, b)
(ii) Imperfect implementation (Willan and Eckermann 2010)
(iii) Pricing under uncertainty (Willan and Eckermann 2012)
(iv) Global value of information across jurisdictions and risk sharing (Eckermann and Willan 2009, 2013)

We consider decision contexts (i) and (ii) in this chapter and (iii) and (iv) in Chaps. 6 and 7, respectively.

### 5.5.1 *Appropriately Allowing for Within Jurisdiction Decision Contexts Applying the CLT*

Within a jurisdiction for cases of interest where new therapies have evidence of positive while uncertain net benefit potentially optimal joint research and reimbursement decisions locally (Eckermann and Willan 2007) are between options to:

(i) Delay the decision and undertake a trial (DT) where opportunity costs of delay are faced while value of information is an option value of delay;
(ii) Adopt without further research – no trial (AN); and
(iii) Where feasible adopt the new intervention and undertake a trial (AT), in which case reversal costs are faced in considering the expected value of additional research and hence assessing EVSI.

Note that where INB of a new therapy is negative with current evidence, then the new therapy should *a priori* be rejected in the absence of a lower price enabling positive INB given:

(i) The opportunity cost of researching therapies with current expected negative INB not expected to be adopted is research into promising therapies and programs with positive INB that are; and

(ii) Appropriate incentives are created for manufacturers to lower price where feasible to enable positive INB and potential for adoption but also appropriately discourage further research where they do not expect positive INB given expected effects and costs including price of the intervention.

Note also that for usual cases of interest with positive while uncertain incremental net benefit driven by expected net clinical benefit for patients, adopting and trialling within a jurisdiction becomes infeasible as well as unethical. This is because informed patients would prefer to be outside the trial setting with certainty of adopting the new therapy rather than face a random chance of the new therapy in the trial setting.

Nevertheless, adopting and trialling within a jurisdiction can be feasible and ethical for cases where positive while uncertain INB is driven by lower cost rather than expected net clinical benefit. Potentially adopting and trialling could also be feasible within a jurisdiction if there were no access to healthcare outside a trial setting, while ethically more questionable.

Importantly, adopting and trialling does become feasible moving beyond the within-jurisdiction trial setting (e.g. UK, USA or Australia) to consider global trial designs across jurisdictions. Indeed the ability for individual jurisdictions to adopt and trail is highlighted as a distinct advantage of optimal global societal decision-maker trials in Chap. 6. Further, global manufacturer trials in Chap. 7 are shown to also create the potential for robust risk sharing under such arrangements. Nevertheless, in this chapter, we restrict consideration to within-jurisdiction analysis where one should note AT is often not feasible.

## 5.6 How Can VOI Methods Inform the Choice Between AN, DT and AT Where Feasible?

VOI methods invoking the CLT which provide a closed-form solution to EVSI and enable simple optimising of the ENG across RCT designs, have been established (Willan and Pinto 2005a, b) for cases of interest with positive INB and extended to allow for key decision-maker contexts (Eckermann and Willan 2007, 2008a, b, 2009, 2013; Willan and Eckermann 2010, 2012).

To algebraically consider these methods in what follows, let

- $k$ = the rate of incidence of the condition in question
- $a$ = the accrual rate in to the trial
- $n$ = the number of patients per arm recruited to the trial
- $\hat{b}$ = the estimate of INB from trial data
- N$t$ = the number of patient that can benefit from the new health-care intervention at time $t$, where the current time is $t = 0$
- $\tau$ = time for follow up and analysis
- $C_f$ = the fixed cost of doing the trial

- $C_v$ = variable cost per patient of being on trial incremental to that of being treated with sae arm outside of trial (assumed to be the same by treatment arm for simplicity)
- $C_r$ = the cost of reversing a decision to adopt

Further, let $b_0$ be the current estimate of mean incremental net benefit (INB), with associated variance $v_0$, where $b_0, v_0 > 0$. Let $\mathrm{ENG_D}$ be the expected net gain in the comparison of DT and AN, i.e. $\mathrm{ENG}_D$ is the difference between the expected value less cost with delay of the trial (sample) information. Let $\mathrm{ENG_A}$ be the expected net gain in the comparison of AT and AN, i.e. $\mathrm{ENG_A}$ is the difference between the expected value of the trial (sample) information minus the cost of that trial with adoption. If $\mathrm{ENG}^*_D$ and $\mathrm{ENG}^*_A$ are the corresponding maximum values with respect to trial sample size, then the following are optimal decision rules in maximising ENG of the decision:

$\mathrm{ENG}^*_\mathrm{D} < 0 \,\text{and}\, \mathrm{ENG}^*_\mathrm{A} < 0, \text{then AN is preferred.}$

$\mathrm{ENG}^*_\mathrm{A} < \mathrm{ENG}^*_\mathrm{D} > 0, \text{ then DT is preferred.}$

$\mathrm{ENG}^*_\mathrm{D} < \mathrm{ENG}^*_\mathrm{A} > 0, \text{then AT is preferred.}$

Alternatively, optimising across designs can be identified as the option with highest positive rate of return on direct research funding. That is, ENG* could be determined for adoption or delay as the trial that maximises

$$\mathrm{ENG} / \left(C_\mathrm{f} + 2nC_\mathrm{v}\right).$$

Assuming bivariate normality of INB under the CLT, as in Figs. 5.1, 5.2 and 5.3, then

$$\dot{b} \sim N\left(b_0, v_0 + 2\sigma^2 \big/ n\right),$$

where $\sigma^2$ is the between-patient variance of incremental net benefit.

Hence, the per patient current EVPI or expected opportunity loss relative to perfect information can be calculated as

$$E_0\left[L(b)\right] = \int_0^{-\infty} -bf_0(b)\,db$$

$$= e^{-b_0^2/2v_0}\sqrt{v_0/2\pi} - b_0\,\phi\left(-b_0 / v_0^{1/2}\right),$$

where Φ(.) is the cumulative distribution function for a standard normal random variable.

Following a trial of size $n$, the updated estimate of $b$ following the CLT is

$$b_1 = v_1\left(b_0 / v_0 + n\breve{b} / 2\sigma^2\right)$$

with variance,

$$v_1 = \left(1/v_0 + n/2\sigma^2\right)^{-1}$$

where $b_1$ is also normally distributed.

The associated per patient posterior expected opportunity loss is calculated as

$$E_{\tilde{b}}\left\{\int_0^{-\infty} -bf_1(b)\,db\right\}$$

a formulation for which is provided in Willan and Pinto (2005a, b). EVSI per patient is prior EVPI or expected opportunity loss less the expectation of this posterior expected opportunity loss.

Importantly the closed-form solution for EVSI per patient in Willan and Pinto (2005a, b) applying VOI methods under the central limit theorem (CLT) (Nixon 2010; Willan and Briggs 2006; Briggs et al. 1999) enables EVSI estimation and ENG optimisation with knowledge of only the prior mean and variance of INB, the time horizon over which evidence is expected to be valuable (Eckermann and Willan 2008b) and usual data on expected patient incidence, trial accrual and direct trial costs (fixed and variable costs).

## 5.7 Expected Value and Cost of Trials with a Delayed Reimbursement Decision (DT Versus AN)

The value of trial information undertaken with delay now compared to adopting now based on current expected INB (DN vs AN) is not affected by costs of reversal, as delaying avoids reversal with what turn out with further research to be bad decisions to adopt a strategy with current evidence of expected while uncertain INB. However, there is an opportunity cost of delay in undertaking further research for cases of interest with positive expected while uncertain INB. The expected value of this opportunity costs is the prior expected INB that would otherwise have arisen with adoption now for the patients on standard therapy within and outside the trial arising during recruitment, follow-up and analysis time, that is, until a revised decision is made.

Opportunity costs of delay with DT arise for patients both outside and inside trial receiving standard care and hence contribute to both

(i) Variable opportunity costs dependent on size of the trial ($n$) given time of recruitment of $b_0(k \times 2n/a - n)$ or equivalently $n(b_0(2k/a - 1))$ as a function of trial size; and
(ii) Fixed opportunity costs independent of trial size associated with the fixed time period of follow-up and analysis of $b_0\tau k$.

These opportunity costs with DT are additional to direct costs of trialling, which include the fixed ($C_f$) and the variable cost of trialling each patient ($C_v$) which for a standard 2 arm RCT of $n$ patients per arm results in a direct cost of $C_f + 2nC_v$. Consequently total costs to decision-makers of trialling with DT are

$$C_f + b_0 \tau k + n\left(2C_v + b_0(2k/a - 1)\right)$$

Population EVSI is found by multiplying the per-patient EVSI by the size of the population to whom information from the trial is valuable, which, following Eckermann and Willan (2008b), will be reduced when accounting for the duration of trial accrual, follow-up, analysis and reporting. Time needs to be explicitly modelled allowing for study accrual rate ($a$), time for follow-up and analysis ($\tau$).

In determining population EVSI, the expected time for information updating is

$$\frac{2n}{a} + \tau$$

This time reduces patient population over the time horizon ($T$) given an expected patient population incidence rate ($k$). Hence, the remaining patient population within jurisdiction who can benefit from trial information from a trial of size $n$ per arm is

$$k\left(T - \left(\frac{2n}{a} + \tau\right)\right)$$

Multiplied by EVSI per patient, this represents population EVSI.

Given a time horizon $T$, population EVSI is shown in Fig. 5.4 to initially increase while the proportional increase in the value of additional evidence per patient is greater than the reduction in the patient population with additional recruitment time. Population EVSI reaches its highest level at the point where the proportional increase in EVSI per patient is equal to the proportional reduction in

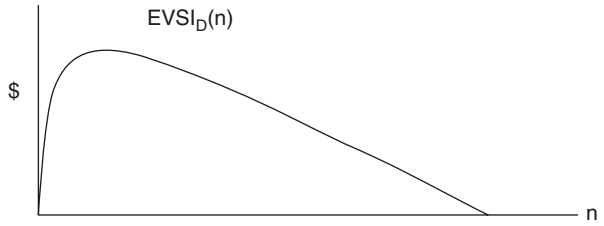


**Fig. 5.4** Population EVSI as a function of study size ($n$) with related reduction of time horizon

the patient population and then reduces eventually to 0 where there is no population left, corresponding to the trial size where trial reports at the time horizon $T$. That is, population EVSI reduces to 0 where the trial has eaten up the time horizon when it reports, as there is no population left to benefit from additional evidence, where

$$T = \left( \frac{2n}{a} + \tau \right);$$

or equivalently at the trial size where

$$n = \frac{a\left(T - \tau\right)}{2}$$

However, note that given positive expected costs of undertaking research (direct and opportunity cost with delay and trialling) in maximising ENG, we are only interested in trial sizes from Fig. 5.4 where EVSI is still increasing with $n$. Given EVSI and total costs as a function of $n$, the ENG of proposed trials with delay relative to adopting now can then be found as population EVSI minus the expected cost of that trial size per arm, including the opportunity cost of delay expected to be incurred by patients who receive the old technology while the trial is performed. It is not optimal to undertake trials where ENG is negative, while it is potentially optimal to trial where ENG is positive. Hence, a positive ENG is necessary while not sufficient to establish the optimal trial design, where ENG or return on investment is maximised.

However, a negative ENG for a given trial design does not imply that there is no research design for which ENG is positive and hence current evidence sufficient to AN. For example, if EVSI is less than the expected costs of a frequentist-designed trial of given size, this does not mean that there will not be another trial or research study where EVSI is greater than expected cost and hence have positive ENG. Other sized trials or low-cost alternative research to reduce decision uncertainty can have EVSI greater than their expected cost to decision-making. Hence, consideration of EVSI relative to expected costs of frequentist-designed trials can only ever partially inform questions in relation to optimal trial design and decision-making to the extent that certain research designs are excluded where ENG is negative.

In optimising ENG of trial design and decision-making, we have noted that the expected value of sample information EVSI($n$) per patient, and the population to which that EVSI applies to find population EVSI as their product, as well as expected direct and opportunity costs depend on the sample size ($n$). Hence, expected net gain ENG($n$) = EVSI($n$) − ETC($n$) or return on research (ENG per dollar allocated to research funding) is also a function of the sample size, $n$.

Consequently, ENG can be optimised as a function of sample size, with sufficient evidence to adopt now for cases of interest with positive while uncertain INB if ENG is always negative. That is, if EVSI is never greater than expected cost of

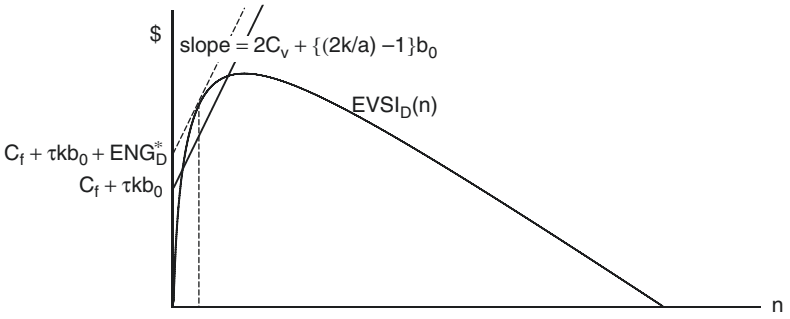


**Fig. 5.5** The research option value and opportunity cost of delay (DT vs. AN) (Source: Eckermann and Willan (2007))

alternatively sized trials. Otherwise, over some range of trial sizes EVSI will be greater than expected costs, and hence ENG with DT will be positive and optimised at a trial size as in Fig. 5.5 where marginal value from trialling (slope of the population EVSI curve) is equal to the marginal cost of trialling (slope of the expected total cost function). For trials smaller than this, at the margin, the expected value is greater than expected cost of increasing the trial size, while for trial sizes larger than this the expected cost is greater than the expected value of expanding the trial size at the margin.

Alternatively, return on research (ENG per research dollar invested) can be optimised where

$$\mathrm{ENG} / \left( C_{\mathrm{f}} + 2nC_{\mathrm{v}} \right)$$

is maximised and compared with other options, including AT where feasible but more generally other options for investment in research or services (Eckermann, Karnon and Willan 2010).

In interpreting Figure 5.5, dealying the decision about whether to adopt while trialling (DT) ensures feasibility of the collection of further evidence. This additional evidence has expected value in reducing the expected opportunity loss of adopting the strategy maximising expected net benefit with current evidence and represents the EVSI with DT. The expected value of sample information for DT versus AN represents a research option value of delay (Eckermann and Willan 2008a) and is not affected by costs of reversal.

The expected value of sample information per patient with delay and trialling (DT) at time t is given by

$\mathrm{EVSI}_{\mathrm{D}} = N_t \left\{ \int_0^{-\infty} -b \left\{ f_0(b) - f_1(b) \right\} db \right\}$, where $f_i(.)$ is the probability density function for a normal distribution with mean $b_i$ and variance $v_i$, $i = 0, 1$.

Population EVSI for a trial of size $n$ with delay ($EG_D(n)$), following Eckermann and Willan (2007) is then given by

$$\mathrm{EG}_{\mathrm{D}}(n) = E_{\tilde{b}}\mathrm{EVSI}_{\mathrm{D}}$$

$$= E_{\tilde{b}}\left[ N_t \left\{ \int_0^{-\infty} -b\left\{ f_0(b) - f_1(b) \right\} db \right\} \right]$$

$$= \left\{ N_0 - (\tau + 2n/a)k \right\} \left[ \int_0^{-\infty} -bf_0(b)\,db - E_{\tilde{b}} \left\{ \int_0^{-\infty} -bf_1(b)\,db \right\} \right]$$

where $N_0 = kT$ is the expected population over the time horizon now, reduced by the incident population arising over the time until the trial reports ($(\tau + 2n/a)k$), to represent the remaining population to which the trial evidence has value. In considering the EVSI per patient components to which this population estimate applies in estimating population EVSI:

$$\int_0^{-\infty} -bf_0(b)\,db = (v_0/2\pi)^{1/2} \exp\left(-b_0^2/2v_0\right) - b_1\phi\left(-b_0/v_0^{1/2}\right)$$

where Φ(.) is the cumulative distribution function for a standard normal random variable.

A closed-form solution formulation for

$$E_{\tilde{b}}\left[ \int_0^{-\infty} -bf_1(b)\,db \right]$$

is provided in Willan and Pinto (2005a, b) as previously indicated.

Now, the total cost of delay and trialling as a function of trial size per arm ($n$) given opportunity cost of delay arises for all patients except the $n$ in the active arm in the trial is

$$\mathrm{TC}_{\mathrm{D}}(n) = C_f + 2nC_v + (tk - n)b_0$$

$$= C_f + 2nC_v + \left\{ (\tau + 2n/a)k - n \right\} b_0$$

The expected net gain from delay and trialling with a trial of size $n$ is consequently

$$\mathrm{ENG}_{\mathrm{D}}(n) = \mathrm{EG}_{\mathrm{D}}(n) - \mathrm{TC}_{\mathrm{D}}(n)$$

Hence, given current estimates of uncertainty in relation to incremental net benefit ($b_0$, $v_0$), the incidence and accrual rate of patients ($k$, $a$), fixed and variable costs of trialling ($C_f$ and $C_v$) and the time horizon ($T$) over which evidence has value, the expected net gain of trialling while delaying $\mathrm{ENG}_{\mathrm{D}}(n)$ can be optimised with respect to trial size ($n$).

For positive ENG and trial size ($n$), the optimal sized trial and ENG with delay is

$$\mathrm{ENG}_{\mathrm{D}}^{*} = \mathrm{ENG}_{\mathrm{D}}\left(n_{\mathrm{D}}^{*}\right) = \max_{n>0} \mathrm{ENG}_{\mathrm{D}}\left(n\right)$$

where $\left(n_{\mathrm{D}}^{*}\right)$ is the optimal sample size for DT versus AN if $\mathrm{ENG}_{\mathrm{D}}^{*} > 0$.

If $\mathrm{ENG}_{\mathrm{D}}^{*} < 0,$ the optimal sample size is 0.

The decision rule for DT versus AN is seen graphically in Fig. 5.5. The $\mathrm{TC}_{\mathrm{D}}(n)$ line has intercept

$$C_{\mathrm{f}} + \tau k b_0$$

and slope

$$2C_{\mathrm{V}} + \left\{\left(2k/a\right) - 1\right\} b_0$$

If the $\mathrm{EG}_{\mathrm{D}}(n)$ curve lies below the $\mathrm{TC}_{\mathrm{D}}(n)$ line for all $n$, then $\mathrm{ENG}_{\mathrm{D}}^{*} < 0$ and the optimal sample size is 0.

## 5.8 EVSI where Adopting and Trialing is Feasible

New trial evidence has value in reducing the expected value of losses from decisions made under uncertainty. An assumption made in calculating EVSI of trial information based on reduction in expected value of losses avoided (EVPI) is that the avoiding of losses with further information is costless. This assumption is valid where decisions have been delayed while trials are undertaken (DT). However, where adopting the new intervention and undertaking a trial (AT) is feasible and undertaken at time 0 then, reversal of adoption is not costless. For cases of interest with expected positive while uncertain INB adopting while trialling (AT) is only feasible and ethical with local trials where INB is driven by expected cost savings, given where INB relates to net clinical benefit informed patients would prefer certainty of adoption outside of trial to a chance within trail. Costs of reversing decisions are faced with AT, reducing the expected likelihood and value of changing decisions and hence the EVSI of trials undertaken with adoption.

EVPI and EVSI calculations until now with delay as the decision context for trialling have implicitly assumed that avoiding losses with perfect or further information is costless. However, expected costs of reversal ($C_{\mathrm{r}}$) are faced if the new therapy is adopted at the same time a trial is undertaken (AT). Costs of reversal ($C_{\mathrm{r}}$) include costs of reversing public health messages and unamortised costs of technology and training at time of decision reversal.

Reduction in the value of information with costs of reversing decisions becomes most obvious in the case where the costs of reversal are high enough that the decision to adopt becomes irreversible (Bernanke 1983, Tirole 1988). As considered in

Eckermann and Willan (2008b), decisions become irreversible where given costs of reversal, it would always be better to live with the decision rather than reverse it. Hence, where a decision to a adopt a new intervention have high enough cost that it is irreversible in this sense, then there is no expected value of sample information once the new intervention is adopted, and hence EVSI should be 0, as discussed in Eckermann and Willan (2008b).

More generally where costs of reversal are faced with AT, the expected value of sample information for a trial of given size is less where the decision has been made to adopt rather than delay. Consequently, in considering optimal trial design where there is positive but uncertain incremental net benefit and AT is feasible, the expected costs and value of information of planned trials within a jurisdiction are conditional on whether the new intervention is adopted or not while such trials are undertaken. The expected cost and value of trial information depends on the simultaneous decision to adopt or delay as:

(i) Where the new intervention is adopted, the expected value of sample information from a trial is reduced, since costs of reversal are faced.
(ii) Where the decision is delayed, there is an additional expected opportunity cost for patients receiving the standard intervention outside the trial until trial information updates evidence.

Importantly then, where AT is feasible to find the optimal strategy and trial design requires identifying optimally sized trials and expected net gain in comparison of both:

(i) Delay the decision and undertake a trial (DT) versus adopt the new intervention without further research (AN), allowing for opportunity costs of delay.
(ii) Adopt the new intervention and undertake a trial (AT) versus adopt the new intervention without further research (AN), allowing for costs of reversal.

In comparing AT and AN, decision-makers face costs of reversal $C_r$ with trialling in attempting to avoid expected losses associated with negative incremental net benefit in the remaining population

$$N_t = k\left(T - t\right) = k\left(T - \left(\tau + 2n / a\right)\right)$$

where trial evidence updates decision-making at time

$$t = \tau + 2n / a$$

For values of posterior incremental net benefit ($b_i$) between $-C_r/N_t$ and 0, the cost of living with negative INB is less than cost of reversal ($C_r$), and hence it is optimal not to reverse.

Hence, for $b_i$ between $-C_r/N_t$ and 0, the optimal decision does not change with this trial evidence, and hence the expected value of information is 0.

For $b_i$ less than $-C_r/N_t$, the optimal decision is reversal to the standard intervention, since the cost of reversal $C_r$, is less than the expected loss of not reversing, $-N_t b_t$.

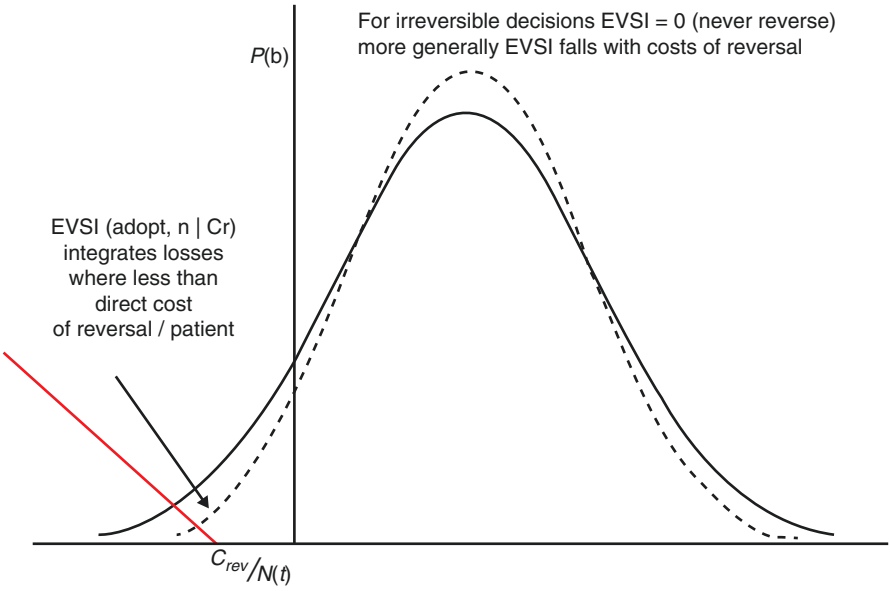


**Fig. 5.6** EVPI and EVSI per patient given costs of reversal with adopting and trialling (Source: Eckermann and Willan (2007))

However, while the expected net benefit of reversal is positive for INB less than $-C_r/N_t$, the expected value of information is reduced in comparison to that of delay by the cost of reversal per patient.

The combined effect of costs of reversal on expected value of information with adoption rather than delay can be modelled with the opportunity loss function shifted to the left by $C_r/N_t$ , as shown in Fig. 5.6.

The opportunity loss function for adopting the new intervention conditional on costs of reversal is given by

$$L(b) = 0 \; where \; b \geq -C_r / N_t$$

$$L(b) = -(b + C_r / N_t) \; for \; b < -C_r / N_t$$

The expected value of sample information (EVSI) at time t is given by

$$\text{EVSI}_D = N_t \left\{ \int_{-C_r/N_t}^{-\infty} -(b + C_r / N_t) \{ f_0(b) - f_1(b) \} db \right\}$$

$$\text{EVSI}_D = N_t \left\{ \int_{0}^{-\infty} -b \{ f_0^A(b) - f_1^A(b) \} db \right\}$$

where $f_i^{\mathrm{A}}(.)$ is the probability density function for a normal distribution with mean

$$b_i + C_{\mathrm{r}} / N_{\mathrm{t}}$$

and variance $v_i$, $i = 0, 1$.

Note as per delay that

$$N_{\mathrm{t}} = N_0 - tk = N_0 - (\tau + 2n/a)k$$

Making this substitution and taking the expectation with respect to the expected gain at time $t$, denoted EGA($n$), as a function of $n$, is given by

$$\mathrm{EG}_{\mathrm{A}}(n) = \left\{N_0 - (\tau + 2n/a)k\right\}\left[\int_0^{-\infty} -bf_0^{\mathrm{A}}(b)\,db - E_{\check{b}}\left\{\int_0^{-\infty} -bf_1^{\mathrm{A}}(b)\,db\right\}\right]$$

where $N_0 = Tk$ and

$$\int_0^{-\infty} -bf_1^{\mathrm{A}}(b)\,db = (v_0/2\pi)^{1/2}\exp\left(-(b_0 + C_{\mathrm{r}}/N_{\mathrm{t}})^2/2v_0\right) - b_1\phi\left(-(b_0 + C_{\mathrm{r}}/N_{\mathrm{t}})/v_0^{1/2}\right)$$

A formulation for

$$E_{\check{b}}\left[\int_0^{-\infty} -bf_1(b)\,db\right]$$

is derived by substituting

$$b_0 + C_r / N_t$$

for prior INB in the closed-form formulation in Willan and Pinto (2005a, b) for posterior expectation of EVPI.

The total cost of the trial is

$$\mathrm{TCA}(n) = C_{\mathrm{f}} + 2nC_{\mathrm{v}} + nb_0$$

which is the sum of the financial cost and the opportunity cost for the $n$ patients in the trial who receive the standard intervention, noting that those outside trial setting has adopted new therapy. Thus, ENGA($n$) = EGA($n$) – TCA($n$) and ENG can be optimised across trial size $n$ per arm for adoption and trialling (AT) as

$$\mathrm{ENG}_{\mathrm{A}}^{*} = \mathrm{ENG}_{\mathrm{A}}(n_{\mathrm{A}}^{*}) = \max_{n>0}\mathrm{ENG}_{\mathrm{A}}(n)$$

If $\mathrm{ENG}_{\mathrm{A}}^{*} < 0$, the optimal sample size is 0. The decision rule for AT versus AN graphically in contrast to Fig. 5.5 for DT versus AN has total cost function with

intercept $C_f$ (i.e. no fixed opportunity cost of delay during analysis and trial follow-up time) and slope $2C_v + b_0$ (i.e. opportunity cost relative to AN only arise for patients in the standard arm with AT). If the population EVSI curve falls below the TCA($n$) function for all $n$, then $\mathrm{ENG}_A^* < 0$ and the optimal sample size is 0.

## 5.9 Illustrating Optimising of Joint Optimising Research and Reimbursement Decisions – Early Versus Late External Cephalic Version

In a pilot study, 232 pregnant women presenting in the breech position were randomised between early (34 weeks with new intervention) versus late (37 weeks with standard intervention) external cephalic version (ECV) (Hutton et al. 2003). ECV is an attempt to manipulate the foetus into a cephalic rather than breech presentation. Elective caesarean section is accepted practice for breech presentation, and the primary outcome for the trial was non-caesarean delivery. In the early ECV arm, 41 of 116 (35.3%) patients had a non-caesarean delivery compared with 33 of 116 (28.4%) in the late ECV arm.

Let avoiding a caesarean section delivery be valued at $1000 due to reduced health system cost (while potentially also reflecting patient preference for non-caesarean delivery). Then with no other difference in effects or cost between early and late ECV, the estimate of INB is $b_0\Delta_e 1000$, where $\Delta_e$ is the incremental probability of a non-caesarean delivery for early relative to late ECV. Hence, the prior distribution for $b$, given the pilot data, has mean

$$b_0 = \left(41/116 - 33/116\right)1000 = 68.97$$

and variance

$$v_0 = \left( \frac{\frac{41}{116}\left(1-41/116\right)}{116} + \frac{\frac{33}{116}\left(1-33/116\right)}{116} \right) 1000^2 = 3724.78$$

The mean per-patient INB of $68.97 is a little more than one standard deviation (SD = $\sqrt{3724}$ = 61.0) greater than 0, reflected in the INB distribution shown in Fig. 5.7.

For a US incidence rate of 50,000 per year and a time horizon of 20 years, then the patient population at time 0 is 20 × 50,000 = 1,000,000. Based on a total budget for the planned frequentist 730 patients per-arm trial of $2,836,000, the fixed cost of setting up the trial is estimated as $C_f$ = $500,000 and the variable cost per patient as $C_v$ = $1600. It should be emphasised that $C_v$ is the incremental variable cost per patient of being on the trial relative to expected costs of treatment with the same intervention outside the trial. Additional costs of treating patients with a new

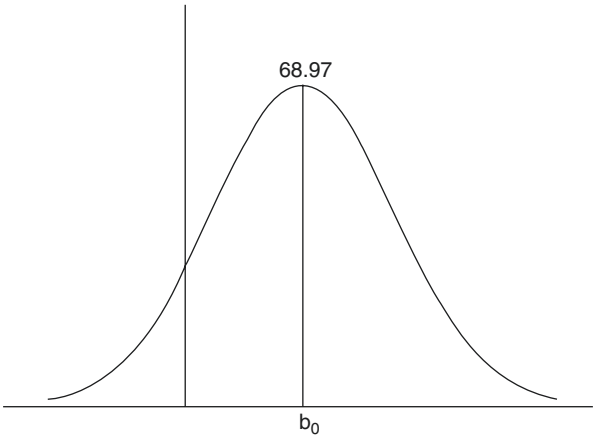


**Fig. 5.7** INB distribution from pilot evidence for early versus late ECV (Source: Eckermann and Willan (2007))

intervention are already implicitly included in expected incremental net benefit and should not be double counted by also being included in the variable cost of being on trial, $C_v$.

For an overall non-caesarean delivery rate of (41 + 33)/(116 + 116) = 74/232, an estimate of the between patient variance is

$$\sigma^2 = 74/232\left(1-74/232\right)1000^2 = 217,227$$

Using these values for $b_0$, $v_0$, $\sigma^2$, $N_0$, $C_f$ and $C_v$, an accrual rate of 500 per year (i.e. $a = k/100$) and allowing for a six-month period post accrual for intervention follow-up and data collection and analysis, the expected decision-maker value and costs for DT versus AN as a function of trial size ($n$ = 0–1000 patients per arm) are shown in Fig. 5.8. The expected cost is greater than expected value of a trial with delay (DT) at all sizes and becomes increasingly so as $n$ increases, due primarily to the opportunity cost incurred by the large number of patients who would receive the standard intervention during the accrual and the six-month data collection/analysis period.

However, even if unrealistically all patients are accrued to the trial (i.e. $a = k$), the optimal sample size is still 0, due to the large opportunity cost that still arises during the 6-month follow-up and analysis time, independent of accrual rate. Of the fixed cost of trialling of more than 2.2 million with DT (at $n$ = 0) in Fig. 5.8, $500,000 related to direct fixed costs of undertaking the trial and more than $1.7 million ($1,724,500) related to fixed opportunity cost of delay, from 25,000 patients receiving late rather than early ECV (NB $68.97 less per patient) during the 6-month analysis and follow-up study time. Indeed, the fixed opportunity cost of delay with DT during such time is more than $9400 per day.

Note that based on the same pilot data (Hutton et al. 2003), the frequentist-designed trial by investigators was a larger trial of 730 patients per arm. Given this represents a

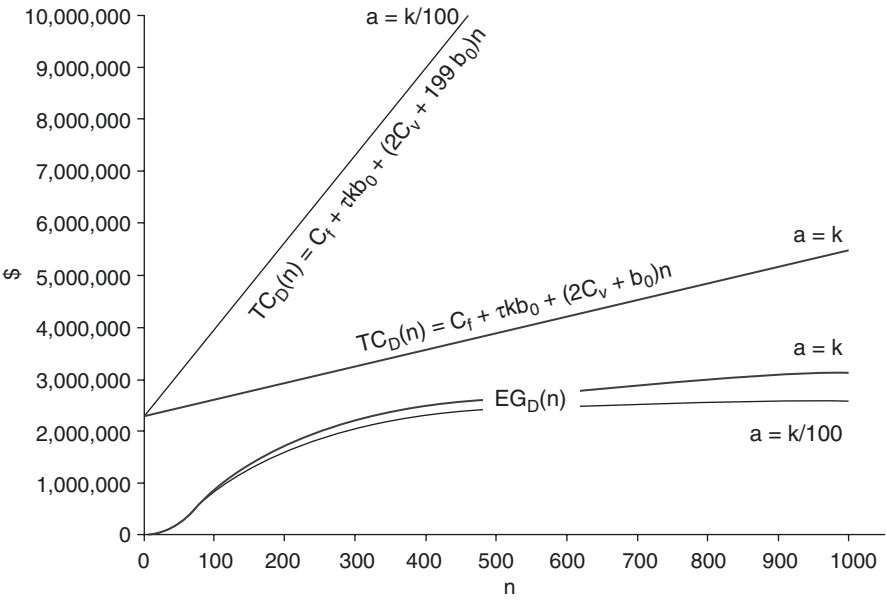


**Fig. 5.8** Expected value and cost of ECV trials (size *n*/arm) with delay, for realistic ($a = k/100$) as well as unrealistic ($a = k$) accrual rates (Source: Eckermann and Willan (2007))

trial with delay, the estimate of costs to decision-makers in USA of such a trial would be more than $4.5 million even if $a = k$. That is, even unrealistically assuming all incident patients in the US were recruited to that study (in 1460/50,000 = 0.0292 years, or just over a week and a half), a cost already more than the $3 million EVSI estimated for the population over a 20-year time horizon, with an expected net loss of more than $1.5million. More realistically with 1% of incident patients recruited across USA (500 per year and hence 2.92 years recruiting), the trial cost to decision-makers of a 730 patient per arm trial is estimated as more than $14.5 million, while population EVSI falls to nearer $2.5 million with the 2.63-year longer accrual time eating up part of the population to benefit. Overall this results in an expected net loss of more than $12 million projected in Fig. 5.8. This expected net loss is almost entirely attributable to additional opportunity cost with delay of $11.836 million, with $1.7245 million related to opportunity costs over the 6 months follow-up and analysis and $10,091,272 to opportunity costs of delay over the 2.92 years of recruitment for all those except trial patients with early ECV having late ECV. These opportunity costs dwarf EVSI but also the expected direct trial cost of $2.836 million.

The very large opportunity costs of delay while trialling with DT could however be almost eliminated with AT (adopting early ECV while trialling) if feasible, given opportunity cost of AT relative to adopting now with no trial is then limited to trial patients in the late ECV arm. In the case of early versus late ECV, adoption while trialling is feasible given net benefit is driven by health system cost savings of avoiding caesarean delivery. Further, in considering costs of reversal with AT there

are no additional equipment or technical training required for early versus late ECV implementation in practice. Reversal costs with AT in the absence of equipment and training are limited to public health guideline changes and messaging of provider networks for adoption. Hence, for the ECV case, adopting early ECV while trialling would avoid expected opportunity cost of delay, and the very real health system cost of additional caesarean sections arising with a DT design such as that of a frequentist trial, while facing limited potential costs of reversal.

More generally, adopting and trialling within a jurisdiction is only usually feasible and ethical for cases where positive while uncertain INB is driven by lower cost rather than necessarily expected net clinical benefit. Otherwise, where INB is driven by net clinical benefit patients prefer to stay outside the trial setting to ensure access to the expected clinical benefit of the new therapy rather than having a chance of the new therapy on trial. However, even where AT is feasible the avoiding of opportunity costs of delay does not necessarily indicate AT is optimal, given the impact of costs of reversal on EVSI and subsequently the ENG of AT versus AN needs to be considered in such cases. That is, EVSI conditional on cost of reversal needs to be compared with direct trial costs and the opportunity costs for patients receiving standard therapy with lower expected INB across trial designs to establish whether AT is preferable to AN at any trial size and if so optimise trial design (ENG or return on research). Hence, while for ECV adopting without a trial (AN) is preferred to delaying and trialling (DT) for any feasible trial given current evidence, if adopt and trialling (AT) is considered feasible, AT should also be compared with AN, and ENG of trial design optimised conditionally on costs of reversal.

For the case of ECV comparing AT versus AN with an expected cost of reversal of $2 million – associated purely with reversing public health messages if early ECV were not supported post trial – $ENG_A(n)$ is maximised at a sample size of 284 per arm, with expected net gain of $361,442. This results from an EVSI of $1,798,882 with that sized trial a financial cost of $1,408,800 and an opportunity loss from trial patients having late ECV of $19,586, as illustrated in Fig. 5.9, which is drawn to the same scale as Fig. 5.8 (comparison for trials with DT vs AN). The expected return on research from the trial with highest ENG is $361,442/$1,408,800 or 25.7%. Given the marginal cost approaches marginal EVSI at this size, if the objective were to optimise ENG per dollar of direct funding (i.e. ENG/(Cf+2nCv)) then this expected return would be maximised at a somewhat smaller trial with marginally higher expected return, while lower absolute ENG. For example with early vs late ECV an expected return of 29% with a trial of 220 per arm, with trial direct costs reduced by 15% while ENG reduced by 5%. Note this implies that in optimising funding across alternative research investments, or more generally across investment alternatives (research, reimbursement and regulation), optimal trial design for ECV would be somewhere between 220 and 284 patients per arm.

For research funders such as the Australian National Health and Medical Research Council (2009) or UK Medical Research Council (2009), consideration of return (ENG) per research dollar rather than maximising ENG in each research proposal is worthwhile given such comparisons at the margins will always allow them to stretch their funding to fund more research while still increasing the ENG

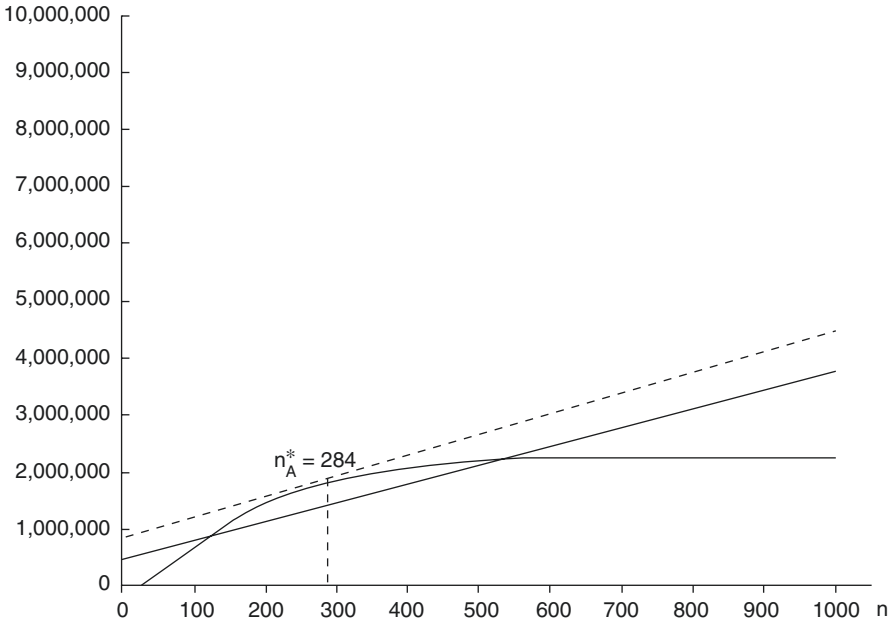


**Fig. 5.9** Expected decision-maker value, cost and expected net gain of ECV trials (size *n* per arm) with adoption in the USA (Source: Eckermann and Willan (2007))

from their funding pool. This is because for each competitive research proposal with positive ENG return on research return can be increased at the margins with somewhat smaller trial than that suggested by maximising ENG, as illustrated in the case of ECV. Hence, information on the size of trial where ENG return per research dollar is maximised as well as where ENG is maximised is valuable to the research funder and indeed could be mandated for provision, as part of a process of optimisation. It is an empirical issue for each research proposal as to how far below ENG maximisation in terms of trial size expected return, ENG per research dollar, is maximised - an empirical issue which reflects the marginal ENG and direct trial cost reductions. Never the less in the case of ECV where return on research is optimised (at a rate of return of about 29%) around 220 rather than 284 per arm, direct trial costs are reduced from 1.410 million to 1.204 million, given variable costs per patient of 1600 dollars.

For research funders when making decisions about how big trials should be and how many trials their constrained budget funding can optimally stretch to, those trials with highest ENG returns per research dollar, apart from being the first to be funded, should also be funded closer to their ENG maximising level than other proposals that end up being funded from the research funding pools with lower returns. This is the case given their marginal ENG return per research dollar from an extra patient on trial will be higher at the ENG per research dollar maximising point than those funded with lower returns. Optimisation will also clearly depend on

available alternatives reflecting local proposals within a jurisdiction. For example, if the ECV proposal was one of the lowest return projects funded then one would expect it to be funded at or close to 220 patients per arm, while somewhat higher if one of the highest return projects.

Such mechanisms used by the funder naturally rewards those proposals with highest expected ENG per research dollar in the size of their trails which naturally arise in optimising across funded trials their marginal advantages relative to proposals with lower return.

The above general marginal considerations in terms of maximising ENG across funded activities are naturally also considerations for funders more generally across research, reimbursement and regulation activities in Sect. 5.3 and general health system budget constrained optimising with local decision contexts in Chap. 11.

Returning to the case of ECV, the EVSI curve for AT versus AN in Fig. 5.9 is lower than that for the equivalently sized trial for DT versus AN (Fig. 5.8), due to costs of reversal ($2 million). However, while $ENG_D(n)$ is negative for all values of $n$, $ENG_A(n)$ is positive for values of $n$ from 122 to 530. The total cost function intercept (where $n = 0$) for AT versus AN, compared to that for DT versus AN, is smaller by $1.7245 million. This represents the opportunity cost avoided by AT for patients who receive the early ECV (new intervention) during the data collection/analysis period ( $\tau \times k \times b_0$). The coefficient of $b_0$ for opportunity cost of delay, part of the slope term of the cost line for AT versus AN, is one 199th of that for DT versus AN, with 198 times as many patients outside the trial on late ECV as on late ECV within the trial in the case of DT, given an expected 1% accrual rate.

While AT with 284 patients per arm maximises ENG for costs of reversal of $2 million, at some higher level of costs of reversal, AN would be preferred to AT. EVSI per patient decreases with costs of reversal for AT versus AN at any $n$. A cost of reversal of $5 million is sufficient to lower the expected value of information curve for AT versus AN below the cost line for all $n$. Consequently, AN would therefore be preferred to AT for costs of reversal of more than $5 million.

In summary, for the ECV example in the US, AN is preferred to DT for any feasible trial, and AT is preferred to AN for costs of reversal below $5 million, otherwise AN is the optimal strategy. For an expected cost of reversal of $2 million, the optimal treatment strategy in maximising ENG is to adopt and undertake a trial of 284 patients per arm.

In considering VOI analysis within a jurisdiction, the appropriate decision context has been established where positive while uncertain incremental net benefit is driven by net clinical benefit as comparing the ENG of delaying and trialling versus adopting now with no trial. Trialling while adopting is a consideration, but only where trialling is feasible and ethically undertaken in informed patients while adopting (Eckermann and Willan 2007, 2009, 2013; Willan and Eckermann 2010). Where there is expected net clinical benefit of a new therapy, patients would prefer certainty of adoption outside of the trial setting and, hence, trialling while adopting is often neither feasible nor ethical within a jurisdiction.

### 5.9.1 Comparing AT, AN and DT

Where adopt and trial (AT) is feasible as in the case of ECV locally, as INB is driven by cost savings, AT reduces EVSI with costs of reversal; however, delay and trial (DT) faces additional opportunity costs of delay for patients treated outside the trial setting. Hence, a trade-off arises between the value and cost of trial information in comparing DT and AT, while current evidence is sufficient and adopt and no trial (AN) preferred if expected net gain (value less cost) is not positive for any feasible trial (AT, DT). Consequently optimal decision-making requires joint consideration of research and reimbursement, comparing ENG of designs for:

(i) DT versus AN conditional on opportunity costs of delay; and
(ii) AT versus AN conditional on costs of reversal if AT is feasible (usually not locally).

AN is preferred if ENG is not positive for any feasible trial.

These potential options within a jurisdiction for cases such as the ECV example illustrate that it is important to appropriately understand whether AT is feasible or not, but also costs of adoption, delay and reversal.

### 5.9.2 Distinguishing Between Costs of Adoption, Delay and Reversal

Costs of adoption for new therapies encompass costs of learning by doing, and additional costs of training and equipment, costs which should be included in NB assessment. Where costs of adoption haven't been included in INB assessment, they reduce prior INB, but also shift down the whole INB distribution and hence increase EVSI and ENG of trialling. Costs of reversal include costs of reversing public health messages and unamortised costs of training and equipment which reduce EVSI with AT (while not faced with DT). Opportunity cost of delay arise with DT vs AN for patients treated with standard therapy outside as well as within the trial setting to the extent of positive prior INB, which increase overall costs with DT.

## 5.10 More General Implications for Optimising Joint Research and Reimbursement Decisions

In considering research and reimbursement decisions, applications of value of information methods prior to Eckermann and Willan (2007) in cases of interest with evidence of positive but uncertain net benefit of a new intervention framed the decision context as being whether to undertake a trial or not based on value and costs of information and, if so, a trial of what size (Claxton 1999; Claxton and Thompson

2001, Claxton et al. 2002). A fixed context implicit in identifying optimal trial design had been to adopt the intervention in order to avoid opportunity cost of patients remaining on standard intervention with delay. As a result, previous VOI methods incorrectly:

(i) Separated simultaneous decisions of whether to research and reimburse (adopt);
(ii) Assumed adoption where prior E(INB) > 0, and made statements or drew implications to the effect that uncertainty doesn't affect the decision to reimburse;
(iii) Did not consider optimal trial design for delay and trial (DT), nor the ENG of DT versus AN; and
(iv) Did not allow for impact of cost of reversal on EVSI or optimal trial design with adopting and trialling (AT), implicitly assuming AT was always feasible.

The framework for optimal joint research and reimbursement decisions identified in Eckermann and Willan (2007) established that DT should only be compared with AT if DT is preferred to AN and AT is feasible and preferred to AN. If this were the case, then comparison of DT versus AT should be at the optimal trial design for each. That is, where their respective expected net gain in comparison with AN is maximised allowing for opportunity costs of delay in the case of DT and costs of reversal in reducing EVSI in the case of AT. This framework allows optimal joint research and reimbursement decisions by societal decision-makers in a jurisdiction of interest in cases of interest with evidence from a new intervention or strategy with positive while uncertain INB, as illustrated in this chapter following Eckermann and Willan (2007, 2008a, b) and Eckermann et al. (2010).

This VOI framework for joint research and reimbursement decisions improved on previous VOI approaches in:

(i) Allowing optimal trial design for delay, which was not considered previously in comparing DT versus AT;
(ii) Allowing for costs of reversal in optimal trial design for AT versus AN, where previously expected net gain and trial design for AT versus AN remained the same regardless of reversal costs;
(iii) Comparing AT versus DT at the 'optimal' trial design for each, rather than at the optimal size trial for AT versus AN; and
(iv) For AT with costs of reversal allowing optimal decision-making in the situation where it is better to live with a negative incremental net benefit than incur costs of reversal.

More generally, comparison of ENG for DT versus AN and AT versus AN where feasible acts as a circuit breaker to enable appropriate joint optimisation in considering the simultaneous decisions of whether to trial and/or adopt in comparing viable strategies of AN, DT and AT where feasible.

Adopting early ECV while recruiting for trialling was considered feasible for the case of ECV where advantages of avoiding caesarean section relate to reducing health system costs. However, AT within a jurisdiction is often expected to be infeasible and unethical where the new therapy has positive while uncertain incremental

net clinical as well as overall net benefit, as informed patients prefer certainty of treatment outside trial to chance of new therapy in a trial setting. Hence, feasible joint research and reimbursement options within a jurisdiction for these cases will be restricted to DT versus AN, noting that DT versus AN is still a joint reimbursement and research decision with trialling only possible with delay in that case.

Nevertheless even where AT is infeasible within jurisdiction, AT remains a valuable option in moving beyond within jurisdiction analysis in this chapter to consider optimal global trial design (Chaps. 6 and 7). That is, for promising strategies with positive while uncertain net benefit driven by net clinical benefit, where AT is infeasible within jurisdiction, adoption can be undertaken within a jurisdiction where patient recruitment for translatable trial evidence is undertaken elsewhere as part of a global trial. For such cases, the ECV illustration in this chapter highlights the value to societal decision-makers of global societal decision-maker trials (considered in Chap. 6) being able to avoiding expected opportunity costs of delay, a result which is shown to also extend to manufacturer trials similarly avoiding opportunity costs of delay for societal decision makers and manufacturers alike, in Chap. 7.

In leading into Chaps. 6 and 7, it is worth noting that estimating expected value and cost of trial research within a jurisdiction and associated optimal decision-making and trial design within jurisdiction implicitly assumes new information is derived within but not outside jurisdiction. Where this assumption is appropriately relaxed and value of new information from outside of jurisdiction are considered, then making a side payment to another jurisdiction to increase the size of a new or existing trial becomes potentially optimal. Such side payments may be particularly attractive to jurisdictions where adopt and trial (AT) is optimal, but randomised control trials are infeasible and unethical, due to expected while uncertain net clinical benefit as well as net benefit. More generally, the broader question which consideration of new information outside of jurisdiction points to is: what are optimal decisions and optimally sized trials across jurisdictions?

To address this broader question, we extend the framework for optimal societal decision-maker trials within jurisdiction in this chapter, to consider optimal research design across jurisdictions in Chaps. 6 and 7. Chapter 6 considers optimal global societal decision-maker trials given local decision-making preferences within jurisdictions and Chap. 7 optimal global manufacturer trials and pricing, in each case explicitly allowing for translatability of evidence across jurisdictions as part of optimal global design.

### *5.10.1 VOI Advantages over Frequentist Designs in Enabling Efficient Research Design for Joint Research and Reimbursement Decisions*

Optimal trial design and joint research and reimbursement decision-making require considering the efficiency – expected value relative to costs – of additional research to societal decision-making. For cases of interest with positive while uncertain INB value of information (VOI), methods appropriately employed can allow

jurisdictions to optimise expected value relative to cost of research decisions and associated trial designs given relevant decisions contexts faced (Eckermann and Willan 2007, 2008a, b, 2009, 2013; Eckermann et al. 2010; Willan and Eckermann 2010, 2012).

In contrast, frequentist trial design methods condition on a hypothesis in relation to a primary outcome alone, rather than prior evidence of INB or the expected value or cost of research in informing trial design. Such methods fail to consider the expected value or cost of further evidence to decision-making in light of current decision (INB) uncertainty. Rather, frequentist methods for trial design consider a type I error (the probability hypothesis tested as negative when true, typically 5% by convention) and type II error (the probability hypothesis tests as positive when false, typically 20% by convention) for a 'clinically important' minimum significant difference of a primary effect alone. The level of type I and type II error is arbitrary by convention, while the choice of primary outcome and minimum significant difference (MSD) for power calculation has no necessary relationship with current evidence of INB let alone expected value of trial design. Hence, such frequentist designs do not consider the expected value relative to costs of trial designs for decision-making and associated questions of:

(i) Whether it is worthwhile undertaking further research (expected value greater than cost); and
(ii) The efficient or optimal design of such research – what design maximises expected value relative to cost or return on research investment.

Given arbitrary while conventional levels of type I and type II error, the one place where frequentist trial design frequently often do end up considering direct costs of trial design is where statisticians are asked to back solve a threshold level of minimum significant clinical difference (MSD) in order for the trial size to stay within a research budget. Trialists typically will then subjectively justify the level of such a back-solved MSD on ground such as a clinical 'consensus estimate' or failing that attempt to find an alternative primary outcome with a justifiable MSD that stays within research budget. Such common practices only serve to further reinforce the arbitrary nature of frequentist design in relation to patient and budget-constrained health system population outcomes and decision-making. Such arbitrary considerations are in stark contrast to systematic consideration of net clinical and economic benefit inherent in INB and the expected value relative to cost or efficiency of trial designs to decision-making given current decision (INB) uncertainty with value of information methods.

In contrast to frequentist trial designs, use of value of information principles and methods can directly and explicitly optimise the expected value relative to cost of research to decision-makers appropriately allowing for key decision contexts. By doing so, they can enable efficient as well as robust trial design and decision-making. That is, systematically encourage trial design to optimally inform research decisions in a way that optimises the expected value of research funding by allocating bodies such as the National Health and Medical Research Council (NHMRC) in Australia or the Medical Research Council and Wellcome Trust in the UK

(Australian NHMRC 2009; UK Medical Research Council 2009). However, the key to VOI methods being able to inform better decisions and trial or research designs in practice is that they allow for the relevant decision contexts and what is expected with and without further research (counterfactual considerations) in estimating the expected value relative to cost and ENG across decisions and designs. In this respect, there are some additional issues arising in application of VOI methods within jurisdiction that we now discuss.

## 5.11 Conclusion and Discussion of Broader VOI Methods Issues Arising for Decision-Making Within Jurisdiction

VOI methods can be applied to allow policy and decision-makers to more efficiently design and prioritise healthcare research. In this chapter, we have applied Occam's razor to alternative types of VOI methods to assess their simplicity and ability to inform decision-making in addressing four natural questions:

(i) Is further research for a specific HTA potentially worthwhile?
(ii) Is the cost of a given (e.g. frequentist) research design less than its expected value?
(iii) What is the optimal research design for a specific HTA?
(iv) How can research funding be best prioritised across alternative HTAs?

In addressing these questions, we have shown that population EVPI, while the simplest measure to calculate, does not:

(i) Provide a necessary condition to justify or exclude further research without considering costs of actual research designs
(ii) Have any necessary link with EVSI, ENG or return to research
(iii) Represent an upper bound for future research without considering value across jurisdictions and degree of implementation

Hence, consideration of population EVPI in isolation, while simple, provides neither necessary nor sufficient conditions to inform policy and decision-making in addressing any research decision or trial design optimisation and can be particularly dangerous if used as the basis for addressing research prioritisation. EVPI, in focusing on size of decision uncertainty rather than whether research is worthwhile, can easily result in priority being given to trials with negative ENG and return to investment and/or excluding research with the highest expected return, as illustrated in Table 5.2.

Questions (i) to (iv) have been shown to be fully addressed where VOI methods are applied to maximise expected return to research in estimating EVSI, expected costs and ENG at each potential trial size. The level of complexity required to undertake this is prohibitive using bootstrapping methods for optimising trial design across all possible trial sizes (Ades et al. 2004), but is simplified to a feasible level applying VOI methods with the Central Limit Theorem (CLT). These CLT methods

have also been shown to allow for relevant decision contexts within and across jurisdictions, such as time, option value and opportunity costs of delay and imperfect implementation.

CLT methods both significantly reduce the complexity that bootstrapping methods require to estimate EVSI (Eckermann et al. 2010) and allow optimisation of ENG across trials. Such use of the CLT has also been shown to outperform bootstrapping where sample sizes are small and data are skewed, while asymptotically approximating each other when sample sizes are large (Nixon 2010; Briggs et al. 1999). Hence, CLT methods enable simple robust estimation of EVSI, conditional on proposed trial size and, consequently, optimise ENG (EVSI less expected costs) in relation to proposed trial sample size.

Furthermore, the extension of these methods to allow for important decision-maker contexts, including time, opportunity costs of delaying trials, the option value of delay, the VOI outside of jurisdictions and imperfect implementation, has clarified and simplified the appropriate framework for using VOI to inform optimal research and reimbursement decisions within and across jurisdictions. Optimising ENG under this framework directly addresses the practical requirements for robustly using VOI methods to inform optimal decision-making across decision contexts identified by Spiegelhalter (2004).

Methods applied under the CLT consequently more than satisfy Occam's razor by having clearly defined advantages in enabling simpler estimation of EVSI required to address questions 1–4, allowing for important decision contexts and outperforming alternative methods in modelling INB with small and skewed samples. That is not to say that CLT methods for estimating EVSI are all that is required for every HTA or that they are without some limitations. These methods include all variables inherent within current evidence of the distribution of INB in estimating EVSI, ENG and optimising proposed trial design, assuming the proposed trial structure reflects that of prior summary evidence. As a result, they do not currently allow partial analysis, where the VOI from research on one or a subset of variables from the overall set of variables behind the distribution of INB is considered. Methods to undertake and integrate such partial analyses have been developed with alternative non-parametric methods such as bootstrapping. In using such methods, Ades et al. (2004) demonstrate that inner and outer loops should be employed in integrating additional expected partial evidence to update the overall distribution of INB (and associated decision uncertainty). While significantly more complex, these partial methods can have additional value to decision-makers beyond methods for optimal overall trial design in pointing towards optimal sub-studies in RCTs or, more generally, components of optimal research design for modelled analysis. Nevertheless, such partial methods have been restricted to EVPI rather than EVSI, and appropriate decision frameworks and contexts remain important considerations, which partial methods have yet to address.

While the use of VOI methods based on the CLT for trial-based analysis has been established in Willan and Pinto (2005a, b), Eckermann and Willan (2007, 2008 a, b; 2009; 2013, 2016) and Willan and Eckermann (2010, 2012), Kent et al. (2013) also illustrate CLT methods with modelled cost effectiveness analysis. The

importance of appropriately allowing for key decision contexts is highlighted in such modelled, as with trial-based, VOI analysis.

It is also important to note that new methods for rapid regression EVSI assessment have started to emerge which while enabling non-normal forms for INB currently do not appropriately allow for decision contexts. For example, Andronis and Barton (2016) do not yet appropriately or consistently consider even the most rudimentary decision context – the counterfactual of what happens without further research in estimating EVSI (Eckermann and Willan 2016). Rather, in attempting to extend the static framework suggested by Fenwick et al. (2008) and partialise and separate expected value of perfection (perfect information and implementation) into expected value of implementation and expected value of perfect information, they repeat the same mistake of attempting to assert that imperfect implementation reduces the expected value of information. While implementation with further research is expected to be imperfect when relaxing the unrealistic assumption of perfect implementation, it is *a priori* expected to be less imperfect than without further research, given a relationship between strength of evidence and adoption (Eckermann and Willan 2016; Willan and Eckermann 2010). Hence, adoption is *a priori* expected to improve with strength of evidence which generally leads to EVSI increasing when allowing for imperfect implementation. In effect, when relaxing the assumption of perfect implementation, additional research has value to EVSI from improving implementation with strength of evidence as well as reducing uncertainty.

Andronis and Barton (2016) like Fenwick et al. (2008) fail to consider the counterfactual cases of what happens with and without further research. Fenwick et al. (2008) asserted that moving form perfect to imperfect implementation will reduce value of information, which Andronis and Barton (2016) manage to replicate despite moving from a static framework to EVSI measures where appropriate counterfactual consideration should make the opposite clear (Willan and Eckermann 2010; Eckermann and Willan 2016).

Feasible options locally for the usual case of positive while uncertain net clinical benefit and INB within jurisdiction with imperfect or perfect implementation are restricted to:

(i) Delaying and trialling - making a future adoption decision conditional on new evidence; or
(ii) Adopting the strategy maximising INB now with no further research. A third option to adopt and trial while feasible across jurisdictions is, as we have discussed, usually not feasible in such cases within jurisdiction.

However, Andronis and Barton (2016) in relation to the counterfactual of what happens without further research make EVSI calculations with perfect implementation as though decision-making without further evidence supports the strategy with lowest INB (in their case gemcitabine plus carboplatin with a 680 pound per patient lower NB). As highlighted in Eckerman and Willan (2016) this leads to their EVSI estimate with perfect implementation of 24.99 million pounds being severely overestimated, in undiscounted terms by 19.3 million pounds, with 680 pounds per

patient for 5680 patients per year seen over a 5-year timeframe, or in the order of 18 million pounds with a 3.5% discount rate. Critically, this implies that EVSI with perfect implementation should have been in the order of 6 million pounds and appropriately less than that estimated with imperfect implementation (8.04 million pounds). Hence, while Andronis and Barton (2016) improve on Fenwick et al. (2008) in appropriately considering improved implementation with strengthening of evidence (they considered implementation increasing from 50% to 75% in supporting strategy maximising INB over 5 years), they make the mistake of not considering optimal decision-making with adoption of the strategy maximising INB for counterfactual cases with perfect implementation.

Appropriately allowing for optimising behaviour in counterfactual cases with and without imperfect implementation along with considering the opportunity cost of research investment reinforces that societal decision-making should restrict VOI analysis with new interventions to cases of interest with positive while uncertain INB in allocating budget-constrained societal research funding. Assessing interventions with negative while uncertain INB subjectively differs from those with positive INB as the counterfactual case without research is rejection and *a priori* have less than 50% chance of positive INB, while facing the opportunity cost (best alternative action) of investing in research for promising therapies with positive while uncertain INB and orders of magnitude higher chances of long-term adoption. Similarly, note that health shadow prices (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014) reflecting opportunity cost of budget-constrained reimbursement (best alternative adoption and financing actions) are the appropriate threshold values to apply to incremental effects in INB assessment for any given jurisdiction, as chapter 11 illustrates.

In optimising more generally across research, reimbursement and implementation, strategies to improve implementation of current practice without collecting further evidence (i.e. based on existing evidence of INB) should also be considered (Fenwick et al. 2008; Willan and Eckermann 2010) as in Chapt. 9 and Sect. 12.6. Strategies aimed at improving implementation include knowledge transfer or creating active incentives for providers to maximise net benefit in practice. For example, using the net benefit correspondence theorem in efficiency comparisons or funding mechanisms for providers (Eckermann 2004, Eckermann and Coelli 2013). Such strategies are expected to have value directly in improving use or existing programs and technologies or promising new interventions and strategies, but also indirectly in acting as a shift factor to improve the relationship between strength of INB evidence and implementation.

In relation to the expected value of information when the assumption of perfect implementation is appropriately relaxed, the value of research should increase allowing for optimising behaviour in societal decision making with implementation improvement expected relative to no further research with *a priori* expected greater strengthening of evidence. Hence, more research should be considered alongside pure implementation strategies (Willan and Eckermann 2010; Eckermann and Willan 2016; Grimm et al. 2017), not less research and smaller trials (Fenwick et al. 2008; Andronis and Barton 2016).

EVSI increasing with imperfect implementation is appropriately conservative in applying VOI frameworks to joint research and reimbursement decisions. While imperfect implementation may mean there are higher return investment options from directly improving implementation, it also means further research has additional value where alongside reducing uncertainty further research is *a priori* expected to improve implementation. More generally this highlights that VOI methods need to think carefully about optimal decision-making for counterfactual cases as well as the *a priori* expected interaction between further evidence and implementation. More promising in this respect are the methods of Grimm et al. (2017) who not only appropriately allow for the counterfactual of what happens without further research but also enable for dynamic impacts of evidence diffusion in implementation with and without further research. Nevertheless, the methods they present also need to allow for the myriad of other key decision contexts that CLT methods have been developed for in relation to opportunity cost and option value of delay, global VOI, etc. (Eckermann and Willan 2007, 2008a, b, 2009, 2013; Willan and Eckermann 2010, 2012), that Spieigelhalter (2004) highlighted as necessary to enable VOI to be useful to real decision-making.

Rapid regression methods emerging in journals currently, like VOI methods based on the CLT, provide the potential to optimise ENG. Such rapid regriession methods have the potential to temper the issues of much higher levels of complexity required for non-parametric bootstrap methods, while still allowing for non-normality. However, such method still currently faces issues of not allowing for relevant key decision contexts. Hence, while such new methods provide potential advantages in enabling consideration of non-normal distributions for INB, the impact on overall design of this needs to be compared with order of magnitude effects of decision contexts. That is, rapid regression methods have the potential to improve on the CLT to the extent that INB distributions may be non-normal, but need to earn their stripes in allowing for key decision factors. The bottom line is that they should be compared with CLT methods to consider impacts of not allowing for decision contexts while considering non-normal INB distributions.

Such comparisons appropriately encourage rapid regression VOI methods to develop their consideration of key decision contexts while keeping decision-makers appropriately informed with CLT based VOI methods of the impacts of these contexts, as required to improve decision-making in practice.

## References

Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. Med Decis Mak. 2004;24(2):207–27.

Andronis L, Barton P. Adjusting estimates of the expected value of information for implementation: theoretical framework and practical application. Med Decis Mak. 2016;36(3):296–307.

Arrow KJ, Lind RC. Uncertainty and the evaluation of public investment decision. Am Econ Rev. 1970;60:364–78.

Australian National Health and Medical Research Council. Strategic plan. 2009. Sydney (NSW): Australian Government, 2007–9 [online]. Available from URL: http://www.nhmrc.gov.au/publications/synopses/nh80syn.htm. Accessed 17 May 2010.

Bernanke BS. Irreversibility, uncertainty and cyclical investment. Q J Econ. 1983;98(1):85–106.
Briggs AH, Mooney CZ, Wonderling DE. Constructing confidence intervals for cost-effectiveness ratios: an evaluation of parametric and non-parametric techniques using Monte Carlo simulation. Stat Med. 1999;18(23):3245–62.
Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ. 1999;18:341–64.
Claxton K, Sculpher M. Using value of information analysis to prioritise health research: some lessons from recent UK experience. PharmacoEconomics. 2006;24(11):1055–68.
Claxton K, Thompson K. A dynamic programming approach to efficient clinical trial design. J Health Econ. 2001;20:797–822.
Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence. Lancet. 2002;360:711–5.
Eckermann S. Hospital performance including quality: creating incentives consistent with evidence-based medicine. PhD Dissertation, UNSW, Sydney; 2004. http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/.
Eckermann, S., Coelli, T., Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice, Soc Sci Med 2013; 76: 159–168. http://dx.doi.org/10.1016/j.socscimed.2012.10.020.
Eckermann S, Pekarsky B. Can the real opportunity cost stand up: displaced services the straw man outside the room. PharmacoEconomics. 2014;32(4):319–25.
Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16:195–209.
Eckermann S, Willan AR. The option value of delay in health technology assessment. Med Decis Mak. 2008a;28:300–5.
Eckermann S, Willan AR. Time and EVSI wait for no Patient. Value Health. 2008b;11:522–6.
Eckermann S, Willan AR. Globally optimal trial design for local decision making. Health Econ. 2009;18:203–16.
Eckermann S, Willan A. Optimal global VOI trials: better aligning manufacturer and decision maker interest and enabling feasible risk sharing. PharmacoEconomics. 2013;31:393–401.
Eckermann S, Willan AR. Appropriately estimating EVSI with imperfect implementation: improving practice and reducing uncertainty with appropriate counterfactual consideration. Med Decis Mak. 2016;36(3):282–3.
Eckermann S, Briggs A, Willan AR. Health technology assessment in the cost-disutility plane. Med Decis Mak. 2008;28(2):172–81.
Eckermann S, Karnon J, Willan A. The value of Value of Information: best informing research design and prioritization using current methods. PharmacoEconomics. 2010;28(9):699–709.
Fenwick E, Claxton K, Sculpher M. The value of implementation and the value of information: combined and uneven development. Med Decis Mak. 2008;28:21–32.
Grimm S, Dixon S, Stevens J. 2017. Assessing the Expected Value of Research Studies in Reducing Uncertainty and Improving Implementation Dynamics. Medical Decision Making published online 7 Jan 2017. http://www.sagepub.com/journalsPermissions.nav. DOI:10.1177/0272989X16686766.
Hutton EK, Kaufman K, Hodnett E, Amankwah K, Hewson SA, McKay D, Szalai JP, Hannah ME. External cephalic version beginning at 34 weeks' gestation versus 37 weeks' gestation: a randomized multicenter trial. Am J Obstet Gynecol. 2003;189:245–54.
Kent S, Briggs A, Eckermann S, Berry C. Are value of information methods ready for prime time? Int J Technol Assess Health Care. 2013;29(4):435–42.
Nixon RM, Wonderling D, Grieve RD. Non-parametric methods for cost-effectiveness analysis: the central limit theorem and the bootstrap compared. Health Econ. 2010;19(3):316–33.
Pekarsky B. Trusts, constraints and the counterfactual: reframing the political economy of new drugs. PhD Thesis University of Adelaide. 2012.
Pekarsky BAK. The new drug reimbursement game: a regulator's guide to playing and winning. New York: Springer; 2015.

Raiffa H, Schlaiffer R. Applied Statistics Decision Theory. New York: Wiley Interscience; 1967.
Ruta D, Mitton C, Bate A, Donaldson C. Programme budgeting and marginal analysis: bridging the divide between doctors and managers. BMJ. 2005 Jun 25;330(7506):1501–3. doi:10.1136/bmj.330.7506.1501.
Schlaiffer R. Probability and statistics for business decisions. New York: McGraw Hill; 1958.
Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty: when is there sufficient evidence? Value Health. 2005;8(4):433–46.
Spiegelhalter DJ. Incorporating bayesian ideas into health-care evaluation. Stat Sci. 2004;19(1):156–74.
Tirole J. The theory of industrial organisation, vol. 307-308. Cambridge, MA: MIT Press; 1988.
UK Medical Research Council. Mission statement. London: UK Medical Research Council; 2009 [online]. Available from URL: http://www.mrc.ac.uk/About/Missionstatement/index.htm. Accessed 17 May 2010.
Willan A, Briggs A. The statistical analysis of cost-effectiveness data. Chichester: Wiley and Sons; 2006.
Willan AR, Eckermann S. Optimal clinical trial design using value of information methods with imperfect implementation. Health Econ. 2010;19(5):549–61.
Willan AR, Eckermann S. Expected value of information and pricing new health care interventions. PharmacoEconomics. 2012;30(6):447–59.
Willan AR, Pinto EM. The expected value of information and optimal clinical trial design [published erratum appears in Stat Med 2006; 25: 720]. Stat Med. 2005a;24:1791–806.
Willan AR, Pinto EM. The value of information and optimal clinical trial design. Stat Med. 2005b; 24:1791-1806. Correction: Statistics in Medicine 2006;25:720.

# Chapter 6
# Globally Optimal Societal Decision Maker Trials

## 6.1 Introduction

Within a jurisdiction adopting and trialling (AT) where a new therapy has positive, while uncertain net clinical benefit and overall net benefit is usually infeasible and unethical, as informed patients prefer certainty of treatment outside trial to chance of new therapy in a trial setting. Hence, 'within jurisdiction' feasible options as considered in Chap. 5 for such promising therapies or strategies will often be restricted to delaying and trialling (DT) or adopting now (AN) in optimising joint research and reimbursement decisions. However, in designing optimal trials across jurisdictions, trials don't need to be undertaken within jurisdiction, and hence adopting and trialling in another jurisdiction becomes a feasible (and often valuable) option for promising therapies or strategies of interest with expected positive while uncertain INB. That is, a jurisdiction can adopt and translate trial evidence from other jurisdictions as part of optimising joint research and reimbursement decision and trial design across jurisdictions, and ideally globally, as this chapter considers following Eckermann and Willan (2009).

Indeed, in this chapter it becomes clear that within jurisdiction, VOI analysis is inconsistent to the extent that synthesising all translatable evidence arising external to jurisdiction in estimating the prior distribution for INB recognises all relevant trial evidence retrospectively yet only considers evidence arising within jurisdiction as having prospective value. Evidence arising from publicly available trials is non-rival and non-excludable across jurisdictions. Hence, provided evidence is translatable, evidence arising in one jurisdiction has value across jurisdictions.

Where prospective VOI from trials across jurisdictions is considered, an additional viable option to undertaking a trial within jurisdiction is for a side payment to influence trial design in another jurisdiction. This avoids replication of fixed trial costs and reduces heterogeneity of evidence, the potential for cherry picking of evidence or Frankenstein's monster issues in evidence synthesis that frequently arise with multiple trials. Hence, provided trial evidence translates, a single optimal



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_6

trial across two jurisdictions improves on separate trials within each jurisdiction. Extending this principle across all jurisdictions raises the question: what is the globally optimal trial design?

## 6.2 Expected Value and Costs Across Jurisdictions for Global Trial Design

To address the question of global optimal trial design, we need to optimise the expected value relative to cost of a global trial across jurisdictions, just as we considered expected value and cost of trials within jurisdiction in Chap. 5. In considering the expected value and cost of a trial across jurisdictions, note that, as in Chap. 5, each jurisdiction has a distribution for prior INB and relevant decision contexts. Hence each jurisdiction in assessing whether to delay or adopt as part of a global trial can calculate their local value of information – EVSI of trials conditional on costs of reversal with adoption and translation of global trial evidence (AT) and EVSI less opportunity costs with delay and recruitment of trial patients (DT).

Given trial evidence has the public good characteristics of being non-excludable and nonrival across jurisdictions, VOI for optimal decisions of whether to adopt or delay in each jurisdiction ($j = 1,\ldots, J$) can be summed across jurisdictions. This enables estimating global VOI at any trial size for jurisdictions locally optimal decisions to delay or adopt on the proviso that global trial evidence translates. Global ENG, the global expected value less cost, can then be maximised in allocating trial sample across jurisdictions for locally optimal decisions at any given trial size explicitly allowing for evidence translation across jurisdictions in optimising global trial design. That is, ensuring coverage in translation from jurisdictions who recruit patients as part of a global trial to jurisdictions who adopt and don't recruit while the trial is being undertaken.

Hence, the globally optimal trial design given optimal local decision making is defined by the global trial size ($n$ per arm) as the sum of the set of $n_j's$, the number of patients recruited across jurisdictions ($j = 1\ldots J$) that maximises:

$$\sum_{j=1}^{J} \max\left(\mathrm{ENG}_{Dj}\left(n, n_j\right), \mathrm{ENG}_{Aj}\left(n, n_j\right)\right) - \sum_{j=1}^{J}\left(C_{fj} + 2n_j C_{vj.}\right)$$

Note that the decision to delay or adopt is chosen by each jurisdiction to maximise local ENG, excluding direct trial costs which are shared globally. Where AT is infeasible locally nj's will be 0 for jurisdictions who adopt. To ensure coverage of evidence translation in satisfying conditions for local adoption and globally optimisation, pareto optimal side payments to strategic jurisdictions may be required to move their preference from AT to DT. More generally, jurisdictions should only be willing to AT (adopt while the trial is being undertaken) if evidence translates to their jurisdiction.

## 6.3 Illustrating Methods: Globally Optimal Trial Design (The USA, UK and Australia)

To illustrate optimising trial design across rather than within jurisdiction, we extend the ECV within jurisdiction example for the USA in Chap. 5 to consider optimal trial design and decision making across the USA, UK and Australia with relevant ECV variables for the USA, UK and Australia in Table 6.1.

The expected global population EVSI and expected costs of an optimal global trial across these three jurisdictions are shown in Fig. 6.1, where we initially, as in Chap. 5, assume AT is feasible for the ECV case where INB of early versus late ECV is driven by cost savings of avoiding caesarean sections. Note also that as in

**Table 6.1** Jurisdiction-specific parameters for VOI trial design

| | | US | UK | Australia |
|---|---|---|---|---|
| Annual incidence | $k_j$ | 50,000 | 10,000 | 3000 |
| Patient horizon at baseline[a] | $N_0$ | 1,000,000 | 200,000 | 60,000 |
| Annual accrual rate | $a_j$ | $k_j/100 = 500$ | $k_j/20 = 500$ | $k_j/6 = 500$ |
| Fixed cost[b] | $C_{fj}$ | 500,000 | 500,000 | 375,000 |
| Variable cost[b] | $C_{vj}$ | 1600 | 1600 | 1200 |
| Cost of reversal[b] | $C_{rj}$ | 2,000,000 | 1,000,000 | 500,000 |

[a]Assuming a 20-year time frame
[b]In US dollars

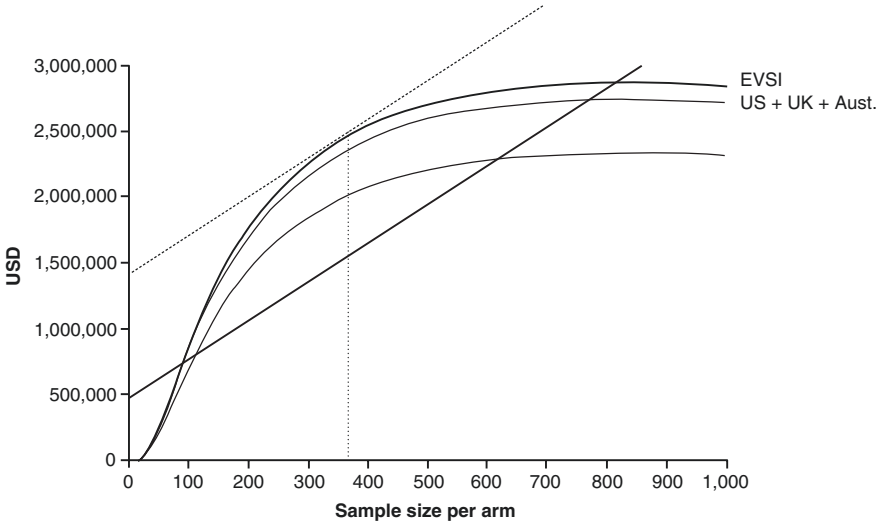


**Fig. 6.1** Globally optimal VOI trial design maximises global ENG – expected value across jurisdictions less shared global trial cost (*Source:* Eckermann and Willan 2009)

Chap. 5 and following Eckermann and Willan (2007), the threshold value in the USA to the health system of avoiding a caesarean section is set as 1000 USD, with associated prior expected incremental net benefit of $68.97 per patient (and variance 3725). Threshold values in the UK and Australia are 1000 USD and 750 USD, respectively, following Eckermann and Willan (2009) with associated expected incremental net benefit $68.97 and $51.73 (variance 2095), respectively.

ENG is maximised with a trial of 372 patients per arm in Australia, where expected trial costs reflect those for Australia (25% less than those in the UK and USA). Table 6.2 compares the EVSI, expected cost and ENG of globally optimal trial design with that from locally optimal trial design, where with AT feasible locally for ECV only, the USA would have positive ENG for a trial of 284 patients per arm, as shown in Chap. 5.

The ENG across jurisdiction increases with an optimal global trial of 372 patients in Australia to $1,142,625 from $724,349 with a locally optimal trial of 284 patients in the USA under the same assumption of being able to AT within jurisdiction. The increase in ENG of $418,246 is mainly due to recognising the increased EVSI across jurisdictions and also in reducing the direct costs and opportunity costs of trialling in optimising choice of where to trial. The increase in ENG implies many Pareto optimal solutions that are possible in funding the trial to improve ENG of each jurisdiction. In Eckermann and Willan (2009), we suggest the simplest and arguably fairest arrangement would see all jurisdictions ENG from that of locally

**Table 6.2** ENG[a] for optimal local versus global trial design for ECV across Australia, the UK and US if AT were locally feasible

| | | Locally optimal ($n1$, $n2$, $n3$) = (284, 0, 0) | Global optimal ($n1$, $n2$, $n3$) = (0, 0, 372) |
|---|---|---|---|
| EVSI | US | 1,789,828 | 2,018,030 |
| | UK | 310,941 | 352,199 |
| | Australia | 51,966 | 59,440 |
| *Total EVSI* | | *2,152,735* | *2,419,669* |
| Opportunity cost | US | 19,586 | 0 |
| | UK | 0 | 0 |
| | Australia | 0 | 19,244 |
| | Total | 19,586 | 19,244 |
| Financial cost | US | 1,408,800 | [b] |
| | UK | 0 | [b] |
| | Australia | 0 | [b] |
| | Total | 1,408,800 | 1,267,800 |
| *Total cost* | | *1,428,386* | *1,287,044* |
| ENG | US | 361,442 | [b] |
| | UK | 310,941 | [b] |
| | Australia | 51,966 | [b] |
| *Total ENG* | | *724,349* | *1,142,625* |

[a]All figures in US dollars
[b]By negotiation

optimal solution increase proportional to their incidence. That is, the additional US$ 418,276 would be 'distributed' across USA/UK/Australia in ratio of 50:10:3 and hence US$ 331,965:US$ 66,393:US$ 19,197.

In practice the USA would pay for the trial to be undertaken in Australia and an additional US$ 31,687 to account for opportunity cost to Australia of $19,244 and Australia's proportional share of additional ENG ($19,197) to the extent that this does not arise naturally from the $7474 in greater EVSI with the larger globally optimal trial. Such an agreement with these side payments is in the interest of the USA given, relative to the locally optimal solution, the global trial both lowers their trial and opportunity cost associated with prior INB for trial patients in the late ECV arm (by US$ 160,586) and increases their expected value of research (by US228,202) with a trial of 372 versus 284 patients per arm.

Note that this initial comparison of globally versus locally optimal VOI trials has assumed that adopting and trialling is feasible within a jurisdiction for ECV where incremental net benefit of avoiding caesarean section is associated with health system cost savings. In the case of ECV, an argument for net clinical benefit of early versus late ECV and hence infeasibility of attempting to trial while adopting locally could be made where women generally indicated a preference for avoiding caesarean delivery. If that were the case, then informed women would be expected to prefer certainty of early ECV outside the trial setting to chance of early ECV on trial, and recruiting to a trial while adopting could become infeasible.

If adopting and trialling (AT) within a jurisdiction were infeasible, then the locally optimal trial design in the case of ECV would be no trial in any of the US, the UK or Australia; with ENG of 0. The globally optimal solution if adopting and trialling in the same jurisdiction were infeasible would be to adopt in the UK and USA and delay with a trial of 339 patients per arm in Australia, with global ENG of $920,590, as shown in Table 6.3.

**Table 6.3** ENG[a] for optimal global trial design for ECV across Australia, the UK and USA with locally feasible delay while trialling

| Global Optimal $(n_1, n_2, n_3) = (0, 0, 339)$ | | |
|---|---|---|
| EVSI | US | 1,976,287 |
| | UK | 344,914 |
| | Australia | 58,210 |
| *Total EVSI* | | *2,379,411* |
| Costs | Opp. cost delay (Australia) | 270,221 |
| | Direct trial cost | 1,188,600 |
| *Total Cost* | | *1,458,821* |
| ENG | US | [b] |
| | UK | [b] |
| | Australia | [b] |
| *Total ENG* | | *920,590* |

[a]All figures in USD
[b]By negotiation

The smaller globally optimal size trial (339 vs. 372 per arm) and ENG associated with the trial ($920,590 vs. $1,142,365) reflects the additional opportunity cost of delay faced with delaying and trialling rather than adopting and trialling in Australia.

However, note that while ENG increases by US$ 418,276 if AT is feasible within jurisdiction (US$ 1,142,625 vs. US$ 724,349), the increase in ENG is US$ 920,590 (920,590 vs. 0) if AT is infeasible. This greater ENG gain from optimal global trial where AT locally within a jurisdiction is infeasible reflects the additional advantage of being able to delay and trial (DT) in one or more jurisdictions and translate evidence to other jurisdictions who adopt (AT).

Hence, limitations locally to feasibly AT become a further advantage globally where the option of jurisdictions to adopt while a global trial is undertaken is a feasible and often valuable option, avoiding opportunity cost of delay for promising new therapies or strategies.

The additional USD 920,590 of ENG with a globally optimal trial versus locally optimal trials where AT is infeasible can as previously be equitable shared across jurisdictions. A 50:10:3 ratio between the USA, UK and Australia resulting in ENG shared as USD 730,627:USD 146,125:USD 43,838. In practice in addition to funding the expected USD 1,188,600 direct costs of the trial to reach this fair distribution of benefit, the USA would need to contribute US$ 57,060 and the UK USD 198,789 towards the opportunity cost of delay faced by Australia (USD 270,221). For each of the jurisdictions the USD 1,976,287, USD 344,914 and USD 58,210 expected value of the trial to the USA, UK and Australia mean it is worth their while for each to support the optimally sized trial in Australia, rather than have no trial locally or indeed any other global trial.

Further, as we show in Chap. 7 for manufacturer funded global trials, such global trial options are valuable both to societal decision makers and manufacturers alike, better aligning their interest for early adoption of promising therapies and globally optimal further evidence. Opportunity cost of delay is avoided by both, while at the same time globally optimal research is collected to best inform decision making across jurisdictions – with evidence translation explicit as part of optimising global design.

## 6.4 Explicitly Addressing Imperfect Translation in Optimal Global Trial Design

In considering whether side payments are required to influence individual jurisdictions to DT for global trial optimisation, note that making translation explicit in optimising global trial design improves on current highly imperfect translation. Currently, in the absence of explicit consideration of evidence translation across jurisdictions, local trials are typically undertaken in the USA but nowhere else. The degree of translatability of trial evidence across jurisdictions depends on the extent to which local populations, practice and relative prices and preferences differ. On each of these counts, the USA is different to the rest of the world. Hence, a locally

optimal trial in the USA represents the one place where trial evidence may have limited translatability and value to the rest of the world. That is, with imperfect translation between the USA and the rest of the world, locally optimal solutions of a trial in the USA and no trial elsewhere can have limited VOI to decision makers outside the USA.

In contrast where evidence translation is explicitly considered as part of global optimal trial design, recruiting patients in the rest of the world is expected on grounds of minimising direct and opportunity costs of delay with trial recruitment but also required to ensure and optimise global evidence translation. Hence, imperfect translation increases the scope for gains in ENG from globally versus locally optimal trial design in allowing for improvement of translation with optimal global design where ensuring translation is explicitly considered.

As a general principle, globally optimal trials which explicitly allow for translation in design will have greater ENG than local trials unless there is no translatability anywhere – where locally optimal would become globally optimal.

Optimally designed global trials explicitly allowing for evidence translation in providing a first best solution also identify how suboptimal locally optimal trials are. They overcome market failure from free-rider effects (small trials) and suboptimal spreading of fixed costs (too many trials). That is, one large global trial reduces heterogeneity of evidence across multiple trials and associated cherry picking of evidence and 'Frankenstein's monster' (O'Brien 1996) issues highlighted in Chap. 3, while also increasing expected homogeneity of practice (implementation) within and across jurisdictions. Additionally, global trials can also aid in standardising evidence required by manufacturers across jurisdictions, while enabling higher-quality evidence to inform regulators and decision making across jurisdictions.

## 6.5 Global Trials for Existing Technology

Robustly evaluating existing programs and technology would benefit from RCT evidence that cannot be feasibly or ethically collected within jurisdictions where it has already been adopted. Within jurisdictions where a strategy, therapy, program or existing technology has already been adopted, trialling will usually be infeasible and evaluation limited to observational evidence from practice – one arm or selected patients by arms. However, a global trial allows feasible trial recruitment in jurisdictions where the program or technology has not already been adopted and hence enables robust evidence of relative treatment effect. Consequently, societal decision maker global trials can also enable robust evaluation of existing programs as well as new technologies in jurisdictions who AT as part of a global trial, something we return to explore in greater detail in Chap. 11. The ability to robustly evaluate existing programs with global trials is suggested as particularly valuable in establishing optimal investment options with current technology and the health shadow price and opportunity cost of investing in new technology as explored in detail in Chap. 11, following Pekarsky (2012, 2015) and Eckermann and Pekarsky (2014).

## 6.6 Conclusion: Optimal Global Trial Design as First Best Solution

The bottom line is that optimal global trial design provides a first best solution, increasing ENG *c.f.* local trials by:

(i) Recognising global VOI;
(ii) Minimising trial cost and heterogeneity of evidence;
(iii) Overcoming technical infeasibility of AT; and
(iv) Explicitly allowing for evidence translation and avoiding market failure with no or too many trials.

Further, while locally optimal trial designs may not provide 'sufficient evidence' for frequentist notions of robust powering, globally optimal trials are generally more efficient and bigger in recognising global value of evidence as well as singular in their findings. Hence, global trials support the large simple trial (or KISS – keep it simple stupid) principles and minimise issues of Frankenstein's monster (Chap. 3) and cherry picking of evidence that plague areas of research with multiple trials. In doing so they also increase strength of evidence and improve expected implementation (Willan and Eckermann 2010). Hence, optimal global value of information trial designs provide a valuable approach to address Frankenstein's monster issues for Bayesian or frequentist alike. They provide a first best solution to joint research and reimbursement decisions with the most robust evidence as well as the most efficient design given expected value and cost of evidence globally. Such optimal global trials with adopting and trialling and evidence translation explicitly allowed for as part of optimising design also ensure coverage with evidence development. That is, ensuring robust translatability of evidence to jurisdictions who adopt early (AT) as part of a global trial while avoiding opportunity costs of delay as part of optimising (Eckermann and Willan 2013).

Further, optimal global trial designs also provide advantages over locally optimal trials with the ability to robustly risk share in jurisdictions who adopt and trial (Eckermann and Willan 2013). Such risk sharing arrangements are considered in detail in Chap. 7 for globally optimal manufacture trials alongside key issues of optimal processes for new technology pricing and research between manufacturers and societal decision making under uncertainty (Willan and Eckermann 2012).

## References

Eckermann S, Pekarsky B. Can the real opportunity cost stand up: displaced services the straw man outside the room. Pharmacoeconomics. 2014;32(4):319–25.

Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16:195–209.

Eckermann S, Willan AR. Globally optimal trial design for local decision making. Health Econ. 2009;18:203–16.

Eckermann S, Willan A. Optimal global VOI trials: better aligning manufacturer and decision maker interest and enabling feasible risk sharing. Pharmacoeconomics. 2013:31;393–401. http://www.ncbi.nlm.nih.gov/pubmed/23529209.

O'Brien B. Economic evaluation of pharmaceuticals. Frankenstein's monster or vampire of trials? Med Care. 1996;34(12 Suppl):DS99–108.

Pekarsky BAK. The new drug reimbursement game: a regulator's guide to playing and winning. New York: Springer; 2015.

Pekarsky B. Trusts, constraints and the counterfactual: reframing the political economy of new drugs. PhD thesis, University of Adelaide. 2012.

Willan AR, Eckermann S. Optimal clinical trial design using value of information with imperfect implementation. Health Econ. 2010;19:549–61.

Willan AR, Eckermann S. Expected value of information and pricing new health care interventions. PharmacoEconomics. 2012;30(6):447–59.

# Chapter 7
# Value of Information, Pricing Under Uncertainty and Risk Sharing with Optimal Global Trial Design

## 7.1 Introduction

In this chapter, we consider optimal processes for joint reimbursement and research decisions between societal decision makers and manufacturers while also allowing for the role of uncertainty in pricing of new technologies, interventions or strategies. This forms the basis for robust risk-sharing arrangements as part of global trials, where it is feasible to adopt and trial. Consequently global trials with appropriately explicit coverage of evidence translation are shown to allow feasibly undertaking robust risk-sharing arrangements in jurisdictions who adopt and trial as part of optimal joint research and reimbursement processes. Such arrangements can mitigate against potential impacts of cost of reversal with adopting and trialling while avoiding opportunity costs of delay for promising new therapies for societal decision makers and manufacturers alike, while having appropriate incentives to create globally sufficient evidence and prices supporting early adoption.

Risk sharing relies on continuing collection of evidence to support meaningful contracts with respect to future contingencies. Hence, it is natural to consider risk sharing in concert with VOI methods when addressing what evidence can and should be collected to inform such arrangements. Robustly informing risk-sharing arrangements requires prospective information regarding the incremental net benefit of promising new interventions or technology and their use in practice. However, where an intervention is adopted in a particular jurisdiction while a trial is undertaken, randomized clinical trials within that jurisdiction are likely to be infeasible and unethical for such interventions in the cases where they would be most helpful, where there is current evidence of positive while uncertain incremental net clinical benefit and economic net benefit. Informed patients in these cases would usually be reluctant to participate in a trial, preferring to receive the new therapy with certainty outside the trial setting rather than participate in a trial with a chance of either therapy. In general, adopting and trialling (AT) within a jurisdiction only becomes



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_7

potentially feasible and ethical where positive, while uncertain, INB is the result of lower costs rather than improved net health outcomes.

The bottom line is that adopting and trialling and informing risk-sharing arrangements are usually problematic and infeasible within a jurisdiction for the case of most interest – promising therapies with expected positive while uncertain net clinical benefit and overall net benefit. In addressing such risk-sharing issues within a jurisdiction, optimally designed global trials are shown in this chapter to facilitate trialling post adoption leading to feasible, more complete and robust risk-sharing arrangements that in turn can often mitigate impacts of costs of reversal with adopting and trialling. More generally as Chap. 6 discussed, global trials offer distinct advantages over locally optimal designs and decision making in providing globally optimal evidence in maximising global value relative to costs, including allowing opportunity cost of delay to be avoided in jurisdictions who adopt. As these opportunity costs of delay are avoided by societal decision makers and manufacturers alike with adopting and trialling, optimal global manufacturer trial designs and risk sharing arrangements are also shown to better align societal decision maker and manufacturer interests, for early adoption of promising therapies and robust, translatable evidence coverage across jurisdictions.

## 7.2 Pricing Under Uncertainty

To address optimal processes for risk sharing, we first need to consider adoption and pricing decision between societal decision makers and manufacturers for interventions with uncertain evidence. Hence, we begin following Willan and Eckermann (2012) by considering two interrelated perspectives in processes of adoption and pricing negotiation under uncertainty:

(i) A societal decision making perspective where given current evidence of INB under uncertainty, their maximum acceptable price (threshold) can be determined for adoption now (reimbursement) rather than demanding further evidence – where all other factors being equal (ceteris paribus) a lower threshold is expected with greater uncertainty; and
(ii) A manufacturer perspective, where given a societal decision maker's threshold price, they need to assess if they should propose at that price given current evidence or gather more evidence with an *a priori* expectation of reduced INB uncertainty and an associated increase in the societal decision maker's threshold price under uncertainty.

From a societal decision making perspective where direct cost of trials for their new interventions are born by manufacturers, then societal decision making costs are limited to opportunity costs. Hence, whether there is sufficient evidence becomes a question of whether opportunity costs of trialling outweigh the expected value of sample information.

Now, to consider threshold prices under uncertainty, we separate out the manufacturer's price of the intervention from other incremental costs in order to see the

impact of price on the INB distribution and subsequently the societal decision making value and cost of undertaking further research, that is,

$$\Delta C = \Delta C_{HS} + P$$

where $\Delta C_{HS}$ are health system costs expected with the intervention excluding the interventions price.

This, in turn, leads to incremental net benefit having the expression

$$INB = \lambda \Delta E - \left(\Delta C_{HS} + P\right) = \lambda \Delta E - \Delta C_{HS} - P$$

The other incremental health system costs of the therapy relative to an appropriate comparator, often standard care, may of course be negative, where net cost offsets are expected from associated treatments and net incremental effects. Clearly, these other incremental costs are more likely to be negative to the extent the new therapies price has been removed.

Importantly, this division of incremental costs makes explicit and clear the role of the new therapies price as opposed to other net cost impacts in determining where INB is located. Hence, for example, if, without loss of generality, we consider a low and high price for the intervention as shown in Fig. 7.1, then it is clear that the whole INB distribution shifts to the left with a higher price relative to a low price.

Importantly such moving of the whole INB distribution with change in the intervention price within a jurisdiction is expected to have a large influence on both expected value of further research (EVSI) as well as opportunity costs of delay.

In relation to the expected value of research to societal decision making with alternate pricing, remember from Chap. 5 that the expected value of further trial evidence involves integrating across the expected shrinkage of the tail distribution for INB less than 0. The extent of this tail distribution will be greater with a higher price as shown in Fig. 7.1 and consequently the expected value of research, as per patient and population EVSI, increases with price of the intervention.

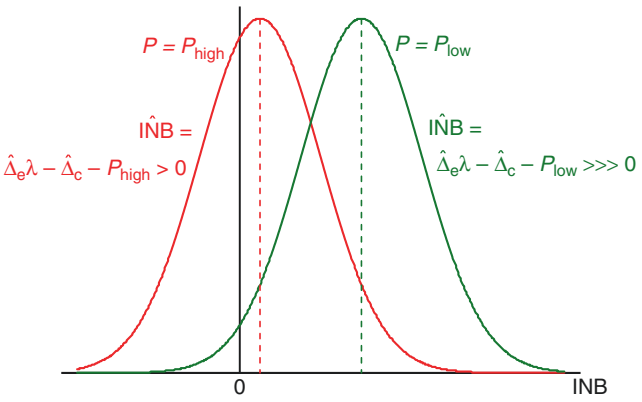


**Fig. 7.1** The impact of price on INB and the expected value and cost of evidence

In relation to expected cost of manufacturer trial research to societal decision making, opportunity cost of delay arise to the extent of the expected positive INB foregone with usual care for patients treated outside as well as in the trial setting.

Hence, with a high price for the new intervention in Fig. 7.1 and subsequent low level of expected INB, the opportunity cost of delay from the manufacturer trial will be less than with a low price and greater expected INB. Expected costs of trialling to societal decision making are hence expected to fall with a price increase.

Combining the expected opportunity cost and expected value of sample information impacts, it is clear that expected net gain of trialling, the expected value less expected cost of trialling, increases with price at any trial size, both because the value of information increases and the opportunity cost of delay while trialling falls. Price via its impact on the location of the whole INB distribution is therefore a key factor in societal decision maker consideration of whether there is sufficient evidence to support adoption now or whether further evidence should be required.

In the extreme case where a new technologies price is high enough that INB is 0, then the opportunity cost of delay per patient not receiving the new therapy is 0, while the expected value of sample information per patient integrating across half the INB distribution where INB is negative. Hence, where expected INB is 0, the ENG to societal decision making from demanding further evidence (a manufacturer trial) will be unequivocally positive, and societal decision makers should demand more evidence and/or a lower price. More generally, in order for societal decision making to support adoption now rather than delaying and trialling as optimal, this implies that expected INB needs to be positive as well as the INB distribution concentrated enough such that the expected opportunity cost of trialling is greater than the expected value.

If the manufacturer's proposed price falls, then the distribution for INB shifts rightward implying societal decision making EVSI falls, while expected INB becomes positive and hence their opportunity cost increases.

If the price is low enough and the resulting expected INB as well as location of the INB distribution more generally are high enough, then current evidence becomes sufficient, at the price where societal decision making ENG of any further manufacturer trialling is just never positive at any given trial size. That is, the societal decision maker threshold price given current evidence is where the price is just low enough that EVSI as a function of trial just never becomes greater than (is tangent to) the expected total cost as a function of trial size (Willan and Eckermann 2012), at $P = P_d^0$ in Fig. 7.2.

Hence, given current evidence, there will be a societal decision maker threshold price which is just low enough associated with a location of the INB distribution just positive enough that decision maker EVSI remains equal to or less than the expected opportunity cost of trialling. Adopting now is preferred at or below this price since there is sufficient evidence at that price, with the expected value of trialling never greater than the expected cost and hence ENG of further research is never positive.

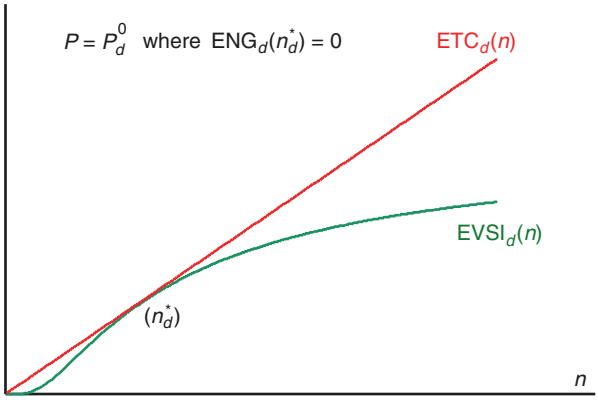


**Fig. 7.2** Threshold price where current evidence is sufficient to adopt now with ENG of delaying and trialling never positive

## 7.3 Illustrating Threshold Pricing Under Uncertainty

To illustrate pricing under uncertainty, consider the case of the CADET-Hp (Canadian Adult Dyspepsia Empirical Treatment – Helicobacter pylori positive) trial (Chiba et al. 2002, 2004). In this trial adding 500 mg metronidazole and 250 mg of clarithromycin to standard 20 mg omeprazole had a 13.71% higher success rate in successful dyspepsia management at 1 year (50.70% vs. 36.99%). For a threshold value of successful dyspepsia management of $Can 500, the prior expected incremental net benefit excluding the price of additional treatment is $Can143.85, given a $Can68.55 value of the 13.71% dyspepsia management improvement and $Can75.30 expected cost savings for other health system treatment costs to 12 months.

However, given INB uncertainty, the threshold price at which it becomes optimal for societal decision maker to adopt now (evidence is sufficient given the expected value relative to expected cost of further research) is $Can 106.53 per patient per year (Willan and Eckermann 2012). While expected INB excluding price is $143.85 per patient, for prices between $Can 106.54 and $Can 143.85, the expected value of further research-given population EVSI of optimal trial size from a societal decision maker's perspective remains greater than the expected opportunity cost of delay given current INB uncertainty.

At a price of $Can 106.53 or lower, the opportunity cost of trialling becomes just equal to or lower than the expected additional value of reducing uncertainty. Figure 7.3 illustrates the optimal-sized trial societal decision making should advise manufacturers' undertake given the expected cost and value of further trialling for prices between $Can 106.53 and $Can 143.06.

At the price threshold, $P_d^0 = 106.53$ or lower, the optimal trial size is 0, and the optimal societal decision is to adopt. For prices above 106.53 to 143.06 trial sizes

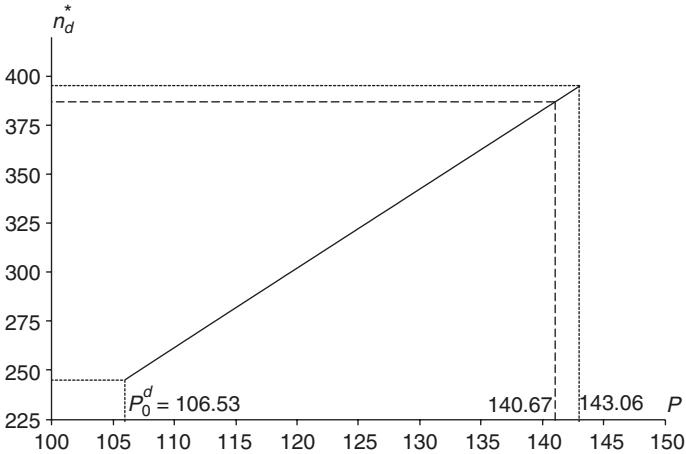


**Fig. 7.3** Societal decision maker threshold price for adoption and optimal delay trial size for prices above this until rejection level (INB < 0) (Source: Willan and Eckermann 2012)

with DT optimal from a societal decision making perspective, recommended trial size range from below 250 to almost 400 per arm.

In general the societal decision making threshold price that results from equipoise between AN and DT within a jurisdiction involves a trade-off between the value of the additional evidence (EVSI) and the opportunity costs of delay while it is collected. Hence, where trials are financed by manufacturers, INB must be greater than 0 for AN to be preferred to DT by the decision maker. This is because for the societal decision maker:

(i) EVSI is an increasing function of the proportion of the INB distribution below 0 and hence also price; and
(ii) Opportunity cost of delay is 0 at a price where INB is 0 and an increasing function with mean INB and, hence, a decreasing function of price.

Now, given the societal decision making threshold price, the question arising for the manufacturer with an expected profit-maximising objective is whether to: (i) Propose adoption at that threshold price; or (ii) Undertake further trial research with the *a priori* expectation of reduced uncertainty and hence a higher expected future societal decision maker threshold price. That is, the societal decision making threshold price given current INB uncertainty, with further evidence is expected to increase towards that where INB is 0 given INB uncertainty is expected to reduce, the extent dependent on the size of trial. Hence, the manufacturer needs to assess the expected value (expected price increase for future patients after evidence is updated) relative to trial costs – both direct trial costs and opportunity cost (lost revenue while trialling from not accepting the Social Decision Maker's (SDM) threshold price for adopting now).

From a manufacturer perspective (noted as company c in notation), the expected value of a trial of size $n$ given current SDM threshold price P is the expected increase

in SDM threshold price from that trial size across the remaining expected patient population:

$$\mathrm{EVSI}_c(n) = B(n)\left\{E\left(P_d^1\right) - P\right\}$$

where $E\left(P_d^1\right)$ is the decision maker's expected post-study threshold price and $B(n)$ is the remaining post study population.

The expected total cost of trialling for the manufacturer includes direct trial financial costs and opportunity costs of delay:

$$\mathrm{ETC}_c(n) = \mathrm{Financial}(n) + D(n)P$$

where $D(n)$ patients would otherwise have received the new therapy at the SDM threshold price with current evidence ($Can 106.53 for the CADET-Hp study).

Now from a manufacturer's perspective, as P increases, their

$$\mathrm{EVSI}_c(n) = B(n)\left\{E\left(P_d^1\right) - P\right\}$$

decreases and

$$\mathrm{ETC}_c(n) = \mathrm{Financial}(n) + \mathrm{D}(n)\mathrm{P}$$

increases. Hence their manufacturer

$$\mathrm{ENG}_c(n) = \mathrm{EVSI}_c(n) - \mathrm{ETC}_c(n)$$

decreases with $P$.

Assuming a manufacturer (company) objective of maximising expected profit where ENG of trialling is positive it becomes optimal to trial, and hence there exists a company threshold price, denoted $P_c^0$ where manufacturer ENG is 0. Below this threshold it is optimal for manufacturers to trial in expectation of a higher price (assuming companies are risk neutral) and equal to or above which it is optimal to submit at the societal decision maker threshold price.

Now, considering interaction between societal decision making and manufacturer threshold prices, if the societal decision making threshold price given current INB evidence (uncertainty) is equal to or greater than the manufacturer threshold price, then the manufacturer should submit at the societal decision maker's threshold price and adopt now. However, if the societal decision making threshold price given current INB evidence (uncertainty) is less than the manufacturer threshold price, then it will be optimal for the manufacturer to trial. Note that this assumes manufacturers are risk neutral and only interested in maximising their expected long-term profits, while if they were risk averse, this would generally lower their threshold price and make them more likely to accept the certainty of a societal decision maker threshold price now.

In the case of the CADET-Hp study within the Canadian jurisdiction, the manufacturer threshold price was $Can 113.06, greater than the societal decision maker threshold price of $Can106.53 (Willan and Eckermann 2012). Consequently, within jurisdiction, the best expected action by the manufacturer (assuming risk neutrality) was to trial. The optimal trial from a manufacturer perspective was 137 patients per arm, which had an expected future threshold price of $140.67 and an ENG of $Can 6,451,162 to the company. Nevertheless, if the manufacturer were somewhat risk averse or interested in certainty with more immediate profits, then they might well accept the SDM threshold price of $Can 106.53 and forgo this expected longer-term profit from the higher expected future threshold price. The *a priori* expectation of a future price of 140.67 and higher profit with a trial of 137 patients per arm, while expected, is not guaranteed, unlike the revenue stream accepting adoption at the SDM threshold price of 106.53 given current evidence.

The manufacturer's risk-aversion considerations in deciding whether to trial or price up to the societal decision maker's threshold highlight trade-offs within a jurisdiction for promising new therapies between their certainty of current revenue versus expectation of future higher prices with trialling while facing opportunity costs of delay.

Nevertheless, note that for manufacturers and decision makers alike for promising new therapies with positive while uncertain net clinical and overall net benefit, they would both like to be able to avoid opportunity costs of delay with early adoption and have further trial evidence in reducing uncertainty for their respective interest. While societal decision makers want to reduce uncertainty in order to avoid bad decisions and manufacturers in order to increase future prices, note that in practice, additional trial evidence is also expected to improve implementation (Willan and Eckermann 2010; Eckermann and Willan 2016). Adopting and trialling can hence be particularly valuable to societal decision makers and manufacturers alike in joint research, adoption, implementation and pricing processes where available as a feasible option.

## 7.4 Pricing Under Uncertainty with Adoption in a Global Trial

The threshold price in any jurisdiction above which it is optimal for the societal decision maker to delay rather than adopt given the current level of uncertainty in INB should not differ in a global trial from that locally. That is the case given there is always the option for jurisdictions to delay locally in the absence of a global trial if the price were above this level.

Hence, the current societal decision maker threshold price in any jurisdiction for a global manufacturer trial should be the same as that with a local manufacturer trial (Willan and Eckermann 2012) given the same current INB distribution and uncertainty. For example, in the case of the CADET-Hp trial in Canada, $Can 106.53 with current evidence, as established in Sect. 7.3. From the perspective of

a manufacturer for cases of interest with a promising new therapy – typically with positive while uncertain incremental net clinical benefit – a global trial with such pricing in jurisdiction adopting enables a current revenue stream, but also increasing strength of evidence and hence expected future implementation as well as pricing under uncertainty (Willan and Eckermann 2010, 2012), when evidence is updated post trial.

Nevertheless, whether it is optimal for societal decision makers to adopt now with a global trial at this threshold price requires considering the impacts of costs of reversal with AT just as they were with AT where feasible with local and global trials in Chaps. 5 and 6. While opportunity costs of delay arise with DT, EVSI with AT is reduced by costs of reversal. Hence, the local decision of whether to adopt or delay and recruit as part of a global trial requires considering the trade-off between the impact of costs of reversal in reducing EVSI with AT and the impact of opportunity costs of delay in reducing EVSI less opportunity costs (local EVSI for global trials) with DT.

Costs of reversal faced when reversing a decision to adopt (while avoided with delay) include direct costs of reversing public health messages, unamortized capital and training costs as well as potentially less tangible costs, such as loss of confidence in adoption processes. Such costs of reversal reduce VOI and EVSI where the new intervention is adopted rather than delayed (Eckermann and Willan 2007, 2008a, b). However, the impact of costs of reversal in reducing EVSI can be diminished with:

(i) Risk-sharing arrangements where incremental net benefit is maintained with prices conditional on the evidence of INB to avoid the need for reversal and;
(ii) Insurance provisions for unexpected outcomes to compensate decision makers for when costs of reversal arise which prices can't compensate for.

Importantly, this implies that if the manufacturer as part of a global trial wants to ensure AT is preferred to DT in any given jurisdiction, they need to both:

(i) Ensure translation of evidence from global trial design with data collected in other jurisdictions; and
(ii) Mitigate costs of reversal with risk-sharing and/or insurance arrangements.

Societal decision makers in any given jurisdiction will prefer DT if evidence does not translate or costs of reversal reduce the expected value of evidence enough. Note that the direct cost of trialling should not influence the decision for each local jurisdiction for a global manufacturer trial as this cost is borne by the manufacturer. This lack of influence on local decisions is also true for optimal global societal decision maker trials considered in Chap. 6 to the extent that as part of arranging equitable ENG improvement from local trials in each jurisdiction, trialling costs are shared across jurisdictions, independent of each local decision to adopt or delay (Eckermann and Willan 2009). In Chap. 6 we showed that global societal decision maker trials designed to maximise ENG across jurisdictions have distinct advantages over any combination of local trials, following Eckermann and Willan (2009).

For manufacturer-designed trials, we have illustrated how joint research, reimbursement and pricing decisions can be determined efficiently within a jurisdiction (Canada) given current evidence of INB and the interaction between optimal manufacturer and societal decision making following Willan and Eckermann (2012). We now combine and extend these approaches to illustrate efficient joint reimbursement, research, pricing and risk-sharing arrangements for global manufacturer-funded trials and consider the potential for further advantages of such arrangements over locally optimal trials. To illustrate such potential, we extend the CADET-Hp optimal local pricing arrangement example from Canada to consideration of a global trial (Eckermann and Willan 2013).

In the CADET-Hp (Canadian adult dyspepsia empirical treatment – *Helicobacter pylori* positive) example, prior incremental net benefit of adding 500 mg metronidazole and 250 mg of clarithromycin is driven by a 13.71% higher success rate in successful dyspepsia management at 1 year (50.70% vs. 36.99%).

Given this expected net clinical benefit, the option to adopt and recruit patients to a trial within Canada is infeasible, while as we have shown, comparison of DT and AN leads to higher expected value to manufacturers of delaying and undertaking a trial of 137 patients per arm. Locally, this trial with an expected post trial threshold value of $140.67 has $Can 6,451,162 higher expected net value to the manufacturer than adoption at a societal decision maker threshold price of $Can 106.53 now.

However, it would be advantageous for both societal decision maker and manufacturer alike if it were possible to adopt at $Can 106.53 now and undertake a trial to avoid opportunity cost of delay with a promising new therapy, improve evidence and implementation (Willan and Eckermann 2010) and revise prices conditional on evidence to mitigate against cost of reversal with adoption.

For societal decision makers, if it were possible to adopt and trial now at a price of $Can 106.53 rather than delay and trial, this would have $Can 37.32 expected INB for each patient treated the promising new therapy given current evidence and result in higher population INB of $Can 5.5 million over the expected time taken for a trial of 137 patients per arm. Hence, societal decision making ENG in Canada would increase by $Can 5.5 million.

From a manufacturer's perspective, adopting and trialling would enable a revenue stream while the trial was undertaken, and avoiding a $15.7 million opportunity cost in Canada that otherwise arises with delay over the time taken to update evidence with trial of 137 patients per arm. Hence, for the manufacturer a global trial, even at the locally optimal-sized trial (137 patients per arm) would increase their ENG by at least $15.7 million in Canada alone. This is an underestimate to the extent that this does not reflect a larger globally optimal trial design and size.

While a trial when adopting within Canada is not feasible, a global trial with evidence collected in other jurisdiction(s) and translated to Canada is feasible as well as risk-sharing arrangements where prices adjust with INB evidence under uncertainty, to mitigate against costs of reversal in jurisdictions who AT. Further, such risk-adjusted pricing would also be consistent post-trial with that expected in

jurisdictions who adopt after delaying during the trial. Indeed, optimally designed global trials enable globally optimal trial evidence while adopting as well as risk sharing, thus avoiding opportunity costs of delay while mitigating against costs of reversal impacts for decision makers and companies alike.

From a manufacturer's perspective, adopting while undertaking a global trial generally has advantages over delaying and trialling in any jurisdiction as it allows a revenue stream while obtaining more evidence. Additional evidence is also *a priori* expected to increase strength of evidence and, hence, ceteris paribus, the future degree of implementation (Willan and Eckermann 2010).

The CADET-Hp illustration has thus far considered a trial of the same size as the locally optimal trial as well as the same price as the local threshold.

However, optimal global trials will generally be larger than locally designed VOI trials because of the greater VOI across jurisdictions to both the societal decision maker (Eckermann and Willan 2009) and the manufacturer (Eckermann and Willan 2013). The increase in optimal size of global trials will also be reinforced from a societal and manufacturer perspective with expected improvement in the degree of implementation from a larger trial and hence associated higher relative expected uptake and revenue per period.

In summary, the CADET-Hp example illustrates that while DT is locally optimal, a global trial with AT and translatable evidence priced at the same local threshold price offers distinct advantages for both the manufacturer and decision maker in avoiding their respective opportunity costs of delay while collecting further trial evidence and mitigating against costs of reversal with feasible risk sharing arrangements.

## 7.5 Circuit Breaker Advantages in Bringing Societal Decision Maker and Manufacturer Interests Closer Together

With the value of information approach to pricing proposed in this chapter, Willan and Eckermann (2012) also provide an efficient HTA application process with appropriate incentives for reducing clinical uncertainty and/or price. Manufacturers should only propose when they have enough evidence to justify their price (Eckermann and Willan 2007; Willan and Eckermann 2012) as otherwise societal decision makers should request further evidence and/or price reduction to be competitive with alternative investment options.

Locally, manufacturers only have an incentive to bring their claim when evidence of INB and its distribution is sufficient to support their offer price (Willan and Eckermann 2012). Globally, they have an incentive to ensure translatable evidence and risk share to avoid costs of reversal in both enabling and encouraging early access with AT. More generally such global optimal VOI trials with explicit consideration of translation across jurisdiction also mitigate potential for bias in design of manufacturer trials, given societal decision makers require trial design that enables

translation of evidence across jurisdictions to enable early adoption under a global trial. That is, by holding the keys to manufacturer benefits from early adoption with AT societal decision makers are in a strong position to ensure evidence translation required for AT but also that biases are minimised in relation to factors such as defining appropriate populations, practice and comparators, as well as covering key outcome measures and resource use.

In relation to such considerations, Eckermann and Willan (2013) show advantages of optimal global trials and evidence translation informing risk sharing with AT are particularly clear for two typical cases of interest in practice, where currently there is evidence of positive while uncertain INB and:

(i) Insufficient evidence to support statistically significant efficacy; and/or
(ii) Inadequate scope and duration of effects and/or resource evidence to fully inform INB assessment.

In case (i), decision makers locally, and particularly where they may have a history of clinically conservative frequentist primary hurdle assessment, may be unlikely to support adoption without a trial regardless of how far prices are reduced beyond the point of equipoise in relation to sufficient evidence. Hence, locally they may prefer DT over AN and incur substantial opportunity cost of delay even if manufacturers offered a price at which the ENG of any further research design with DT relative to AN was substantially negative. A global trial with risk sharing in this circumstance acts as a circuit breaker between economic and clinical decisions by enabling adoption while a globally optimal trial is undertaken (AT).

In case (ii), an advantage of a global trial and risk sharing arises in allowing increased evidence coverage while adopting the new technology, valuable evidence which would either not be collected with AN, a situation unlikely to be satisfactory to payers, or be at the expense of facing opportunity cost of delay with DT.

Consequently an optimally designed global trial can again effectively act as a circuit breaker in enabling feasible AT with earlier access and increased evidence. Hence, provided global evidence translates (necessary for early adoption with AT) and pricing reflects strength of evidence for INB as per local delay, then where the impact of cost of reversal are mitigated with risk-sharing arrangements, global AT will be preferred to DT locally for both manufacturers and decision makers alike in cases of interest with positive while uncertain INB.

Jurisdictions that opt for DT in a global trial will be no worse off than if they chose DT with a local trial in terms of opportunity costs of delay, but share in other benefits from global trials including increased EVSI and ENG and expected higher implementation and faster knowledge transfer with greater strength of evidence. The one key proviso in relation to optimal local decision making with a global trial is to ensure translatability of evidence for a Pareto optimal solution in sharing ENG across jurisdictions. As we saw in Chap. 6, optimal global societal decision maker trial design across jurisdictions may consequently require compensation for opportunity cost of delay particularly to make DT locally optimal in globally strategic jurisdictions. Such compensation as part of risk sharing with global manufacturer

trials considered in this chapter in addition to side payments could take the form of lower prices for patients treated during the trial, or more generally, in these jurisdictions.

Alternatively, if adopt with no trial (AN) had been locally optimal at the decision maker threshold price, then an option to AT across jurisdictions with a manufacturer-sponsored trial would enable the same price while additional evidence is collected. When evidence is updated after a global trial with AT, there is an *a priori* expectation to have better informed decision making and improve levels of implementation (Willan and Eckermann 2010). While such additional evidence and expected reduced uncertainty is valuable to societal decision making in terms of robust decision making and expected implementation, the societal decision making threshold price is expected to increase with such reduction in INB uncertainty. Hence, if AN had been optimal, a larger proportion of the advantages expected from undertaking global manufacturer trials accrue to manufacturers in comparison to societal decision making given the *a priori* expectation of increased prices moving toward where expected INB is 0 with reduced INB uncertainty.

### *7.5.1 Deeper Implications for Implementation and Practice*

The joint nature of research and reimbursement decisions with consideration of ENG for DT versus AN locally and DT versus AN and AT versus AN globally can be refined to more generally consider degree of implementation of new and existing interventions given cumulative strength of evidence (Willan and Eckermann 2010). Choice of DT, AN and AT may act as shift factors on this relationship.

However, the primary shift factor is incentives faced by providers and whether they align with maximising net benefit in practice. In this respect when considering the extent to which implementation of decisions in practice can be improved, note that in Chap. 9, net benefit correspondence theorem methods for efficiency comparison are illustrated that enable net benefit-maximising incentives in practice (Eckermann 2004; Eckermann and Coelli 2013).

It should also be noted that while local trial designs reinforce differences in standards and regulations for evidence across jurisdictions, optimal global trials explicitly support recruitment to allow translation of evidence across jurisdictions and rationalization of regulatory differences. As a general principle, the only case where optimal global trials do not have advantages over optimal local solutions is where globally optimal solutions are the same as locally optimal solutions. This can arise when there is enough current evidence that AN at the current threshold price is the optimal solution globally as well as locally in each jurisdiction. At this point, evidence is globally sufficient. Therefore, empirically optimal global trials should differ from, and improve on, locally optimal trials whenever a trial is optimal.

## 7.6 Bottom Line for VOI Methods

Efficient research design follows a principle of maximising EVSI less expected cost of research (Schlaiffer 1958; Raiffa and Schlaiffer (1967), Claxton 1999). The central limit theorem enables simple calculation of EVSI and ENG for optimisation (Willan and Pinto 2005). Locally optimal research and reimbursement decisions within jurisdiction require joint consideration as, their joint nature is inherent in DT versus AN, while for the limited cases locally where adopting and trialling is feasible (INB is driven by cost savings alone), costs of reversal reduce EVSI with AT versus AN (Eckermann and Willan 2007, 2008a, b; Eckermann et al. 2010). Globally optimal trials for locally optimal decisions which explicitly consider evidence translation in satisfying coverage increase ENG across jurisdictions from that locally and particularly if translatability is currently limited (typically with trials undertaken in the US) and in enabling feasible AT and risk sharing (Eckermann and Willan 2009, 2013; Willan and Eckermann 2012).

Evidence from a global trial more generally acts as a circuit breaker in better aligning societal decision making and manufacturer interests for optimal further evidence and early adoption of promising therapies, and is likely to be particularly valuable in typical cases where either: (i) a clinical evidence hurdle has not been overcome; (ii) evidence coverage can be improved; or (iii) robust evaluation of existing technology adopted in a jurisdiction is required. The first two of these key circuit breaker or game-changing advantages of global trials have been explored at length in this chapter in considering new technologies and their pricing under uncertainty. The third of these advantages of global trials is shown to be an important finding in aiding the research evidence required in relation to better use of existing technology in both informing for the health shadow price of investment and identify optimal investment options for budget-constrained optimisation in Chap. 11 following Pekarsky (2012, 2015) and Eckermann and Pekarsky (2014).

What should be clear from section II (Chaps. 5, 6 and 7) is that for VOI methods to enable robust and optimal joint research and reimbursment decision making and trial design, they need to provide the ability to optimise ENG (EVSI less expected cost) allowing for relevant decision contexts. Of current VOI methods, only those based on the central limit theorem (CLT) have been illustrated in Chap. 5, 6 and 7 to both enable optimisation of ENG across designs and allow for appropriate key decision contexts. That is, enable consideration of key practical factors such as trial recruiting rate, follow-up and analysis time, option value and opportunity cost of delay, costs of reversal faced with adoption, expected implementation conditional on strength of evidence and whether global trial designs and risk-sharing arrangements are an option in optimising the expected value relative to expected cost or expected net gain (Willan and Pinto 2005; Eckermann and Willan 2007, 2008a, b, 2009, 2013; Eckermann et al. 2010; Willan and Eckermann 2010, 2012).

## References

Chiba N, van Zanten SJ, Sinclair P, et al. Treating Helicobacter pylori infection in primary care patients with uninvestigated dyspepsia; the Canadian adult dyspepsia empiric-treatment Helicobacter pylori positive (CADET-Hp) randomized control trial. BMJ. 2002;324:1012–6.

Chiba N, Veldhuyzen van Zanten SJ, Escobedo S, et al. Economic evaluation of Helicobacter pylori eradication in the CADET-Hp randomized control trial in primary care patients with uninvestigated dyspepsia. Aliment Pharmacol Ther. 2004;19(3):349–58.

Claxton K. The irrelevance of inference: a decision making approach to the stochastic evaluation of health care technologies. J Health Econ. 1999;6(17):341–64.

Eckermann S. Hospital performance including quality: creating incentives consistent with evidence-based medicine. PhD Dissertation, UNSW, Sydney. 2004. http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/.

Eckermann S, Coelli T. Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice. Soc Sci Med. 2013;76:159–68. http://dx.doi.org/10.1016/j.socscimed.2012.10.020.

Eckermann S, Karnon J, Willan A. The value of Value of Information: best informing research design and prioritization using current methods. Pharmacoeconomics. 2010;28(9):699–709.

Eckermann S, Pekarsky B. Can the real opportunity cost stand up: displaced services the straw man outside the room. Pharmacoeconomics. 2014;32(4):319–25.

Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16:195–209.

Eckermann S, Willan AR. The option value of delay in health technology assessment. Med Decis Making. 2008a;28:300–5.

Eckermann S, Willan AR. Time and EVSI wait for no patient. Value Health. 2008b;11:522–6.

Eckermann S, Willan AR. Globally optimal trial design for local decision making. Health Econ. 2009;18:203–16.

Eckermann S, Willan A. Optimal global VOI trials: better aligning manufacturer and decision maker interest and enabling feasible risk sharing. Pharmacoeconomics. 2013:31;393–401. http://www.ncbi.nlm.nih.gov/pubmed/23529209.

Eckermann S, Willan AR. Appropriately estimating EVSI with imperfect implementation: improving practice and reducing uncertainty with appropriate counterfactual consideration. Med Decis Making. 2016;36(3):282–3.

Pekarsky BAK. The new drug reimbursement game: a regulator's guide to playing and winning. New York: Springer; 2015.

Pekarsky B. Trusts, constraints and the counterfactual: reframing the political economy of new drugs. PhD Thesis University of Adelaide. 2012.

Raiffa H, Schlaiffer R. Applied statistics decision theory. New York: Wiley Interscience; 1967.

Schlaiffer R. Probability and statistics for business decisions. New York: McGraw Hill; 1958.

Willan AR, Eckermann S. Optimal clinical trial design using value of information with imperfect implementation. Health Econ. 2010;19:549–61.

Willan AR, Eckermann S. Expected value of information and pricing new health care interventions. Pharmacoeconomics. 2012;30(6):447–59.

Willan AR, Pinto EM. The value of information and optimal clinical trial design. Stat Med. 2005;24:1791–806.

# Part III
# Regulating Strategies and Providers in Practice: The Net Benefit Correspondence Theorem Enabling Robust Comparison of Multiple Strategies, Outcomes and Provider Efficiency in Practice Consistent with Net Benefit Maximisation

The methods presented in Parts I and II and each of Chaps. 2, 3, 4, 5, 6 and 7 have been for two-strategy comparisons where a new strategy, intervention or technology is compared with a single comparator, ideally representing current usual care and where analysis was restricted to one effect domain (ideally QALYs). This was the case whether estimating incremental net benefit of individual or community interventions, strategies or programmes (Chaps. 2, 3 and 4) or in identifying optimal trial design and decision making locally and globally with value of information methods given current INB uncertainty in Chaps. 5, 6 and 7. Such two-strategy comparisons reflect typical trial design (intervention vs. control) and simplify modelled analysis to a comparison of a new versus existing strategy. However, there are many situations where two strategies and one domain of effect comparisons are inadequate and lead to overly partial or reductionist analysis and more generally fail to satisfy coverage principles in relation to the scope of potential strategies and effects considered.

In reality for many health-care settings, there are multiple potentially optimal pathways, strategies and modalities of care in practice that should be considered, as well as settings where multiple domains of effect should be considered. This is the case whether in undertaking health promotion, prevention, treatment or palliative care (where multiple domains of interest other than QALYs are key as highlighted in Chap. 4) or comparing community programmes, screening programmes, diagnostic or genetic tests. It should therefore be clear that in satisfying coverage and comparability principles introduced in Chap. 1, we need methods that can robustly inform societal decision making when comparing the joint cost and effects or net benefit under uncertainty of:

(i) Multiple strategies (Chap. 8);
(ii) Multiple health-care providers and their efficiency (Chap. 9); and
(iii) Multiple effect domains of interest (Chap. 10), particularly important in areas such as palliative care where key domains of palliative preferences not amenable to integration with patient survival include process of death factors in finalising personal and financial affairs, family and carer impacts (distress and care burden), and environment (location and community) of choice for palliative care and death, as the second half of Chap. 4 highlighted.

Each of these considerations is made explicit as part of satisfying coverage and comparability conditions with evidence synthesis and translation and regulation in practice in Fig. p3.1, the optimal decision cycle diagram.

In Chap. 8 methods for robustly comparing cost-effectiveness of multiple strategies under uncertainty are presented that overcome distinct limitations of two-strategy fixed comparator methods. Presentation on the cost disutility plane with flexible axes is shown to overcome fixed comparator limitations of the cost-effectiveness plane to enable robust graphical cost and effect inference with multiple strategies (Eckermann, Briggs and Willan 2008; Eckermann and Willan 2011). The use of net loss statistics and expected net loss curves and frontiers is similarly identified as having distinct advantages to societal decision making associated with appropriately flexible comparators of net loss statistics. These summary measures directly address issues and limitations in attempting to extend two-strategy fixed comparator methods with incremental benefit statistics, INB curves and CEA curves. They enable robust multiple strategy comparison at any plausible threshold value for effects including:

(i) Robust comparison of expected net benefit across multiple strategies with ENL curves;
(ii) Identification of optimal strategies with the ENL frontier in minimising expected net loss (equivalent to maximising net benefit under the net benefit correspondence theorem (NBCT) Eckermann 2004); and
(iii) Representing EVPI across multiple strategies, with the ENL frontier as the expected opportunity loss without perfect information of adopting the strategy minimising ENL (max ENB) under uncertainty.

ENL curves and the ENL frontier are consequently shown to address the primary concerns of societal decision making under the Arrow-Lind theorem (asymptotically risk-averse or somewhat risk-averse preferences) and provide the most useful summary measures in comparing costs-effectiveness of multiple strategies under uncertainty for reimbursement and research decisions.

When relative performance or efficiency measures across multiple providers or health systems in practice are considered, we would similarly like to be able to identify net benefit maximising peers and the relative efficiency and sources for improvement of other providers. In Chap. 9 the methods in Chap. 8 for robust multiple strategy comparison are shown to naturally extend to multiple provider efficiency comparisons (Eckermann 2004; Eckermann and Coelli 2013). The net benefit correspondence theorem (NBCT) underlying robust comparison on the C-DU plane across multiple strategies in Chap. 8 naturally extends to enable efficiency measures across multiple providers or health systems in practice consistent with maximising net benefit in Chap. 9. While the correspondence itself allows inclusion of quality of care effects in efficiency measurement consistent with maximising net benefit, NBCT coverage and comparability conditions are shown to also provide a robust framework to prevent cost- and effect-shifting and cream-skimming incentives.

Methods introduced in Chaps. 2 and 3 for cost effectiveness analysis relied on a single domain of effect. While the use of quality adjusted life years (QALYs) can in many cases provide appropriate coverage of patient health effects in integrating survival with morbidity and health-related quality of life, they can be a black box in relation to accommodating alternative event rates or preferences across jurisdictions

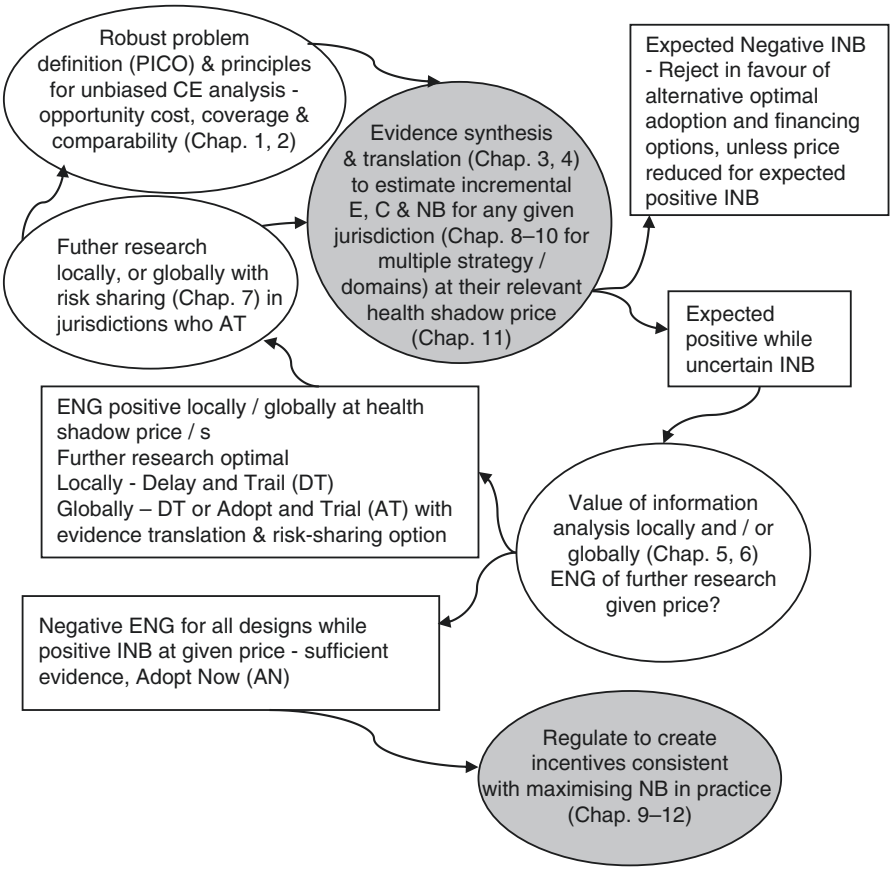


**Fig. p3.1** Optimal decision making cycles for joint research, reimbursement and regulatory processes locally and globally

and more generally do not allow integration of the primary domains of interest to areas such as palliative care patients as highlighted in Chap. 4. In Chap. 10 NBCT methods are shown to also naturally extend to robust multiple effects comparison following McCaffrey (2013) and McCaffrey et al. (2015). Radial properties in C-DU space enable robust comparison of multiple domains under uncertainty, while multiple domain summary measures are shown to naturally extend expected net loss curves and frontiers with one effect and costs to analogous expected net loss planes and surfaces with multiple effects.

## References

Eckermann S. Hospital performance including quality: creating incentives consistent with evidence-based medicine. PhD Dissertation, UNSW, Sydney; 2004. http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/.

Eckermann S, Briggs A, Willan A. Health technology assessment in the cost-disutility plane. Med Decis Mak. 2008;28:172–81.

Eckermann S, Willan A. Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies. Pharmacoeconomics. 2011;29(7):563–77.

Eckermann S, Coelli T. Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice. Soc Sci Med. 2013;76:159–68. http://dx.doi.org/10.1016/j.socscimed.2012.10.020.

McCaffrey N. Modelling joint cost and outcomes uncertainty on the cost-disutility plane – case studies in palliative care. PhD PhD, Flinders University; 2013.

McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Better informing decision making with multiple outcomes cost-effectiveness analysis under uncertainty in cost-disutility space. PLoS One. 2015;10(3):e0115544.

# Chapter 8
# Best Informing Multiple Strategy Cost Effectiveness Analysis and Societal Decision Making: The Cost Disutility Plane and Expected Net Loss Curves and Frontiers

## 8.1 An Introduction to Multiple Strategy Comparison and Limitations of Fixed Comparator Two-Strategy Presentations and Summary Measures

Health economic analysis attempts to inform decision makers of relative effects and costs across potentially optimal alternatives in treating defined patient populations. Comparisons may be between a strategy and single comparator (bilateral) or between multiple strategies (multilateral). In the case of multilateral comparisons, consider, for example, the well-analysed comparison of six alternative strategies for gastro-oesophageal reflux disease (GERD) based on 1-year cost and outcomes in terms of weeks with or without GERD for patients presenting to their physician with endoscopically proven erosive esophagitis (Table 8.1) (Goeree et al. 1999; Briggs et al. 2002; Eckermann et al. 2008).

Multiple strategy cost effectiveness analysis for this example was first considered by Goeree et al. (1999) and then Briggs et al. (2002) on the incremental cost effectiveness plane (weeks without GERD) with a fixed comparator and more recently Eckermann (2004), Eckermann et al. (2005, 2008) and Eckermann and Willan (2011) on the cost-disutility plane with flexible axes (additional weeks of GERD to 1 year relative to the most effective strategy in each replicate).

Such multiple strategy comparisons are becoming increasingly important with diagnostic and genetic testing options as well as treatment of populations across multiple modalities and different strategies for combinations of therapies. Robust methods for multiple strategy comparisons that enable joint comparison of relative costs, effects and net benefit of multiple strategies against each other are required to satisfy coverage and comparability principles, and enable unbiased evidence-based cost effectiveness related decision making. Ideally presentation and summary measures of cost effectiveness evidence that directly inform societal decision making just as they were for two-strategy comparison. For two strategy comparisons, robust methods for joint presentation of costs and effect on the incremental cost



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_8

**Table 8.1** Six alternative strategies for gastro-oesophageal reflux disease (GERD)

| Strategy | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| Cost per patient | 688 | 1088 | 660 | 807 | 747 | 957 |
| Weeks without GERD (to 1 year) | 44.10 | 47.14 | 41.45 | 39.33 | 45.82 | 46.42 |
| Weeks with GERD (to 1 year) | 7.90 | 4.86 | 10.55 | 12.67 | 6.18 | 5.58 |

*A* Intermittent PPI, *B* maintenance PPI, *C* maintenance H2RA, *D* step down maintenance PA, *E* step down maintenance H2RA, *F* step down maintenance PPI

effectiveness plane (relative to the fixed comparator at the origin) and summarising of cost effectiveness evidence with the INB curve (relative to the fixed comparator) and the CEA curve were identified in Chap. 2.

However, comparison of more than two strategies is conceptually and practically very different to two strategy, fixed comparator presentation of incremental costs and effects and cost effectiveness on the cost effectiveness plane and with incremental net benefit and CEA summary curves (Eckermann, Briggs and Willan 2008; Eckermann and Willan 2011). For bilateral (two strategy) comparisons, the appropriate comparator in maximising effect, minimising cost or maximising net benefit is simply the other strategy. This enables a fixed comparator as the origin in the C-E plane and for INB or CEA curves whether considered at expected values or under uncertainty, in any replicate or at any given threshold value. For multilateral comparisons (comparing more than two strategies), the appropriate comparator for any strategy in maximising effect, minimising cost or maximising net benefit can change across replicates, as well as with the threshold value in the case of net benefit.

To begin to see the problems and conceptual and practical issues such fixed comparator presentation and summary measures pose in attempting to accommodate multiple strategy comparison, we first consider the presentation of GERD evidence on the cost effectiveness plane. Conventionally, the cost effectiveness plane with two strategy comparisons presents evidence for a new therapy relative to current usual care. For multiple strategy comparison, Briggs et al. (2002) recognised that graphical interpretation of expected costs and effects on the incremental cost effectiveness plane is aided where the comparator is set to the excepted lowest cost strategy.

In particular, comparison of expected cost and effects of multiple strategies measuring incremental effects and incremental costs relative to the strategy with lowest expected cost strategy as a comparator enables a best practice efficiency frontier to pass through the origin (Fig. 8.1).

More generally the incremental cost effectiveness plane with the lowest cost strategy as the origin (Briggs et al. 2002) at least enables a starting point for considering some useful principles for comparisons of multiple strategies at their expected values in relation to:

(i) Excluding strategies with ‘extended dominance’ – convex combinations of other strategies have lower cost and greater effect (e.g. for GERD strategies D and F can move in SE direction to linear combination of other strategies).
(ii) An efficiency frontier being formed by remaining non-dominated strategies (e.g. strategies C, A, E and B can’t move in SE direction – reduce their cost and increase effect to any convex combination of other strategies);

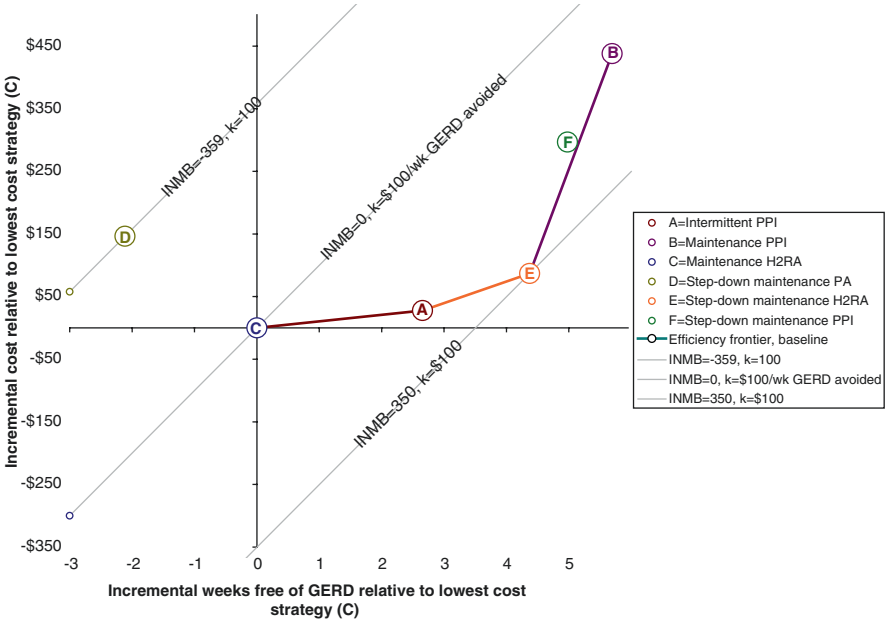


**Fig. 8.1** GERD strategies on the incremental cost effectiveness plane (comparator origin strategy with expected least cost) (Source: Eckermann et al. (2008))

(iii) At a given value for effect, net benefit of strategies can be compared with NB lines reflecting levels of incremental net benefit relative to lowest cost strategy comparator (e.g. $100/week of GERD avoided in Fig. 8.1).

(iv) The appropriate comparator for net benefit consideration changing with threshold values as one moves up the frontier (e.g. for GERD from C to A to E to B).

In relation to (iv), the appropriate comparator at any threshold value is that which maximises NB and hence for each strategy on the frontier reflects the range of threshold values for which they lie on the highest INB line, tangent to the frontier. For example, in the case of GERD strategies on the frontier maximising NB are C from a threshold value of 0 up to $10/week of GERD avoided (the slope of line AC), strategy A from $10 to $36, strategy E from $36 to $243 and strategy B beyond $243.

Nevertheless, presentation on the incremental costs effectiveness plane where performance improves in a south-east direction (cost reduces and effects increase), leads to unbounded consideration of net benefit and more generally poses distinct limitations for informing multiple strategy comparison for analysts and decision makers alike. The south-east direction for identification of dominance and performance improvement inherent in the CE plane does not permit radial properties in contracting to or expanding from a vertex. These radial properties are required to employ standard efficiency measures in identifying a frontier or relative performance of strategies off the frontier relative to the frontier. Efficiency frontiers on the C-E plane are consequently constructed 'by hand' and without being able to compare or interpret relative performance of strategies off the frontier (Eckermann 2004; Eckermann et al. 2008).

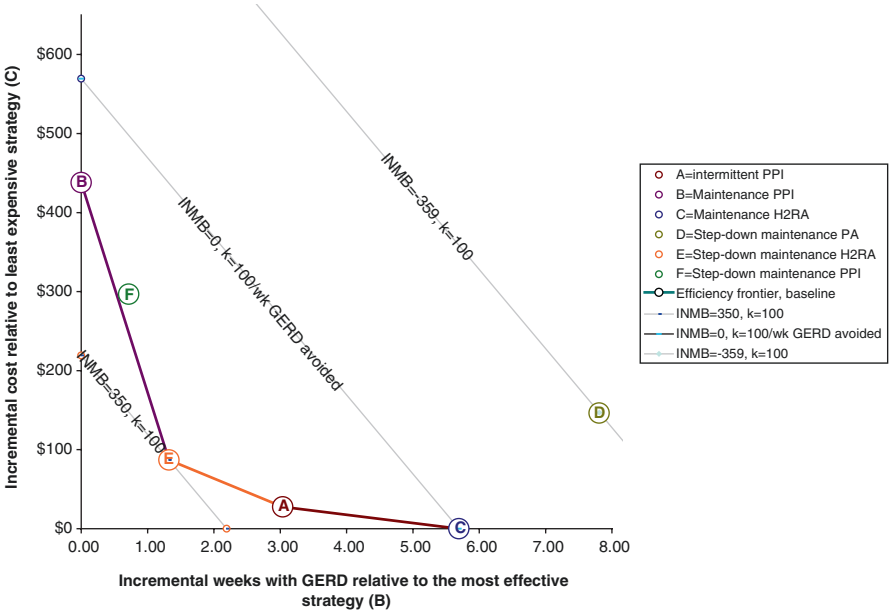


**Fig. 8.2** Frontiers and relative performance with radial properties on the cost disutility plane (Source: Eckermann et al. (2008))

In contrast, comparison on the C-DU plane with flexible axes (Fig. 8.2) provides radial properties with performance improvement where cost and effects framed from a negative or disutility bearing perspective (e.g. mortality, morbidity, weeks with GERD, life years or QALYs lost) reduce and hence both contracting towards the origin or vertex. Combined with flexible axes where cost is measured relative to that of the lowest cost strategy in each replicate and disutility measured relative to the strategy with lowest disutility (highest effect) in each replicate, multiple strategy comparison on the C-DU plane enables:

(i) Technically simpler construction of efficiency frontiers and identification of net benefit-maximising strategies on the frontier and extended dominated strategies off it;
(ii) Degree of dominance (technical inefficiency) to be compared across dominated strategies;
(iii) A bounded comparison of net benefit (iso-net-benefit lines); and
(iv) Cost and effect inference under uncertainty prevented with multiple strategy comparisons on the C-E plane with fixed axes.

Advantages (i) to (iii) of the C-DU plane are explored in greater detail in Chaps. 9 and 10 where they have greater import to comparison and decomposition of provider efficiency following Eckermann (2004) and Eckermann and Coelli (2013) and multiple domain of effect cost effectiveness comparisons following McCafrey et al. (2015), while we focus on (iv) in this chapter following Eckermann and Willan (2011).

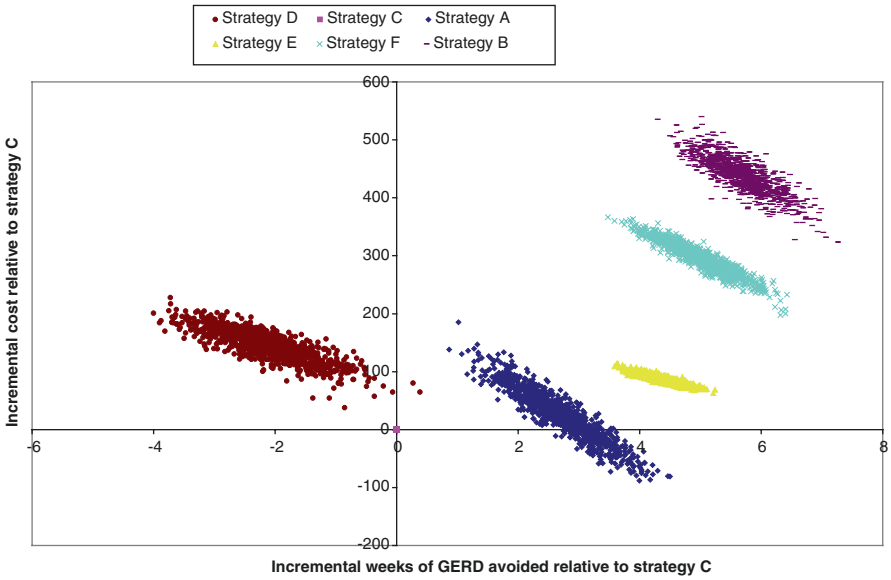


**Fig. 8.3** Multiple strategy cost and effect inference with fixed axes on the CE plane? (Source: Eckermann and Willan (2011))

In relation to cost and effect inference (iv), the fixed nature of all strategies being compared with a single comparator on the CE plane, whether the expected least cost comparator strategy or otherwise, frequently confounds even the simplest graphical inference under uncertainty (Eckermann and Willan 2011). For example, consider on the C-E plane what proportion of replicates each of the 6 GERD strategy maximises effect under uncertainty in Fig. 8.3, constructed from 1000 Mont-Carlo simulation model replicates across these strategies. While strategy B has the highest expected weeks without GERD (47.14 in Table 8.1 and as the highest effect strategy on the frontier in Fig. 8.1), it is not clear which of strategies B, F, E or even A have highest effect in any individual replicate or set of replicates across strategies on the C-E plane in Fig. 8.3. This lack of clarity arises given distributions for each of strategies F, E and even A, considered under uncertainty relative to the fixed expected least cost strategy C, overlap in terms of incremental effect.

This inability of the CE plane to enable even the simplest of graphical inferences reflects the added challenge in comparing multiple strategies of the appropriate comparator changing for cost minimisation and effect maximisation, let alone net benefit maximisation. Indeed, in the case of net benefit the appropriate comparator can change with threshold values for effects, as well as cost and effect evidence across strategies in each replicate, as we later highlight in identifying a solution with the flexible net loss statistic in Sect. 8.3. However, Sect. 8.2 first considers in overcoming more basic fixed comparator problems in considering cost and effect uncertainty and inference on the C-E plane, with flexible axes on the C-DU plane.

## 8.2 Overcoming Fixed Comparator Problems – Multiple Strategy Comparison of Costs and Effects with Flexible Axes on the C-DU Plane

Inference problems arising with fixed comparators on the C-E plane are solved by comparing GERD strategies under uncertainty on the cost-disutility (C-DU) plane using flexible axes of incremental costs relative to lowest costs strategy in each replicate (vertical axis) and incremental disutility relative to highest effect strategy (horizontal axis). As shown in Fig. 8.4, this enables immediate and full graphical effect and cost inference.

On the C-DU plane, the entire distribution of strategy B (maintenance PPI) lies on the horizontal axis with 0 incremental weeks with GERD (note framed from a disutility perspective) relative to the most effective strategy. Hence, strategy B has the lowest number of weeks over 12 months with GERD, or equivalently the highest weeks without GERD, for every one of the 1000 replicates. This stark contrast with the inability to distinguish between effects of strategies B, F, E and A in any given replicate on the C-E plane arises as flexible axes on the C-DU plane ensure that strategies are compared with the appropriate lowest cost and highest effect (lowest disutility) comparator on respective axes and for each replicate.

Similarly, either A or C are the least cost strategies as each have 0 incremental costs relative to the cheapest strategy and hence have replicates lying on the horizontal axis. In general the flexible axes on the C-DU plane, where DU across strategies is measured relative to the lowest DU or most effective strategy in each replicate

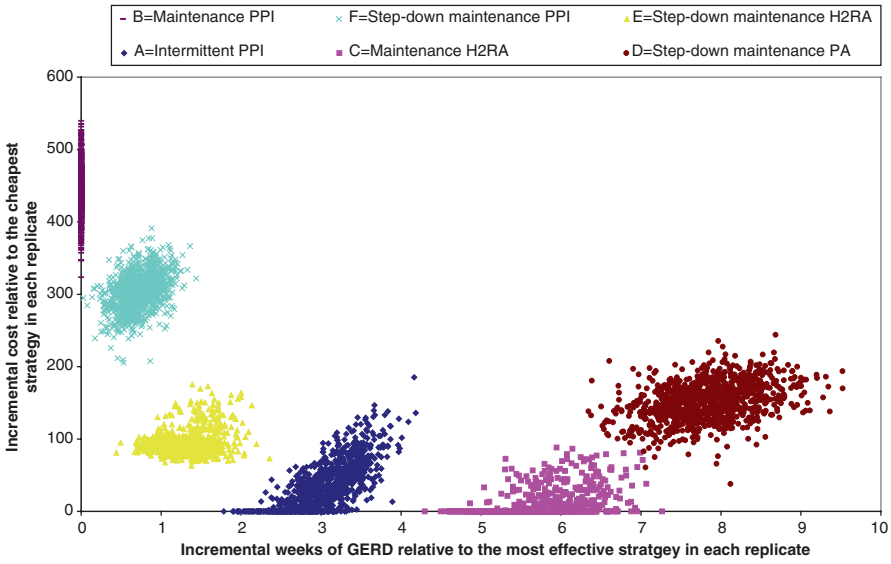


**Fig. 8.4** Multiple strategy cost and effect inference with flexible axes on the C-DU plane (Source: Eckermann and Willan (2011))

and cost across strategies are measured relative to the least cost strategy in each replicate, allow appropriate graphical inference in relation to effects and cost from the proportion of distributions on vertical and horizontal axes. Nevertheless, as on the CE plane, covariance between cost and effects across replicates is still hidden on the C-DU plane, given neither map which replicates and their joint costs and effects link between strategies.

To inform societal decision making in relation to relative net benefit (cost effectiveness) across multiple strategies requires identifying the appropriate net benefit maximising comparator in any given replicate across potential threshold values (as with frontiers for deterministic analysis in Fig. 8.1). However, one needs to first consider what information is required to best inform such societal decision making under uncertainty.

The Arrow-Lind theorem (Arrow and Lind 1970) points to societal decision making (SDM) asymptotically approaching risk neutrality with risk spreading (diversification) across large numbers of government investment decisions and related patient populations, as argued by Claxton (1999). However, Graff, Zivin and Bridges (2002) suggested societal decision making in health care can be somewhat more risk averse than this might suggest, to the extent that patient outcomes may not be completely diversifiable and there are potential effects on private markets of public investment decisions in relation to health technology assessment. Nevertheless, in inform investment decisions under uncertainty, SDM can be characterized across these range of interpretations of the Arrow-Lind theorem for healthcare investment decisions as either asymptotically risk neutral or somewhat risk averse.

If SDM is risk neutral, then the investment objective without further research is to identify strategies maximising expected net benefit (ENB) across plausible threshold values. If SDM is somewhat risk averse, then at any given threshold value, the strategy with highest ENB will still be supported where that strategy also has as high or higher probability of maximising net benefit relative to other potentially optimal strategies. That will be the usual case except potentially over discrete regions of threshold values for effects where strategies vie for ENG maximisation, around which trades-offs can arise between the strategy maximising ENB at any given threshold value and another potentially optimal strategy (with somewhat risk-averse decision making) where they have higher probability of maximising NB at that threshold value.

The key implication is that for summary measures to best inform societal decision making primarily requires showing differences in ENB between strategies, regardless of whether SDM is risk neutral or somewhat risk averse. While expected values alone satisfy risk-neutral societal decision making, somewhat risk-averse decision making can additionally consider incremental probabilities of maximising net benefit across threshold regions but only over discrete threshold regions where trade-offs arise and only between potentially optimal strategies. In this respect, limitations of CEA curves, which only compare probabilities, were recognised by Fenwick et al. (2001) in pointing to the primary need for societal decision making to compare expected net benefit rather than the probability of maximising net benefit across multiple strate-

gies. In Sect. 8.3, we identify expected net loss curves and frontiers (Eckermann et al. 2008; Eckermann and Willan 2011) as providing a first best solution to address the primary need of societal decision making to robustly compare expected net benefit across multiple strategies. Where decision making is somewhat risk averse, this primary need can be supplemented by presenting trade-offs over localised threshold regions where they arise, between incremental ENB and incremental probabilities of maximising NB. However, such probabilities should be informed by relevant bilateral CEA curves in comparison between potentially optimal strategies with the strategy maximising ENB to prevent confounding by other strategies inherent with multilateral CEA curves.

In later discussion, we consider the cost effectiveness acceptability frontier proposed by Fenwick et al. (2001) which presents probabilities of maximising NB for each strategy limited to restricted regions to indicate which strategies maximise ENB at any given threshold value. However, we show that such CEA frontiers are far from a first best solution in informing either primary or potential secondary needs of societal decision making under the Arrow-Lind theorem, facing many problems including: (i) the CEA frontier acting as a black box in lacking the ability to explain ENB optimisation or more generally compare relative ENB of strategies; (ii) decision maker conflation between probabilities presented in CEA curves or frontiers and expected values they want to primarily observe; and (iii) probability confounding in multiple strategy CEA curves or frontiers where societal decision making is somewhat risk averse and considers incremental probabilities alongside incremental ENB in trade-offs between potentially optimal strategies.

To begin identifying a first best solution, we first need to consider how differences in expected net benefit could be presented across multiple strategies, noting that INB curves introduced in Chap. 2 for two strategy comparisons, like the incremental cost effectiveness plane, are restricted to a single fixed comparator. Hence, the fixed strategy comparison that INB curves are based on does not accommodate the need to change comparator to the net benefit maximising strategy across replicates and for alternate threshold values. We need a more flexible comparator statistic than INB for multiple strategy comparisons.

## 8.3 Net Loss Statistics, Expected Net Loss Curves and the Expected Net Loss Frontier

The net loss statistic for any given strategy $i$ from a range of multiple strategies ($i = 1,..,n$) in any given replicate is defined following Eckermann et al. (2008) as

$$NL_i = NB\max - NB_i.$$

This provides the required flexible comparator, appropriately varying the net benefit maximising strategy to compare with across given costs and effects of strate-

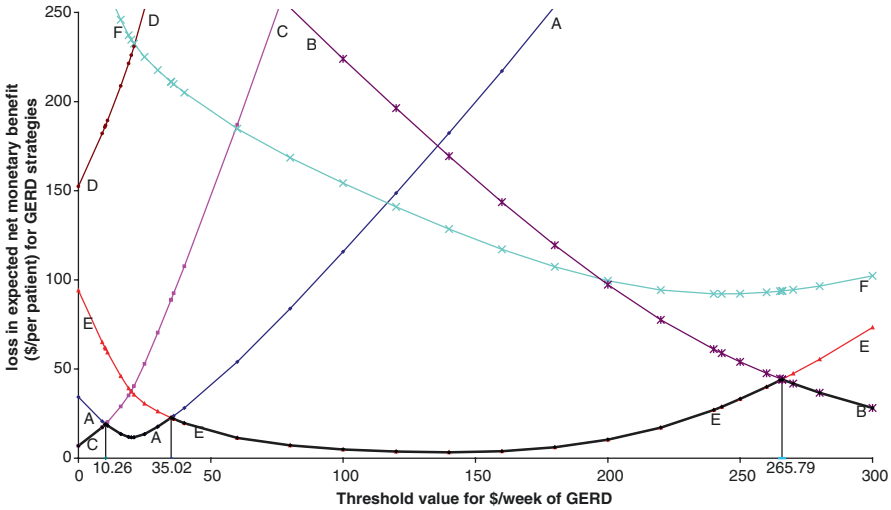


**Fig. 8.5** Expected net loss curves and the expected net loss frontier comparing six GERD strategies (Source: Eckermann and Willan (2011))

gies in each replicate and the range of threshold values for effects ($\lambda$) considered, while being consistent across strategies, retaining the additive separability properties of INB (Stinnett and Paltiel 1997).

Where a strategy is the net benefit maximising strategy in any given replicate at a given threshold value, its NL is 0; otherwise, NL is positive and represents the extent of loss in NB relative to the optimal strategy. Average NL across replicates for each strategy at any given threshold value represents their expected net loss for that threshold value. Mapping these across potential threshold values forms expected net loss curves for each strategy (Fig. 8.5) with the expected net loss frontier as ENL curves lower bound identifying optimal strategies across threshold values.

For GERD strategies ENL curves and their lower bound the ENL frontier show that: strategy C minimises ENL (or maximises ENB) from 0 up to 10.26 per week of GERD avoided; strategy A from 10.26 up to 35.02; strategy E from 35.02 to 265.79 and; strategy B for 265.79 per week of GERD avoided or higher.

Analogous to flexible axes on the C-DU plane, the NL statistic for each replicate and ENL curves for strategies as NL averaged across replicates at any given threshold value provides a flexible comparator systematically ensuring that each strategy in each replicate is compared with the net benefit maximising strategy at any given threshold value. Importantly while appropriately flexible, the consistent comparison with a net benefit maximising strategy in every replicate for any given threshold value also simultaneously provides a consistent benchmark to compare all strategies against within and across replicates. Hence, in comparing across multiple strategies, NL statistics in each replicate and ENL curves summarising evidence across replicates provide the appropriately flexible while consistent comparator. The net benefit maximising strategy in any given replicate as the compara-

tor for NL at any given threshold value overcomes fixed comparator problems of INB statistics. Importantly, this implies the distance between ENL curves at any given threshold value represents the difference in ENB, noting maximising ENB is the same as minimising ENL.

ENL curves are no more difficult to construct than CEA curves. Both identify the strategy maximising net benefit at a given threshold value in each replicate, and then for each strategy, ENL curves calculate expected loss across replicates relative to the NB maximising strategy in each replicate, while CEA curves calculate the proportion of replicates for each strategy compared that they maximise NB. However, unlike CEA curves, ENL curves and their lower bound across threshold values, the ENL frontier both highlight the strategy minimising ENL or equivalently maximising ENB across threshold value and show differences in ENL between strategies at any threshold value, and hence also explain why. Consequently, ENL curves and frontiers clearly and directly inform the primary concern of societal decision makers under the Arrow-Lind theorem. Whether risk neutral or somewhat risk averse they require being informed of differences in expected net benefit or equivalently differences in expected net loss across potential threshold values.

The expected net loss frontier identifies at any given threshold value which strategy minimises expected net loss or equivalently maximises expected net benefit, with distances between ENL curves of strategies represent differences in expected net loss or expected net benefit. Formally this result arises from the net benefit correspondence theorem (NBCT) which shows a one to one correspondence between maximising net benefit and minimising net loss (Eckermann 2004; Eckermann et al. 2008; Eckermann and Coelli 2013, McCaffrey et al. 2015). This result arises in the case of multiple strategy comparison in this chapter, but also more generally, including multiple effect domain comparison (Chap. 10) and comparing efficiency across multiple providers in practice consistent with net benefit maximising quality of care (Chap. 9). Importantly, and as emphasised in Chap. 9 particularly, this result is robust provided NBCT coverage and comparability conditions are satisfied. Comparability and coverage conditions are implicitly or naturally satisfied, where, as in the GERD example, health technology comparisons are based on appropriately randomised control trial evidence in informing relative treatment effects, cover the range of potentially optimal strategies, and cost and effects are compared over common adequate time periods. However, comparability and coverage conditions by necessity need to become more explicit with comparisons in practice such as those required in Chap. 9 across hospitals. Such practice comparisons in order to satisfy coverage and comparability conditions need to standardise for differences in patient risk factors at admission and either model or data link to cover effects and costs over an adequate common time period (e.g. 1 year from date of admission).

In comparing multiple strategies, comparison of relative expected net loss, or equivalently under the NBCT expected net benefit across strategies, is all risk-neutral societal decision making requires in interpreting cost effectiveness evidence. Somewhat risk-averse societal decision making supports the strategy with

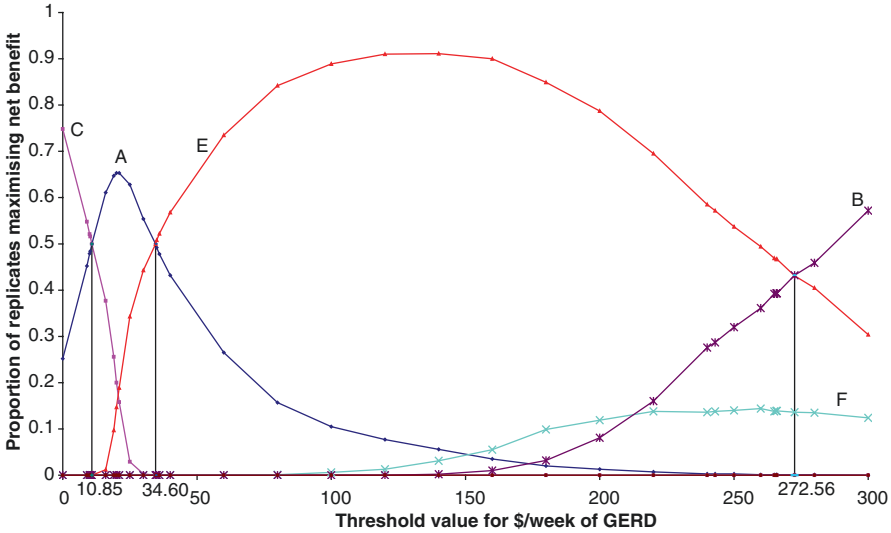


**Fig. 8.6** Multiple strategy CEA curves for GERD strategies

lowest ENL (highest ENB), except for discrete threshold regions where trade-offs between expected value and probability might need to be considered. That is, discrete threshold regions where with lack of symmetry for net benefit distributions between strategies another strategy might have higher probability of maximising net benefit in comparison with the strategy maximising expected net benefit.

ENL curves and the ENL frontier directly inform which strategy is minimizing ENL (maximize ENB) relative to other strategies at any threshold value. Over threshold regions where strategies vie for minimising ENL (maximising ENB), the incremental P(max NB) can be informed by CEA curves over those regions. However, for multiple strategy comparisons this should be restricted to bilateral CEA curves between potentially optimal strategies over such discreet regions given incremental probabilities between these potentially optimal strategies are usually confounded by other strategies with multilateral CEA curves.

For example, consider the case of GERD and comparison between strategies B and E around $265.79 per week of GERD avoided (Fig. 8.5). Below $265.79 E has a lower ENL (higher ENB) and above which B has a lower ENL (higher ENB). Multiple strategy CEA curves across the 6 GERD strategies (Fig. 8.6) show the probability that each strategy maximizes NB in comparison of 6 strategies across the 1000 replicates. These multiple strategy CEA curves suggest that strategy E has higher probability of maximizing NB than B up to $272.56 per week of GERD avoided. Combined with ENL curves, this might suggest a trade-off between higher incremental ENB for B and higher probability of maximizing NB with E over the region from $265.79 to $272.56 per week of GERD avoided, a trade-off region appearing to be $6.77 in size.

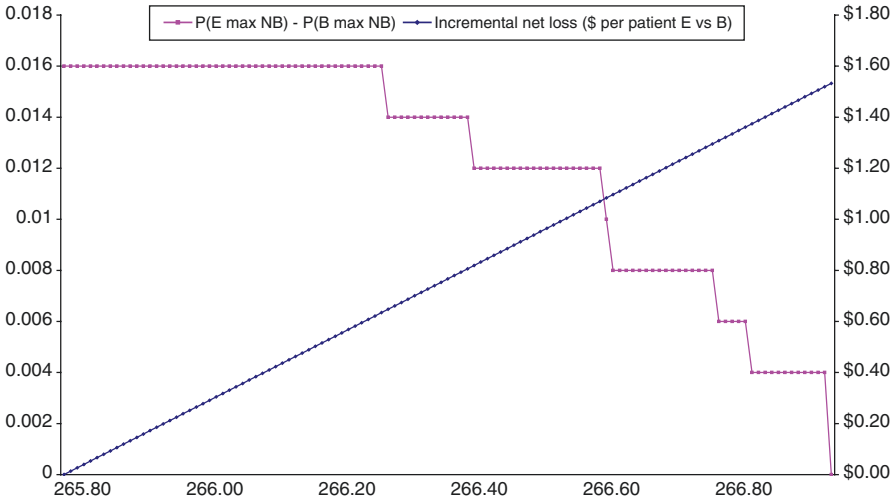


**Fig. 8.7** GERD trade-off region between E having higher probability of maximizing NB than B and lower ENB (higher ENL) (Source: Eckermann and Willan (2011))

However, in comparing potentially optimal strategies B and E, note that the incremental probability of maximizing NB between B and E in Fig. 8.6 in informing such a trade-off for somewhat risk-averse societal decision making is confounded by strategy F over this region. If we remove the confounding impact of F (or more generally other strategies) by restricting CEA curves to bilateral comparison between B and E, then the true picture of the incremental probability of interest between B and E is revealed (Fig. 8.7) with E only having a higher P(max NB) than B in the localized region from $265.79 up to $266.95 per week of GERD avoided (Eckermann and Willan 2011).

Hence, in comparison between potentially optimal strategies, the trade-off region where B has higher ENB and E has higher P(Max NB) is restricted to a discrete region of size $1.16, from $265.79 to $266.95. That is, less than 17% of the $6.77 region suggested with the relationship between B and E confounded by F suggested by multiple strategy CEA curves (Fig. 8.6). Further, in interpreting this already smaller trade-off region, it should be noted that as societal decision making is only somewhat risk averse under the Arrow-Lind theorem, the localized threshold region over which preferences differ from those based on minimising expected net loss is usually expected to be considerably smaller still. For example, consider trade-offs in the region between $265.79 and $266.95 per week of GERD avoided (Fig. 8.7) between strategy B with higher ENB and E with higher probability of maximising net benefit. For somewhat risk-averse decision making under the Arrow-Lind theorem, even at the greatest difference in probability of E maximising net benefit (of 0.016 or 50.8% vs. 49.2% chance of maximising net benefit, between $265.79 and $266.25), a point of indifference with ENB might hypothetically with somewhat risk averse preferences arise at say $0.16 higher ENB for B. Such somewhat risk-averse preferences would imply that only between $265.79

and $265.90 of the trade-off region would E be preferred and diverge from that with risk neutrality.

While the actual discrete region where E might be preferred is an empirical issue where preferences for societal decision making for investment decisions in health care are somewhat risk averse under the Arrow-Lind theorem, the value of higher ENB can be expected to be greater than the value of incremental probability over much of the already discrete trade-off regions. Hence, it should be clear that the impact of societal decision making preferences differing from risk neutrality under the Arrow-Lind theorem is very much at the margins. This is both as ordering of strategies between probabilities and expected values is at the margins and result in very localized threshold trade-off regions, and the impact of those trade-off regions on decision preference will be significantly further diminished by the limited extent of risk aversion. That is, where expected value still largely predominates over probabilities given risk spreading over many decisions and large public health populations impacted.

Finally, an argument could also be made that differences between expected value and probability for any two potentially optimal strategies compared over discrete threshold value ranges may result from random noise of methods employed in estimating INB rather than necessarily being real. That is, if methods were employed where INB distributions between any two strategies were symmetric, then trade-off regions between maximising expected values and considering probability (risk) tradeoffs would disappear altogether. Regardless of where one stands on this argument, it should be clear that societal decision making is predominantly interested in differences in ENB or equivalently ENL in informing investment and reimbursement decisions across multiple strategies based on current evidence.

The bottom line is that ENL curves and the ENL frontier directly inform risk-neutral societal decision making of which strategy is maximising ENB at any threshold region and why.

For somewhat risk-averse decision making, the ENL frontier can be combined with trade-offs over discrete threshold regions where they arise between the strategy maximising ENB and strategies with higher probability of maximising net benefit, informed by relevant bilateral CEA curves to avoid confounding inherent in multiple strategy CEA curves. ENL and relevant bilateral CEA curve evidence is combined in Table 8.2 for GERD strategies.

**Table 8.2** GERD advice for somewhat risk-averse societal decision making

| Optimal strategy | Threshold value ($/wk. GERD) |
|---|---|
| Strategy C | $0 to $10.25 |
| Trade-off between A (higher ENB) and C (higher P(max NB)) | $10.26 to $10.85 |
| Strategy A | $10.86 to $34.60 |
| Trade-off between A (higher ENB) and E (higher P(max NB)) | $34.61 to $35.02 |
| Strategy E | $35.03 to $265.79 |
| Trade-off between B (higher ENB) and E (higher P(max NB)) | $265.80 to $266.95[a] |
| Strategy B | Beyond $266.95[a] |

[a]Trade-offs informed by ENL frontier and bilateral CEA curve to prevent confounding

Further, advantages arise from the ENL frontier in informing societal decision making as the ENL frontier also provides a missing link between research and reimbursement decisions. The ENL frontier simultaneously identifies at any threshold value the strategy maximising ENB (on lowest ENL curve and hence ENL frontier) but also the current per-patient EVPI from adopting that strategy.

## 8.4 The ENL Frontier and EVPI

As we considered in Chap. 5, the current per-patient expected value of perfect information (EVPI) is the expected loss avoided by choosing strategies that maximise NB in each replicate with perfect information rather than adopting the strategy maximising ENB across replicates with current evidence (Eckermann and Willan 2007). The ENL frontier identifies conditional on potential effect threshold values, $\lambda$, the strategy minimising ENL (max ENB) and the expected loss that could be avoided by choosing the optimal strategy in each replicate rather than the strategy maximising ENB (Eckermann, Briggs and Willan 2008; Eckermann and Willan 2011).

Hence, the ENL frontier as the lower bound of ENL curves across strategies compared also represents current EVPI per patient of adopting the highest expected net benefit strategy given current evidence at any threshold value. That is, the lower bound of ENL curves represent the expected value across replicates of opportunity losses relative to the optimum strategy in each replicate as a function of potential threshold values for the strategy maximising expected net benefit or equivalently minimising ENL (Fig. 8.5).

For example, at a threshold value of $US100 per week of GERD avoided, strategy E (step-down maintenance H2RA (histamine H2 receptor antagonists)) has the lowest expected opportunity loss across replicates of $US4.90 per patient (Fig. 8.5). E maximizes net benefit (has 0 net loss or opportunity loss) in 88.9% of replicates, while having $US4.90 expected opportunity loss associated with the remaining 11.1% of replicates where E is not optimal. Hence, at a threshold value of $US100 per week of GERD avoided, E has the lowest EOL across strategies by more than $100 per patient, and the current EVPI per patient is $US4.90.

Considering such per-patient EVPI for multiple strategies comes with all the limitations and caveats of EVPI, identified in Chap. 5 following Eckermann, Karnon and Willan (2010), where the importance of estimating EVSI, comparing with expected cost and optimising ENG across designs, was highlighted. However, note that current EVSI and ENG optimisation methods are restricted to two-strategy comparisons. Hence, considering ENL curves and EVPI across multiple strategies can be useful in focussing on which of the multiple strategies are potentially optimal at any given threshold value and the potential value of further research. For example, in the case of GERD, ENL curves and the ENL frontier (Fig. 8.5) suggest research designs for C versus A around $10 per week of GERD avoided, A versus E around $35 per week of GERD avoided and E versus B around $265 per week of GERD avoided. Around $100 per week of GERD avoided, and indeed over the

range from $100 to $200 per week of GERD avoided, the ENL curves and the ENL frontier suggest E is optimal with or without further research, given the potential value of further research per patient ($3 to $10) is many orders of magnitude less than the expected gain in ENB (minimum reduction in ENL of E compared to any other strategy) with current evidence ($80 to $140). This in turn reflects the very high proportion of replicates in which E is the net benefit maximising therapy over these threshold value ranges for effects as well as the extent of benefit from E in those replicates.

Note that these advantages of the ENL frontier for joint research and reimbursement decision making under uncertainty also arises for two strategy comparisons in representing per-patient current EVPI alongside ENL curves enabling comparison of expected net benefit across strategies at any threshold value. Hence, while use of ENL curves and frontiers is clearly indicated for multiple strategy comparison, their use can also enrich analysis for two strategy comparisons in providing current per-patient EVPI alongside differences in ENB (ENL) across threshold values for effects. For two strategy comparisons, ENL curves provide the same information as INB curves in terms of differences in ENB at any threshold value but simultaneously provide estimation of per-patient EVPI.

## 8.5 Best Presentation and CE Summary Measures to Inform Risk-Neutral or Somewhat Risk-Averse Societal Decision Making with Two and More than Two Strategies

We can now compare and contrast advice for best presentation and summary measures with two and more than two-strategy comparisons to best inform risk-neutral or somewhat risk-averse societal decision making (Table 8.3).

While for two strategy comparisons cost and effect inference is facilitated with distributions presented on the incremental cost effectiveness plane relative

**Table 8.3** Best presentation and summary measures to inform risk-neutral or somewhat risk-averse societal decision making in comparing net benefit of two or more strategies

| Decision making preference/objective | Presentation and summary measures for two strategies | Presentation and summary measures for multiple strategies |
|---|---|---|
| Risk neutral<br>Maximise ENB | C-E plane<br>INB curve<br>ENL curves and frontier | C-DU plane<br>ENL curves and frontier |
| Somewhat risk averse<br>Trade-off ENB and P(max NB) | C-E plane and threshold line<br>CEA curves<br>INB curves and CI<br>ENL curves and frontier | C-DU plane<br>ENL curves and frontier<br>plus trade-offs in regions where they occur between maximising ENB and P(max NB)[a] |

[a]From bilateral CEA curves to avoid confounding in multiple strategy CEA curves

to a single fixed comparator strategy, multiple strategy comparisons are confounded with such a fixed comparator and require flexible axes on the cost disutility plane. Similarly, in summarising cost effectiveness evidence across potentially optimal strategies while bilateral CEA curves relate to a single distribution on the CE plane, incremental probabilities of maximising NB between strategies of interest face confounding by other strategies with more than two strategy comparisons.

Most critically for risk-neutral or somewhat risk-averse societal decision making while INB curves measured relative to a fixed comparator inform societal decision making of relative expected positive or negative expected net benefit across threshold values for effects in two-strategy comparisons, this does not extend to more than two strategy comparison. ENL curves and the ENL frontier are required to identify the optimal strategy and differences relative to this for multiple strategy comparison, while also providing EVPI estimates to inform further VOI analysis. ENL curves and frontiers are consequently the most useful summary measures to start informing joint research and reimbursement decision in two or more than two strategy comparisons. Both risk-neutral and somewhat risk-averse decision making should be interested in whether current evidence is sufficient or further research justified, and if so what is optimal. Methods for undertaking robust VOI analysis between two strategies with EVSI, ENG and return on investment allowing for relevant decision contexts (Chaps. 5, 6 and 7) can be pointed to, as with the GERD example for relevant potentially optimal strategies over relevant decision maker threshold regions for effects, with ENL curves and the ENL frontier.

Table 8.3 clarifies that expected net loss curves and frontiers directly address the primary need of societal decision makers to be informed of differences in expected net benefit (expected net loss) with multiple strategy comparisons as with two-strategy comparisons to clearly identify and explain optimal decision-making. Hence, ENL curves and frontiers are the critical foundation to informing multiple strategy net benefit (cost effectiveness) comparisons for societal decision making regardless of whether societal decision-making is risk neutral or somewhat risk averse. The ENL frontier in simultaneously providing estimation of EVPI per patient also provides distinct advantages more generally over alternative summary measures across two or more than two-strategy comparisons in linking and beginning to jointly address research and reimbursement decisions. Hence, ENL curves and frontiers in general provide valuable joint summary evidence of difference in ENB (ENL) for reimbursement decisions with current evidence and potential value of further research to best inform asymptotically risk-neutral or somewhat risk-averse societal decision making with two or more than two strategies. Where societal decision making is somewhat risk averse, multiple strategy comparison is best informed with ENL curves and frontiers supplemented by trade-offs if they arise over discrete threshold regions between the higher ENB of the strategy with highest ENB (ENL curves and frontier) and the incremental probability of other potentially optimal strategies where they have higher probability of maximising net benefit. Bilateral CEA curves between those two strategies should be considered in such

cases to avoid confounding of incremental probabilities between such potentially optimal strategies by other strategies inherent in multiple strategy curves and the CEA frontier, which we now discuss in greater detail.

## 8.6 Discussion of the CEA Frontier

The CEA frontier (in Fig. 8.8 for the GERD example) suggested by Fenwick et al. (2001) recognised that societal decision makers under the Arrow-Lind theorem primarily need to be informed of expected values rather than probabilities. However, the CEA frontier, in taking a multiple strategy CEA curve presentation (as per Fig. 8.6) and restricting presentation of probabilities of each strategy to indicate the threshold values where that strategy maximises ENB, only presents multiple strategy CEA curve probabilities, while expected net benefit differences are still not presented or compared across strategies.

As a result, the CEA frontier:

(i) Does not explain why a strategy maximises ENB at any given threshold value; while

(ii) The probabilities that are presented (Figs. 8.6 and 8.8) will usually confound the incremental probability between potentially optimal strategies of interest over threshold regions where trade-offs may occur (e.g. between E and B in Fig. 8.7).

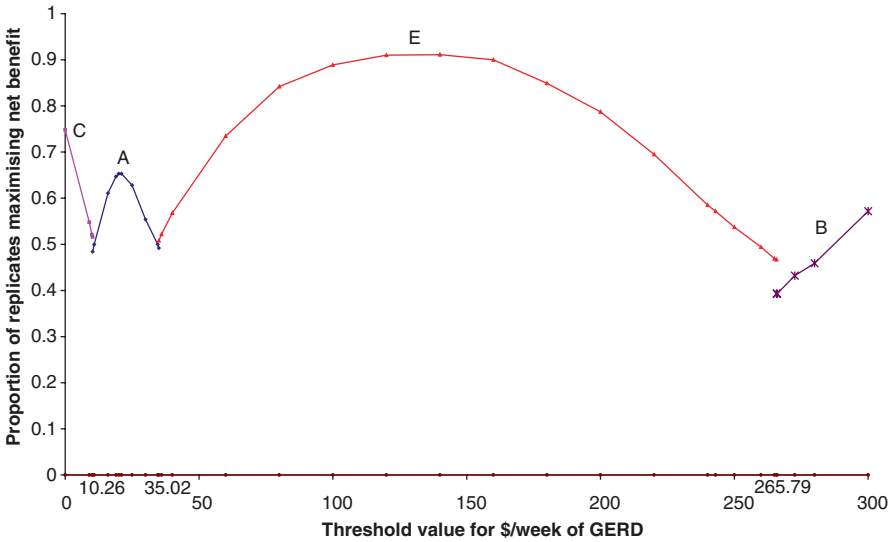


**Fig. 8.8** CEA frontier for GERD strategies (Source: Eckermann and Willan (2011))

In relation to (i), the CEA frontier acts as a black box in indicating why strategies maximise ENB with probability indicators, but not explaining why strategies maximise expected net benefit. Consequently, the CEA frontier can often confuse societal decision making with probability information provided interpreted by societal decision making agencies as the differences in expected net benefit between strategies as that is what they want to be informed of. This black box effect is further reinforced by the probabilities presented in CEA frontiers from multiple strategy CEA curves conflating the relevant incremental probabilities between potentially optimal strategies of interest and the threshold regions over which trade-offs may arise between higher ENB and higher probability of maximising NB. For example conflating of incremental probabilities between E and B around $265.79 per week of GERD avoided in Fig. 8.8, as well as Fig. 8.6.

In contrast ENL curves and the ENL frontier (Fig. 8.5) directly identify optimal strategies in maximising ENB or minimising ENL (ENL frontier) and compare differences in expected net benefit between strategies (ENL curves) conditional on threshold values. Hence, ENL curves and the ENL frontier provide the primary information required for asymptotically risk-neutral or somewhat risk-averse societal decision making. In addition the ENL frontier also represents the per-patient EVPI which compared with differences in ENB with current evidence from ENL curves enables decision makers to start considering whether further research may be valuable and which strategies should be compared with VOI methods.

## 8.7 Conclusion

In conclusion, when comparing the costs and effects of multiple strategies:

(i) The C-DU plane always allows effect and cost inference, the probability of maximising health effects and minimising costs unlike the C-E plane.
(ii) Risk-neutral societal decision making is directly informed of expected differences between strategies by ENL curves and the ENL frontier, unlike CEA curves and the CEA frontier.
(iii) Somewhat risk-averse DM may additionally be informed by trade-offs where they arise over discrete threshold regions between potentially optimal strategies between max ENB (ENL frontier) and the probability of maximising net benefit taken from bilateral CEA curves to prevent confounding by other strategies.
(iv) The ENL frontier also represents per-patient EVPI with current uncertainty at any given threshold value, explicitly linking research and reimbursement decisions. In combination with differences between ENL curves, this can narrow down potentially optimal research designs across threshold values multiple strategies comparing the potential per-patient expected value of future research with differences in expected net benefit with current evidence.

## References

Arrow KJ, Lind R. Uncertainty and the evaluation of public intervention decisions. Am Econ Rev. 1970;60:364–78.

Briggs AH, Goeree R, Blackhouse G, O'Brien BJ. Probabilistic analysis of cost-effectiveness models: choosing between treatment strategies for gastroesophageal reflux disease. Med Decis Mak. 2002;22:290–308.

Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ. 1999;18(3):341–64.

Eckermann S. Hospital performance including quality: creating incentives consistent with evidence-based medicine. PhD Dissertation, UNSW, Sydney; 2004. http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/.

Eckermann S, Coelli T. Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice. Soc Sci Med. 2013;76:159–68. http://dx.doi.org/10.1016/j.socscimed.2012.10.020

Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16:195–209.

Eckermann S, Willan A. Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies. PharmacoEconomics. 2011;29(7):563–77.

Eckermann S, Briggs A, Willan A. The cost-disutility plane: a framework for evidence based medicine in practice. Med Decis Mak. 2005;26:E20.

Eckermann S, Briggs A, Willan A. Health technology assessment in the cost-disutility plane. Med Decis Mak. 2008;28:172–81.

Eckermann S, Karnon J, Willan A. The value of Value of Information: best informing research design and prioritization using current methods. PharmacoEconomics. 2010;28(9):699–709.

Fenwick E, Sculpher S, et al. Representing uncertainty: the role for cost effectiveness acceptability curves. Health Econ. 2001;10(8):779–87.

Goeree R, O'Brien B, Hunt R, Blackhouse G, Willan A, Watson J. Economic evaluation of long term management strategies for erosive oesophagitis. PharmacoEconomics. 1999;16(6):679–97.

Graff Zivin J, Bridges J. Addressing risk preferences in cost-effectiveness analysis. Appl Health Economics Health Policy. 2002;1(3):135–9.

McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Better informing decision making with multiple outcomes cost-effectiveness analysis under uncertainty in cost-disutility space. PLoS One. 2015;10(3):e0115544.

Stinnett AA, Paltiel AD. Estimating CE ratios under second =-order uncertainty: the mean ratio versus the ratio of means. Med Decis Mak. 1997;17(4):483–9.

# Chapter 9
# Including Quality of Care in Efficiency Measures: Creating Incentives Consistent with Maximising Net Benefit in Practice

## 9.1 Overview

Research and reimbursement processes for choosing whether to invest in existing or new health technologies, approaches or programs have focused on net benefit maximisation in evaluating alternative interventions, strategies and programs. However, for net benefit maximising incentives to arise in support of the choice and actual way programs or technologies are used in practice, requires appropriate regulation and incentives for net benefit maximisation in monitoring provider efficiency in practice. That is, efficiency measures of health care providers and institutions such as hospitals allowing for health-care providers and institutions such as hospitals, quality of care and downstream health system impacts consistent with maximising net benefit (Donaldson and Gerard 1993; Eckermann 2004; Eckermann and Coelli 2013). For efficiency measures to create appropriate incentives for health system quality of care in practice:

(i) Quality of care impacts need to be valued in efficiency measures across providers consistent with the appropriate underlying economic objective – maximising net benefit.
(ii) Differences in actual populations cared for across providers (i.e. risk factors of patients at admission) need to be adjusted for to maintain comparability.
(iii) Downstream as well as within-service effects of care (e.g. within admission and post separation effects of quality of care measures) and their associated cost need to be appropriately allowed for in monitoring to maintain appropriate coverage of net benefit impacts.

We show in this chapter that each of these factors is key to appropriate incentives in practice given that perverse incentives are otherwise typically created for cost per service minimising quality of care (e.g. hospital minimum cost per admission quality of care), cream skimming and costs and effect shifting, respectively.

We initially focus on needing to appropriately specify quality of care variables (e.g. waiting times, mortality, morbidity, iatrogenic events, readmission, etc.) in efficiency measures consistent with the appropriate underlying objective of



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_9

maximising health system net benefit as the cornerstone for creating appropriate incentives for quality of care. The net benefit correspondence theorem (NBCT) is identified and illustrated to be the only method for specifying quality variables in efficiency measures which creates incentives for net benefit maximising quality of care (Eckermann 2004; Eckermann and Coelli 2013).

Importantly the NBCT also uniquely provides a robust framework with coverage and comparability conditions to prevent perverse incentives for cost and effect shifting and cream skimming and appropriately qualify analysis where these conditions are not adequately met.

To satisfy coverage conditions and prevent perverse cost and effect shifting incentives, robust quality of care variables in hospitals need to be at a clinical activity (DRG) level with data linkage or modelling to an adequate common time point allowing for post separation health system impacts of care. To satisfy comparability conditions in preventing cream skimming incentives requires standardising cost and effects to control for exogenous within DRG differences between hospital patient population risk factors at admission.

The chapter concludes discussing the NBCT as a generalizable method and its extension to other health care and wider settings, as well as multiple effect domains (Chap. 10) and funding mechanisms (Sect. 12.6). Key Links that the NBCT provides between HTA and practice in jointly addressing optimal research, reimbursement and regulatory decisions are also highlighted.

## 9.2 The Need to Include Quality in Efficiency Measures Consistent with Maximising Net Benefit

Health economics in processes of health technology assessment (HTA) have stressed the importance of jointly comparing the incremental cost ($\Delta C$) and health effects ($\Delta E$) of strategies relative to an appropriate comparator. At a threshold value for effects of care ($\lambda$), this comparison is equivalent to a decision-making objective of maximising incremental net benefit (INB):

$$INB_i = \lambda\left(E_i - E_{\text{comp}}\right) - \left(C_i - C_{\text{comp}}\right)$$

This has been the explicit or implicit economic objective underlying analysis throughout Chaps. 1 to 7, while shown to be equivalent to minimising net loss with multiple stratgey comparisons in Chap. 8.

However, conventional measures of economic efficiency across health-care providers or health systems such as hospitals in practice reflect a cost minimising objective with cost per service based measures. For example, with cost per admission efficiency measures in hospitals (with or without case mix adjustment, Australian Government 2000). Hence, in contrast to HTA, efficiency measures of performance of hospital providers in practice, while including the per admission cost of quality of care, ignore the effects of quality of care. To illustrate such differences, consider comparing 45 NSW hospitals in treating patients at a clinical

activity diagnostic-related group (DRG) level for respiratory infection DRG E62a in Table 9.1, given cost per admission and mortality rate for this DRG.

Presenting this evidence in Fig. 9.1 jointly for costs and mortality rates per admission, what incentives does comparing costs alone create?

**Table 9.1** Cost per admission ($/Ad) and mortality rate per admission (Mort %) for 45 NSW hospitals (Hosp) treating respiratory infection DRG E62a

| Hosp | $/Ad | M% | Hosp | $/Ad | Mort% | Hosp | $/Ad | Mort% |
|---|---|---|---|---|---|---|---|---|
| *1* | $4830 | 40% | *16* | $6199 | 25% | *31* | $5518 | 17% |
| *2* | $9224 | 25% | *17* | *$3858* | *9%* | *32* | $6779 | 27% |
| *3* | $8056 | 8% | *18* | $7411 | 24% | *33* | *$5283* | *3%* |
| *4* | $12,409 | 7% | *19* | $4520 | 12% | *34* | $6977 | 10% |
| *5* | $5123 | 40% | *20* | $6134 | 24% | *35* | $7407 | 24% |
| *6* | $8249 | 6% | *21* | $7484 | 14% | *36* | $5189 | 25% |
| *7* | $4138 | 35% | *22* | $4878 | 26% | *37* | $5820 | 30% |
| *8* | $6000 | 14% | *23* | $5890 | 21% | *38* | $6887 | 23% |
| *9* | $7382 | 13% | *24* | $5296 | 30% | *39* | $6424 | 31% |
| *10* | $6649 | 4% | *25* | $4543 | 21% | *40* | $5921 | 21% |
| *11* | $7545 | 4% | *26* | *$3590* | *17%* | *41* | $5618 | 29% |
| *12* | $8301 | 32% | *27* | $6132 | 6% | *42* | $7057 | 21% |
| *13* | $6052 | 38% | *28* | $7744 | 18% | *43* | $5324 | 34% |
| *14* | $13,128 | 4% | *29* | $5302 | 11% | *44* | $7605 | 27% |
| *15* | $6616 | 10% | *30* | $5920 | 32% | *45* | $6797 | 28% |
| | | | *Industry* | $6332 | 22.4% | | | |

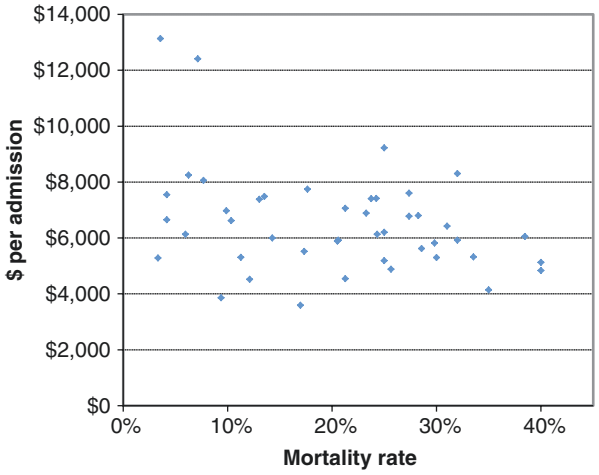


**Fig. 9.1** Cost per admission and mortality rate for DRG E62a in 45 NSW hospitals (Source: Eckermann 2004, "Hospital performance including quality: creating incentives consistent with evidence-based medicine" PhD Dissertation, UNSW, Sydney. http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/)

It is clear that if we only consider cost per admission (the vertical axis in Fig. 9.1) and ignore quality of care (horizontal axis), we make hospitals accountable for the expected average cost of their mix of clinical activities, but not patient quality of care effects such as mortality. Case-mix proponents describe such lack of accountability for patient outcomes as 'clinical neutrality' of case-mix efficiency measures and funding mechanisms (Brook 2002).

However, such partial efficiency measures based on cost per admission alone while creating incentives to minimize cost per admission also create incentives (Eckermann 1994) for:

(i) Minimum cost per admission quality of care;
(ii) Cost-shifting (e.g. high readmission rates); and
(iii) Cream skimming (i.e. choosing less complex patients), to the extent predictive differences are observable between patients within activities compared (e.g. within DRG E62a in Table 9.1 and Fig. 9.1).

In relation to cost shifting (ii), given characteristic incomplete vertical integration of hospitals in health systems (Evans 1981), minimum cost per admission quality of care can more generally be expected to have impacts post separation on higher treatment in other institutional health-care settings, general practice, specialist and aged care services and informal care in non-institutional settings as well as higher hospital readmission rates. The key implication is that minimising cost per admission does not equate to minimum health system costs in considering downstream impacts, let alone maximising health system net benefit in considering the health system cost and value of hospital quality of care (Eckermann 2004).

## 9.3 The Quality of Care Challenge

To create appropriate incentives for quality in practice, economic efficiency measures need to recognise the value as well as cost of quality (Eckermann 1994, 2004; Newhouse 1994). Ideally the value of quality of care should be included in efficiency measures consistent with maximising net benefit to support public policy objectives (Graham 1992). Net benefit from a societal decision making or community perspective is also explicitly or implicitly the objective underlying health economics, health system decision-making and related process whether individual or community-based alternative strategy comparison, health technology or program assessment (Claxton and Posnett 1996; Stinnett and Mullahy 1998; Willan and Lin 2001; Eckermann 2004; Drummond et al. 2005; Willan and Briggs 2006; DeSalazar et al. 2007; Eckermann et al. 2008, 2010; Hawe et al. 2009).

However, historically specifications for including quality of care effects in efficiency measures have not considered the underlying economic objective for health system or provider behaviour in practice. Such specifications suggested for quality can be broadly classified as:

(i) Exogenous specifications (Zuckerman et al. 1994);
(ii) Weakly disposable 'bad output' hyperbolic specifications (Arocena and Garcia-Prado 2007);
(iii) Utility bearing output specifications with the notion of a quality-quantity trade-off (Newhouse 1970); and
(iv) Utility bearing output specification of quality alone (Puig-Junoy 1998; Dawson et al. 2005).

In relation to exogenous specifications of quality, Zuckerman et al. (1994) conditioned cost per admission on standardised mortality rates for hospitals in the lower decile (high quality) and upper decile (low quality), relative to those in 10th to 90th percentile. Such exogenous conditioning improved the efficiency of hospitals with mortality in both upper and lower deciles, as their expected costs were higher for both in comparison to hospitals with standardised mortality rates in the 10th to 90th percentile. The perverse improvement of efficiency measures for those hospitals with low quality (standardised mortality rates in the upper decile) reflects that those hospitals on average had such low quality of care that overall within admission costs increased. To create appropriate incentives for quality of care, hospitals should not be encouraged to reduce their quality of care (increase their standardised mortality rate) in order to have their efficiency performance improved or be paid more. However, exogenous quality of care specifications such as that of Zuckerman et al. (1994) create such incentives, by conditioning on, rather than valuing, quality of care.

More generally exogenous quality of care specifications where efficiency measures condition on quality of care measures:

(i) Do not reflect control of service quality by providers; quality of care by definition is endogenous not exogenous.
(ii) Prevent value of quality being included in economic efficiency measures and hence cannot reflect maximising net benefit.

An endogenous rather than exogenous specification of quality is required for economic efficiency measures to allow value of quality to be included in efficiency measures, let alone reflect an underlying economic objective of maximising net benefit.

Similar issues to exogenous specifications arise in relation to weakly disposable bad output hyperbolic specifications of quality (Färe et al. 1989) where one considers a production technology with one strongly disposable good output – admissions and a weakly disposable 'bad output', say mortality; as considered in detail in Eckermann and Coelli (2013). While there is no cost to disposing of strongly disposable outputs, the weakly disposable 'bad output' implies that one cannot reduce this without cost (Coelli et al. 2005). Considered together a strongly disposable good output (e.g. admissions) and weakly disposable bad output (e.g. mortality) imply a backward bending production function and the need for a hyperbolic direction for efficiency measures. That is, where good outputs and bad outputs are,

respectively, proportionally expanded and contracted at the same rate (Arocena and Garcia-Prado 2007; Tyteca 1996). Apart from being significantly more complex to calculate and conceptualise, such hyperbolic weakly disposable bad output specifications, as highlighted in Eckermann and Coelli (2013), suffer from:

(i) Not preventing projection onto backward bending parts of the efficiency frontier, when estimating efficiency of providers off the frontier;
(ii) Preventing estimation of shadow prices where prices of outputs are unknown (e.g. as in admission in public hospitals); and
(iii) Not allowing a value for quality to be incorporated into economic efficiency measures.

Hence, while significantly more complex than exogenous specifications, hyperbolic weakly disposable bad output specification of quality similarly preclude economic efficiency measures consistent with maximising net benefit.

In relation to less complex while endogenous quality of care specifications, Newhouse (1970) proposed a quantity-quality trade-off (3) where quantity and quality variables were framed as utility bearing outputs, for example outputs of survivors and admissions. However, with such specifications, quality is inextricably related to quantity. Hence, increasing admissions while keeping survivors constant increases mortality and reduces quality. Consequently, specifying quality and quantity variables as endogenous outputs does not support independent Pareto consideration of quality and quantity variables nor allow value or utility to be meaningfully represented in quantity-quality space. Consequently, quality-quantity specifications also cannot support net benefit maximisation.

In attempting to avoid such quality-quantity trade-off problems, research such as that of Puig-Junoy (1998) and Dawson et al. (2005) specify quality alone as an output (4), for example, specifying number of survivors as the only output. This specification reflects an underlying economic objective of minimising average cost per survivor, or more generally minimising average cost per unit of effect framed from a utility bearing perspective. The natural question this raises is whether average cost effectiveness is an appropriate objective for health care?

Many health economists (Arrow 1963; Evans 1981; McGuire et al. 1988; Eckermann 2004; Drummond et al. 2005) have noted that health system costs and effects of care are:

(i) Incremental relative to those of alternative care, hence cost and effect $\neq 0$ with no treatment allowing for downstream health and health system impacts; and
(ii) Patient specific (non-tradable).

Consequently, minimising average cost per unit effect or average cost effectiveness as an underlying objective has been rejected by health economics in HTA in favour of maximising incremental net benefit (INB) in two-strategy comparison (Claxton and Posnett 1996; Stinnett and Mullahy 1998; Willan and Lin 2001; Willan and Briggs 2006). More generally in multiple strategy comparisons (Eckermann 2004; Eckermann et al. 2008; Eckermann and Willan 2011 – see Chap. 8) or multiple

outcome comparisons (McCaffrey 2013, McCaffrey et al. 2015 – see Chap. 10), maximising net benefit has been extended to equivalently minimising net loss.

The challenge then is to enable robust efficiency measurement across providers consistent with maximising net benefit. On face value, this can initially appear insurmountable given robust efficiency measures require radial (ratio) properties (Farrell 1957), where performance improves where all variables increase (output orientation) or decrease (input orientation); while the formulation for INB = $\lambda\Delta E - \Delta C$ doesn't have radial properties. That is, the direction for performance improvement considering INB directly is where incremental effects framed from a utility bearing perspective increase and incremental costs reduce. The lack of radial properties is simply observed on the cost-effectiveness plane for multiple strategy comparison in Fig. 9.2a, where performance improves in a south-east direction as considered in Chap. 8 (Fig. 8.1) comparing six strategies in treatment of gastro-eosophogeal reflux disease (GERD).

Performance improving in moving in a southeast direction in the cost effectiveness plane does not expand from or contract to a vertex and hence prevents radial properties.

However, a linear transformation of net benefit to net loss allows radial (ratio) properties while retaining a correspondence with maximising net benefit, as we saw in Chap. 8 for multiple strategy comparison. This is seen on the cost-disutility plane (Fig. 9.2b) where radial or ratio properties arise with performance improving in

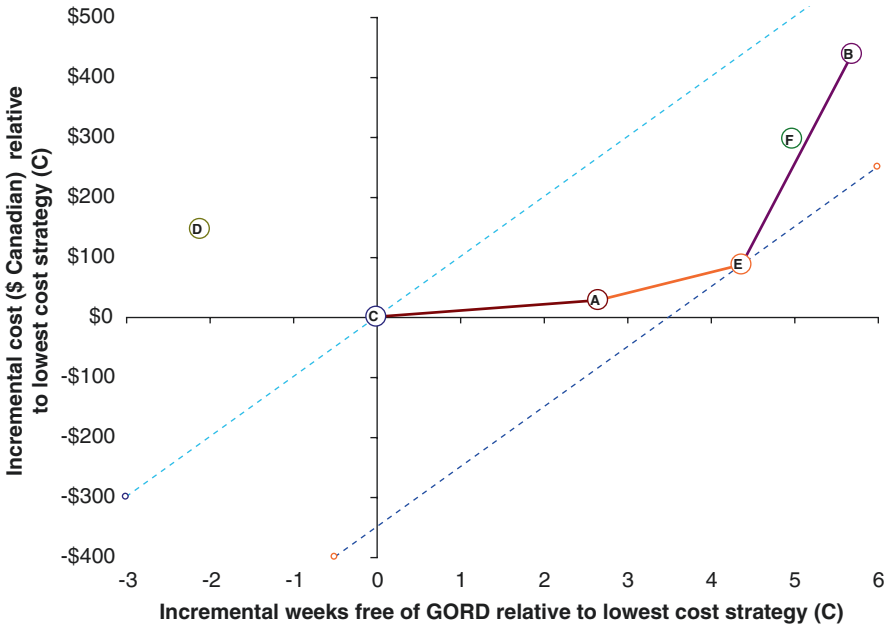


**Fig. 9.2a** Lack of radial properties with direct net benefit comparison on the CE plane (Source: Eckermann et al. (2008))

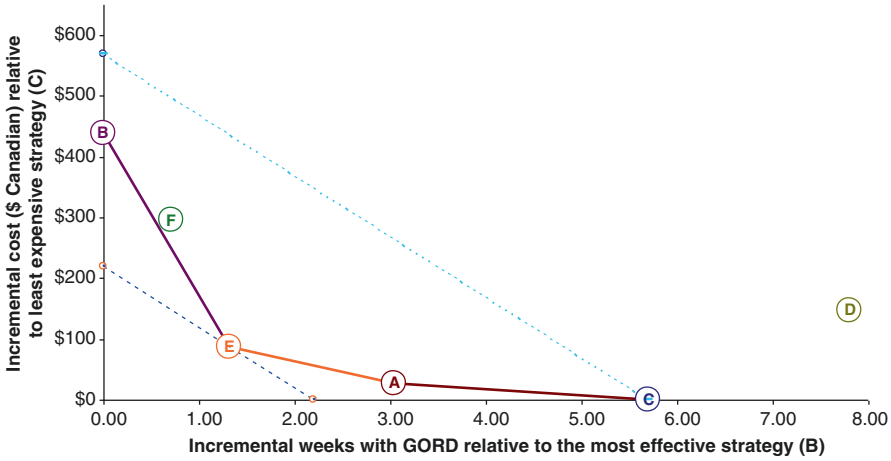


**Fig. 9.2b** Radial properties with comparison on the C-DU plane (Source: Eckermann et al. (2008))

radially (equi-proportionally) contracting costs and disutility framed effects (e.g. mortality) towards the vertex in minimising net loss (equivalent to maximising net benefit).

The ability to reframe net benefit as net loss with a simple transformation allows radial efficiency measures consistent with maximising net benefit on the cost-disutility plane. More formally this is stated as the net benefit correspondence theorem (NBCT), following Eckermann (2004), Eckermann et al. (2008) and Eckermann and Coelli (2013):

There is a one-to-one correspondence betweenmaximising net benefit

$$NB = \lambda E - C$$

and minimising net loss

$$NL = \lambda E - C$$

where

(i) Effects framed from a disutility perspective (DU) cover effects of care in NB (coverage condition); and
(ii) Differences in expected costs and DU are adjusted for (common comparator condition).

In considering efficiency measures of health-care providers in practice, net loss can also be called quality inclusive cost (QIC) as described in Eckermann and Coelli (2013), given that effects are quality of care indicators. In general the one-to-one correspondence underlying the NBCT makes clear that only inclusion of quality effects as inputs framed from a disutility perspective allows efficiency measurement consistent with maximising net benefit.

There are various simple algebraic proofs that derive the one-to-one correspondence between maximising net benefit and minimising net loss or quality inclusive costs under correspondence conditions for one or more effects where they are binary outcomes or otherwise (Eckermann 2004; Eckermann et al. 2008; Eckermann and Coelli 2013). Below we consider the simplest of these proofs for a single binary effect (e.g. mortality) to illustrate how the correspondence arises and why the correspondence conditions are important (and indeed imply a theorem rather than simply a correspondence).

### *9.3.1 NBCT Proof*

Let incremental net benefit per patient for provider *i* relative to a comparator (comp) be given by

$$INB_i = \lambda\left(E_i - E_{\text{comp}}\right) - \left(C_i - C_{\text{comp}}\right)$$

$$= \lambda E_i - C_i - \left(\lambda E_{\text{comp}} - C_{\text{comp}}\right)$$

where $E$ is a binary effect per patient from a utility bearing perspective (e.g. survival rate), $C$ is cost per patient, and $\lambda$ is the net benefit threshold value per unit of effect.

Consider a bilateral comparison between incremental net benefit of providers *i* and *j*.

Without loss of generalisation, let $\text{INB}_i > \text{INB}_j$.

Now, all comparator cost and effect terms cancel under the common comparison condition

$$\Leftrightarrow \lambda E_i - C_i > \lambda E_j - C_j$$

Multiplying both sides by minus 1, the sign changes

$$\Leftrightarrow C_i - \lambda E_i < C_j - \lambda E_j$$

Adding $\lambda$ to both sides of the equation and rearranging, we have

$$\Leftrightarrow C_i + \lambda\left(1 - E_i\right) < C_j + \lambda\left(1 - E_j\right)$$

Under the coverage condition, $(1 - E)$ represents NB effects framed from a disutility perspective, DU (e.g. where E is survival rate and DU is mortality rate)

$$\Leftrightarrow C_i + \lambda DU_i < C_j + \lambda DU_j$$

Hence, there is a one-to-one correspondence between maximising net benefit and minimising quality inclusive cost where coverage and comparability conditions of the NBCT are satisfied. Note that similar proofs arise for the cases of other effect measures with appropriate transformations from utility bearing to disutility bearing effects (e.g. QALYS or life years gained to QALYs or life years lost, etc.), following Eckermann and Coelli (2013).

The one-to-one correspondence between maximising net benefit and minimising quality-inclusive cost underlying the net benefit correspondence theorem implies that the only way of specifying quality of care variables in efficiency measures is from a disutility bearing perspective as inputs. Hence, where maximising net benefit is the appropriate underlying economic objective, efficiency comparison should be undertaken from an input perspective on the cost-disutility plane, or in cost-disutility space with multiple domain of effect comparison as shown in Chap. 10 following McCaffrey et al. (2015).

The correspondence can also naturally be seen graphically on the cost-disutility plane (Fig. 9.3), noting that quality improves in reducing effects framed from a disutility bearing perspective ($E^{DU}$, e.g. readmission, morbidity, waiting time, mortality rate, etc.) and efficiency improves in reducing costs and $E^{DU}$ by radially (equiproportionally) contracting towards the origin. Such a radial direction for performance improvement is critical given Farrell (1957) showed that only radial efficiency measures are scale invariant, the primary property required for efficiency measures. That is, without this radial property, efficiency measures change simply by changing the units or scale of axes. Importantly, the C-DU plane, and more gen-

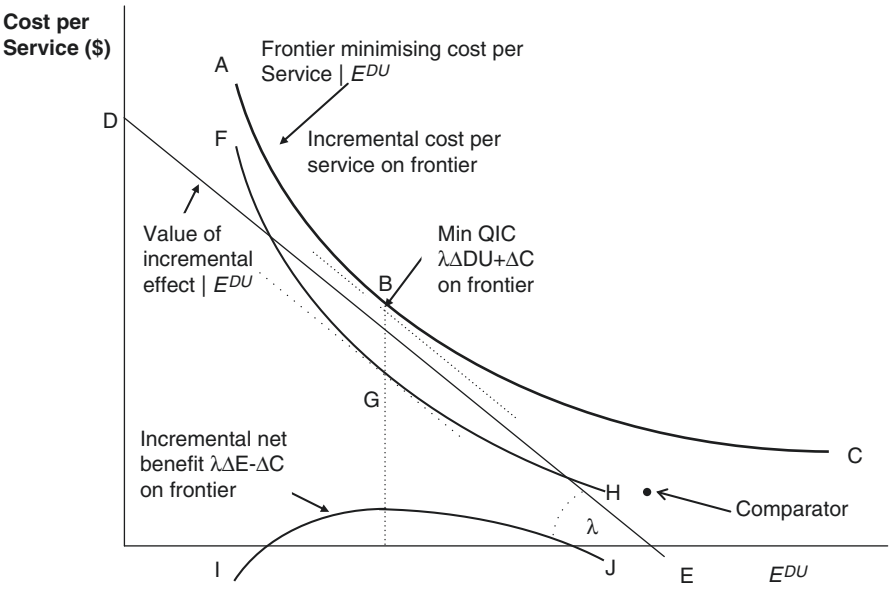


**Fig. 9.3** Graphical representation of the net benefit correspondence theorem (Source: Adapted from Eckermann (2004))

erally C-DU space, has radial properties, and hence efficiency measures do not vary with units of measure or scale of axes.

Figure 9.3 shows graphically why the correspondence between maximising net benefit and minimising quality inclusive cost arises with efficiency frontiers in comparing provider efficiency. For strategies on the frontier ABC relative to a common comparator in Fig. 9.3, the value of their incremental effects (shown by DE) less their incremental costs (FGH) represents their INB (IJ), at their given level of quality, $E^{DU}$. The maximum INB on IJ corresponds to the minimum QIC on the frontier ABC at point B, noting that ABC and FGH are parallel in the vertical plane (constant vertical distance between them is the cost of the comparator). In this respect note that the parallel nature of the frontier (ABC) and incremental cost (FGH) in the horisontal plane (i.e. at any given level of quality of care) requires avoiding the tendency to look at the closest distance between curves. Otherwise, an optical illusion is created by moving from a largely vertical to horizontal plane comparison in going from low to high quality (i.e. from right to left given quality improves with reduction in disutility in Fig. 9.3). Hence, the cost and incremental cost curves actually have the same slope $\lambda$, and quality of care ($E^{DU}$) in the vertical plane of comparison at the point INB is maximised or QIC minimised. In marginal terms at a threshold value for effects of $\lambda$, the value of improved quality of care (reduced $E^{DU}$) along the frontier is greater than the cost at each point in moving from C to B. Beyond that the value of improving quality of care (reducing $E^{DU}$) becomes less than the cost, with each point on AB having greater slope than ED and hence marginal cost being greater than marginal benefit.

While the NBCT has simple algebraic and graphical proofs, the ability that the NBCT provides to undertake radial efficiency measurement consistent with the appropriate net benefit maximising objective function in C-DU space creates a powerful method to allow for including the value of quality in net benefit economic efficiency measures in practice. For example, consider efficiency on the C-DU plane for DRG E62a in comparison of the 45 hospitals costs and mortality rates for DRG E62a (Fig. 9.4). Given performance improves in moving towards the origin (in minimising cost and mortality and hence quality inclusive costs) one can simply and immediately identify hospitals 33, 17 and 26 as peers forming the technical efficiency frontier visually on the C-DU plane, or in applying radial contraction efficiency measurement methods such as data envelopment analysis or index methods (Coelli et al. 2005). That is, these hospitals are technically efficient as they can't radially contract their costs and disutility to that of some convex combination of other hospitals and hence minimise cost for given quality of care ($E^{DU}$) or equivalently maximise quality (minimise $E^{DU}$) for given cost. Hospitals off the frontier are technically inefficient with the ability to radially contract their cost and mortality rate to convex combinations of hospitals 33, 17 and 26 on the frontier.

For any hospital (provider), their technical efficiency is simply calculated using radial methods such as data envelopment analysis or ratio methods which take the distance between the origin and the point on the technical efficiency frontier radially projected onto, as a ratio compared to the distance from the origin to their current cost and disutility rate. For providers on the frontier (33, 17 and 26), their technical efficiency will be 1, while for those off the frontier, this is less than 1, for example, 0.68 for hospital 31 where its radial contraction is depicted in Fig. 9.4.

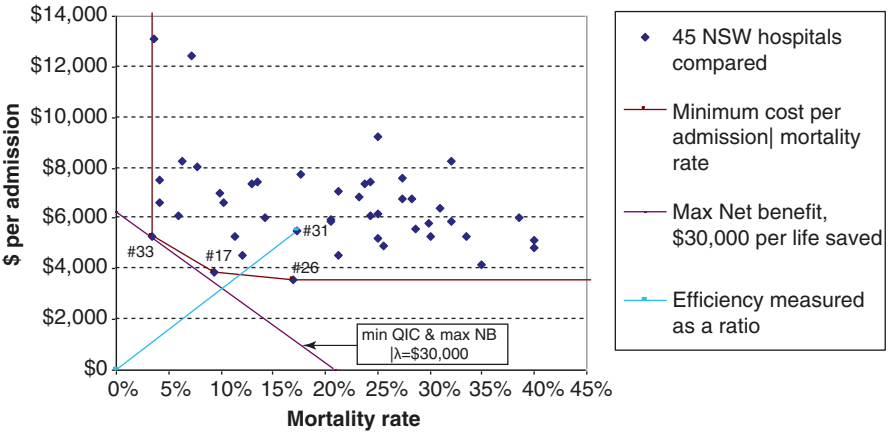


**Fig. 9.4** NBCT efficiency measures given cost and mortality of 45 NSW hospitals for DRG E62a (Source: Eckermann and Coelli (2013))

Technical efficiency under variable returns to scale, where hospitals are effectively compared to a frontier restricted to hospitals of similar size, rather than all hospitals under constant returns to scale can also be simply estimated as shown in Eckermann (2004). This in turn enables scale efficiency (SE) to be estimated as the residual ratio of TE under constant returns to scale and TE under variable returns to scale, i.e. $SE = \frac{TE_{CRS}}{TE_{VRS}}$ (Coelli et al. 2005). Note that technical efficiency under variable returns to scale can only ever be equal to or greater than that under constant returns to scale given the more restrictive nature of peers compared to (those of similar size rather than all hospitals with variable returns to scale).

Considering overall economic (net benefit) efficiency, hospital 33 is shown in Fig. 9.4 at a threshold value of $30,000 per mortality avoided to minimise quality inclusive cost (QIC) or equivalently maximise NB under the NBCT in lying on the net loss line tangent to the convex frontier and hence closest to the origin across hospitals compared at this threshold value. More generally at any threshold value for effects, the quality inclusive cost-minimising (net benefit maximising) hospital can be identified as that on frontier which lies on the QIC line tangent to the frontier, the QIC line closest to the origin at that threshold value.

Net benefit economic efficiency is measured at any threshold value for effects for any hospital as the minimum quality inclusive cost relative to their observed QIC. Given QIC lines have the same slope at a given threshold value by similar triangles this represents the same ratio as the distance from the origin to the point on the lowest QIC line they are projected onto, relative to the radial distance they are from the origin. For the provider that minimises QIC, their economic efficiency is 1 at that threshold value, while other providers have economic efficiency less than 1.

Such net benefit efficiency measures are able to be simply decomposed into:

(i) Technical efficiency, measured relative to radial projection onto the frontier and hence 1 on the frontier while less than 1 off the frontier when using more inputs than required to produce given outputs, (a unit of output in the case of constant returns to scale implicit with per-patient axes); multiplied by
(ii) Allocative efficiency, which is less than 1 where given factor prices providers have suboptimal input factor proportions – value quality too little or too much relative to the relevant threshold value, in the case of net benefit.

Providers off the frontier are economically inefficient and do not minimise QIC (maximise NB) due to technical as well as potentially allocative inefficiency. Those on the technical efficiency frontier when they have economic efficiency less than 1 is due purely to allocative inefficiency.

Such easy decomposition into attributable technical and allocative efficiency components is shown in Fig. 9.5.

To illustrate how economic efficiency simply decomposes, consider the provider at D in Fig. 9.5 who radially projects onto the technical efficiency frontier at $D_1$ and onto the Iso-QIC line of the minimum QIC provider (*C*) at H. Their economic (QIC or net benefit) efficiency is EE = OH/OD, while their technical efficiency under constant returns to scale is TE = $OD_1$/OD and allocative efficiency is AE = $OH/OD_1$. Now, note that $OH/OD = OH/OD_1 \times OD_1/OD$ and more generally EE = AE×TE. Hence, economic (in our case net benefit or QIC) efficiency can be simply decomposed as allocative efficiency multiplied by technical efficiency.

In practice on the C-DU plane, this is achieved with methods such as data envelopment analysis by measuring for each provider economic efficiency in comparing

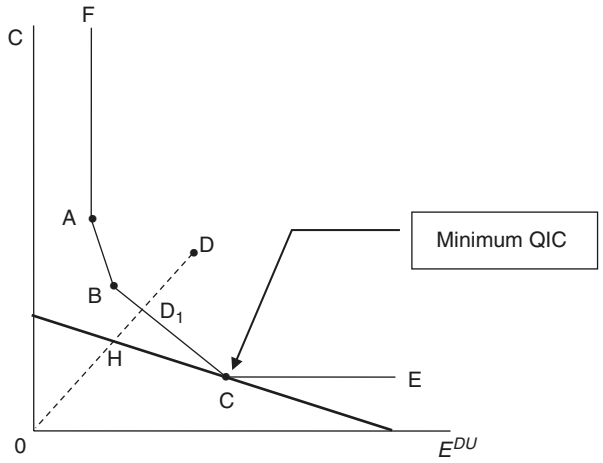


**Fig. 9.5** Net benefit (economic) efficiency decomposition on the C-DU plane (Source: Eckermann and Coelli (2013))

the minimum possible relative to actual QIC ($C + \lambda$DU) and technical efficiency in radially contracting cost and disutility to the convex frontier formed as the lower bound envelope of convex combinations of other strategies (Coelli et al. 2005; Eckermann and Coelli 2013). Allocative efficiency of each provider is then simply calculated as the residual of their economic and technical efficiency, i.e. AE = EE/TE. A decomposition of EE into TE and residual AE is undertaken across our 45 hospitals undertaken in Table 9.2 at a threshold value of $25,000 per mortality

**Table 9.2** Technical, allocative and economic efficiency

| Hospital | TE | AE ($\lambda$ = $25,000) | EE ($\lambda$ = $25,000) |
|---|---|---|---|
| 1 | 0.74 | 0.55 | 0.41 |
| 2 | 0.41 | 0.98 | 0.40 |
| 3 | 0.61 | 1.00 | 0.61 |
| 4 | 0.47 | 0.91 | 0.43 |
| 5 | 0.70 | 0.57 | 0.40 |
| 6 | 0.62 | 1.00 | 0.62 |
| ***7*** | ***0.87*** | ***0.54*** | ***0.47*** |
| 8 | 0.65 | 0.98 | 0.64 |
| 9 | 0.58 | 0.98 | 0.57 |
| 10 | 0.80 | 1.00 | 0.80 |
| ***11*** | ***0.80*** | ***0.89*** | ***0.71*** |
| 12 | 0.44 | 0.86 | 0.38 |
| 13 | 0.59 | 0.66 | 0.39 |
| 14 | 0.93 | 0.47 | 0.44 |
| 15 | 0.67 | 0.99 | 0.66 |
| 16 | 0.59 | 0.83 | 0.49 |
| *17* | *1.00* | *0.99* | *0.99* |
| 18 | 0.51 | 0.88 | 0.45 |
| 19 | 0.85 | 0.95 | 0.81 |
| 20 | 0.60 | 0.83 | 0.50 |
| 21 | 0.57 | 0.98 | 0.56 |
| 22 | 0.74 | 0.73 | 0.54 |
| 23 | 0.63 | 0.89 | 0.56 |
| 24 | 0.68 | 0.71 | 0.48 |
| 25 | 0.79 | 0.78 | 0.62 |
| *26* | *1.00* | *0.78* | *0.78* |
| 27 | 0.80 | 1.00 | 0.80 |
| 28 | 0.51 | 0.98 | 0.50 |
| 29 | 0.76 | 0.99 | 0.75 |
| 30 | 0.61 | 0.72 | 0.44 |
| 31 | 0.68 | 0.91 | 0.62 |
| 32 | 0.54 | 0.83 | 0.45 |

**Table 9.2** (continued)

| | | | |
|---|---|---|---|
| *33* | *1.00* | *1.00* | *1.00* |
| 34 | 0.65 | 1.00 | 0.65 |
| 35 | 0.51 | 0.90 | 0.46 |
| 36 | 0.69 | 0.77 | 0.53 |
| 37 | 0.62 | 0.74 | 0.46 |
| 38 | 0.54 | 0.89 | 0.48 |
| 39 | 0.56 | 0.77 | 0.43 |
| 40 | 0.63 | 0.87 | 0.55 |
| 41 | 0.64 | 0.75 | 0.48 |
| 42 | 0.54 | 0.91 | 0.49 |
| 43 | 0.67 | 0.67 | 0.45 |
| 44 | 0.49 | 0.86 | 0.42 |
| 45 | 0.54 | 0.81 | 0.44 |
| *Industry cost share wtd. Mean* | *0.63* | *0.82* | *0.51* |

avoided, presented with TE and AE columns leading to a EE column as their product to aid simple calculation (EE = TE × AE).

In the same way that net benefit and expected net loss curves and frontiers and CEA curves condition on the threshold value for effects, net benefit efficiency measures can condition on threshold values for quality of care effects, as considered in Table 9.3 for our example with values per mortality avoided from 0 to $100,000. At a threshold value of 0, as with current cost per admission-based efficiency measures (case-mix adjusted or otherwise), net benefit just considers costs and hence reflects cost minimisation. In the case of our 45 hospitals compared hospital 26 minimises cost per admission and hence is identified as the peer with EE of 1 at a 0 threshold value.

As the threshold value for quality of care is increased, Table 9.3 shows that quality of care gradually becomes more important and relative ordering changes. For example, hospital 17 is the peer with EE of 1 at a threshold value of $5000 and $10,000 per death avoided and hospital 33 with the lowest mortality rate at $25,000 per death avoided or higher. Importantly relative economic efficiency of other hospitals are consequently measured relative to the appropriate peers at these threshold values and appropriate incentives are created for net benefit maximisation conditional on the threshold value for quality of care, rather than lowest cost per admission quality of care.

Identifying net benefit maximising, rather than cost minimising, peers and measuring relative performance across hospitals relative to them is key to creating incentives for appropriate quality of care across hospitals.

To determine the exact regions of threshold values for effects where hospitals on the frontier are peers and minimise QIC or equivalently maximise net benefit, the QIC or NB of adjacent providers on the frontier $(i, j)$ simply needs to be equated and solve for $\lambda$. That is, solve for:

**Table 9.3** Economic efficiency conditional on threshold value

| Hospital | $0 | $5000 | $10,000 | $25,000 | $50,000 | $100,000 |
|---|---|---|---|---|---|---|
| 1 | 0.74 | 0.63 | 0.54 | 0.41 | 0.28 | 0.19 |
| 2 | 0.39 | 0.41 | 0.41 | 0.4 | 0.32 | 0.25 |
| 3 | 0.45 | 0.51 | 0.54 | 0.61 | 0.58 | 0.55 |
| 4 | 0.29 | 0.34 | 0.37 | 0.43 | 0.43 | 0.44 |
| 5 | 0.7 | 0.61 | 0.53 | 0.4 | 0.28 | 0.19 |
| 6 | 0.44 | 0.51 | 0.54 | 0.62 | 0.61 | 0.59 |
| 7 | 0.87 | 0.73 | 0.63 | 0.47 | 0.32 | 0.22 |
| 8 | 0.6 | 0.64 | 0.65 | 0.64 | 0.53 | 0.42 |
| 9 | 0.49 | 0.54 | 0.55 | 0.57 | 0.5 | 0.42 |
| 10 | 0.54 | 0.63 | 0.68 | 0.8 | 0.8 | 0.8 |
| 11 | 0.48 | 0.56 | 0.6 | 0.71 | 0.72 | 0.74 |
| 12 | 0.43 | 0.44 | 0.42 | 0.38 | 0.29 | 0.21 |
| 13 | 0.59 | 0.54 | 0.48 | 0.39 | 0.27 | 0.19 |
| 14 | 0.27 | 0.33 | 0.36 | 0.44 | 0.47 | 0.52 |
| 15 | 0.54 | 0.61 | 0.63 | 0.66 | 0.59 | 0.51 |
| 16 | 0.58 | 0.58 | 0.55 | 0.49 | 0.37 | 0.28 |
| 17 | *0.93* | *1* | *1* | *0.99* | *0.81* | *0.65* |
| 18 | 0.48 | 0.5 | 0.49 | 0.45 | 0.36 | 0.27 |
| 19 | 0.79 | 0.84 | 0.84 | 0.81 | 0.66 | 0.52 |
| 20 | 0.59 | 0.59 | 0.56 | 0.5 | 0.38 | 0.28 |
| 21 | 0.48 | 0.53 | 0.54 | 0.56 | 0.49 | 0.41 |
| 22 | 0.74 | 0.7 | 0.64 | 0.54 | 0.39 | 0.28 |
| 23 | 0.61 | 0.63 | 0.6 | 0.56 | 0.43 | 0.33 |
| 24 | 0.68 | 0.64 | 0.58 | 0.48 | 0.34 | 0.24 |
| 25 | 0.79 | 0.77 | 0.72 | 0.62 | 0.46 | 0.33 |
| 26 | 1 | 0.97 | 0.91 | 0.78 | 0.58 | 0.42 |
| 27 | 0.59 | 0.67 | 0.71 | 0.8 | 0.76 | 0.71 |
| 28 | 0.46 | 0.5 | 0.5 | 0.5 | 0 | 0.34 |
| 29 | 0.68 | 0.74 | 0.75 | 0.75 | 0.64 | 0.52 |
| 30 | 0.61 | 0.58 | 0.53 | 0.44 | 0.32 | 0.23 |
| 31 | 0.65 | 0.68 | 0.66 | 0.62 | 0.49 | 0.38 |
| 32 | 0.53 | 0.53 | 0.5 | 0.45 | 0.34 | 0.25 |
| 33 | *0.68* | *0.79* | *0.85* | *1* | *1* | *1* |
| 34 | 0.51 | 0.58 | 0.6 | 0.65 | 0.58 | 0.51 |
| 35 | 0.48 | 0.5 | 0.49 | 0.46 | 0.36 | 0.28 |
| 36 | 0.69 | 0.67 | 0.62 | 0.53 | 0.39 | 0.29 |
| 37 | 0.62 | 0.59 | 0.54 | 0.46 | 0.34 | 0.24 |
| 38 | 0.52 | 0.54 | 0.52 | 0.48 | 0.38 | 0.29 |
| 39 | 0.56 | 0.54 | 0.5 | 0.43 | 0.32 | 0.23 |
| 40 | 0.61 | 0.62 | 0.6 | 0.55 | 0.43 | 0.33 |
| 41 | 0.64 | 0.61 | 0.57 | 0.48 | 0.35 | 0.25 |
| 42 | 0.51 | 0.53 | 0.52 | 0.49 | 0.39 | 0.3 |

**Table 9.3** (continued)

| | | | | | | |
|---|---|---|---|---|---|---|
| 43 | 0.67 | 0.62 | 0.55 | 0.45 | 0.31 | 0.22 |
| 44 | 0.47 | 0.48 | 0.46 | 0.42 | 0.33 | 0.25 |
| 45 | 0.53 | 0.53 | 0.5 | 0.44 | 0.33 | 0.25 |
| Industry | 0.57 | 0.56 | 0.54 | 0.51 | 0.4 | 0.34 |

$$C_i + \lambda DU_i = C_j + \lambda DU_j$$

leading to

$$\lambda = \left(C_j - C_i\right) / \left(DU_i - DU_j\right),$$

which is simply an alternative expression for the ICER threshold between adjacent technically efficient providers on the frontier. Hence, in our example equating QIC or NB of adjacent hospitals 26 and 17 on the frontier, we can solve the threshold value up to which 26 is the peer and beyond which 17 becomes the peer, from data in Table 9.1, as ($3858–$3590)/(0.17–0.09)=($268/0.08)=$3523 allowing for an appropriate number of decimal places in calculation. Hence, the threshold value up to which hospital 26 has EE of 1 and beyond which hospital 17 has EE of 1 is $3523 per mortality avoided. Similarly, the other threshold value from equating QIC or NB for adjacent hospitals 17 and 33 on the frontier can be calculated as $24,356 per mortality avoided. Hence, the QIC minimising or NB maximising peer hospitals are hospital 26 from $0 to $3523 per life saved, hospital 17 from $3523 to $24,356 per life saved and hospital 33 beyond this.

Each of these hospitals have EE of 1 over the respective threshold ranges where they maximise net benefit and are appropriate peers in optimising quality for given cost, at an industry level. Nevertheless a significant policy question of interest is the implicit or shadow value being placed on quality of care across hospital performance generally. On the cost-disutility plane, an industry shadow value for quality is able to be estimated as where the cost share weighted economic efficiency across hospitals is maximised, at $3523 per additional survivor across the 45 hospitals, as shown in Fig. 9.6.

While this industry shadow value for quality of care with current industry behaviour may appear low, it should not be surprising given incentives for a 0 value created by economic efficiency measures and funding mechanisms based on cost alone ignoring quality of care. The implicit industry shadow value reflects the marginal trade-off between cost and mortality over the main regions inefficient hospitals are projected onto with radial contraction to the vertex for 28 or 2/3rds of the 42 hospitals off the frontier who have hospitals 26 and 17 only as their peers.

In comparing hospitals, peer grouping can also be undertaken in attempting to adjust for the severity of within DRG patient populations *a priori* expected by type of hospital. Hence, for example, among the 45 hospitals compared for DRG E62a, the 10 principal referral hospitals could *a priori* be considered to have higher

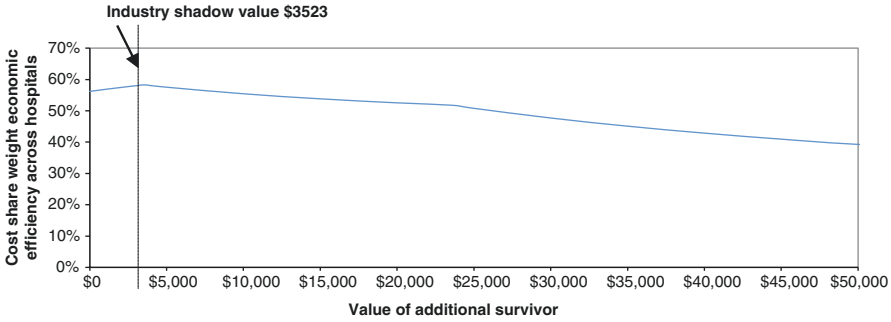


**Fig. 9.6** Industry shadow value where industry economic efficiency maximised (Source: Adapted from Eckermann (2004))

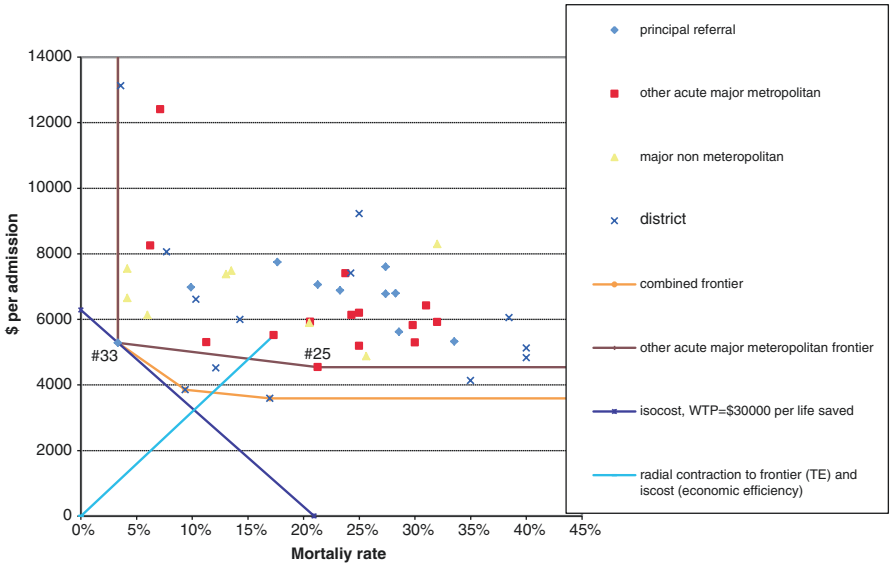


**Fig. 9.7** Ordered peer grouping for other acute major metropolitan hospitals with restricted comparison set and associated frontier (Source: Adapted from Eckermann (2004))

severity of patient within DRG than other hospital types and only be considered with each other – with hospital 33 the only peer (lowest cost and lowest mortality rate among principal referral hospitals). The 14 other major acute hospitals *a priori* could be expected to have the next most severe within DRG complexity and hence could be compared with a frontier formed by both themselves and principal referral hospitals who *a priori* face at least as complex patient populations within DRG.

Figure 9.7 shows how the frontier shifts for other acute major metropolitan hospitals with a more restricted comparison set where principal referral hospital still remains on the frontier, but district hospitals 17 and 26 are no longer in the comparison set, while other acute major hospital 25 now becomes part of the frontier.

Hence, with ordered peer grouping, the other acute major metropolitan hospitals are compared with peers of hospital 25 alone or a combination of hospitals 25 and 33, depending on where they are projected onto this frontier (Fig. 9.7).

Where the eight major non-metropolitan hospitals *a priori* are expected to have the next most severe within-DRG complexity, they would also face this frontier and peers given that no major non-metropolitan hospital lies closer to the origin than this frontier. In the case of the remaining 13 district hospitals as they are *a priori* expected to have least complex patients, they can be compared with all hospitals, and hence their peers are not restricted and face the original general frontier (with hospitals 26, 17 and 33 as peers).

While such ordered peer grouping can aid in addressing comparison conditions of the NBCT based on *a priori* expectations of within-DRG patient complexity, this is crude in comparison to what is required for robust comparability. Robustly satisfying comparability conditions in practice requires adjusting costs and effects for difference in observable patient population risk factors at point of presentation (at inpatient admission in the case of DRGs in hospitals). This is necessary and sufficient to prevent cream skimming incentives to the extent that providers can only cream skim on differences in observable patient population risk factors. Hence, while non-observable risk factors might also be present, providers are not able to cream skim on non-observable factors.

Satisfying coverage conditions in practice requires systematically including effects beyond service with data linkage to, and/or modelling of, post separation effects and treatment costs given health status at service separation, i.e. hospital discharge. Undertaking these processes is necessary and sufficient to prevent incentives for cost and effect shifting, provided coverage of the duration and scope of effects and costs beyond separation is adequate to allow for expected downstream quality of care impacts. In practice this usually needs consideration of 12-month actual and/or modelled impacts, and ideally both to test and improve predictive modelling.

In relation to coverage and comparability conditions, note that our comparison of 45 NSW public hospitals for DRG E62a was based on the 1998–1999 data of within-admission cost and mortality and lacked patient-level risk factor data from hospitals. Hence, under the net benefit correspondence theorem (NBCT), this analysis needs to be appropriately qualified as:

(i) creating incentives for within-DRG cream skimming in not adjusting patient risk factors across hospitals within DRG; and
(ii) creating incentives for cost and outcome shifting in not linking to or modelling effects and costs beyond separation.

The NBCT makes explicit the need for these appropriate qualifications and makes clear the importance of risk adjustment and data linkage or modelling to improve future analysis.

To illustrate what is more generally required to robustly satisfy coverage and comparability conditions of the NBCT, we consider a comparison of three South

**Table 9.4** Logistic regression standardising relative and absolute mortality rates for DRG F10Z

| | Unadjusted | Industry stand. (2) | Adjusted | Adjusted and (2) industry stand. |
|---|---|---|---|---|
| DRG | OR | Risk | OR | Risk |
| *F10Z* | | | | |
| Age | – | – | 1.06 ($p$ = 0.003) | – |
| Charlson comorbidity index | – | – | 1.58 ($p$ = 0.003) | – |
| Hospital A | 0.31 | 0.0153 | 0.26 | 0.0136 |
| Hospital B | 1.00 | 0.0478 | 1.00 | 0.0507 |
| Hospital C | 0.57 | 0.0276 | 0.53 | 0.0267 |
| Total | | 0.0302 | | 0.0302 |

1: No deaths occurred within 12 months of admission at hospital A.
2: Industry-standardized risk rates are back-solved in calibrating actual industry risk to the weighted average of hospital risk from applying the odds ratio for mortality over 12 months in each hospital.

Australian (SA) hospitals in treating cardiac patients with percutaneous transluminal coronary angioplasty (PTCA) for DRGs F10Z and F15Z. Importantly in satisfying coverage and comparability conditions of the NBCT, this analysis undertaken with the SA cardiology network across 1418 index admissions over 2005–2006 included:

(i) Data linkage for mortality and readmission to 12 months from date of index admission; and
(ii) Standardisation of 12-month mortality, readmission rates and associated costs per patient for age and Charlson comorbidity at index admission.

To standardise 12-month mortality rates, logistic regression was undertaken on the odds ratio (OR) for mortality at 12 months allowing for differences in age and Charlson comorbidity index across hospital patient populations at index admission. As established in Chap. 3, standardising risk of binary events using odds ratio (OR) has distinct advantages over relative risk (RR) in ensuring consistent estimation of risk difference with alternative framing and bounding risk between 0 and 1. This the case whenever binary evidence is translated, in indirect comparisons, from trial to jurisdiction or across providers in standardisation (Eckermann et al. 2009, 2011). Table 9.4 reports the raw and adjusted relative OR and absolute standardised rates of 12-month mortality across the three hospitals.

While the SA cardiology network had previously compared surgical hospital mortality rates for PTCA DRGs and found no significant differences between hospitals, they had not previously compared 12-month mortality rates. Hence, when large differences were observed in the unadjusted OR between hospitals for DRG F10Z (OR 0.31 in comparing hospitals A and B), the veracity of these differences was checked internally and externally several times. Once the veracity of raw data had been established beyond any doubt, the network suggested that undertaking adjusted analysis allowing for age and Charlson comorbidity index at index

**Table 9.5** Standardised mortality, readmission and costs per patients across three hospitals for treatment of F10Z and F15Z

| | Std. mortality rate (12 moths) | Std. rate of readmission (12 months) | Std. cost of admissions (12 months) |
|---|---|---|---|
| F10Z and F15Z combined | | | |
| Hospital A | 0.66% | 0.5105 | $10,993 |
| Hospital B | 3.74% | 0.5065 | $10,568 |
| Hospital C | 2.66% | 0.5714 | $11,695 |

admission might substantially diminish observed differences between hospitals. However, risk factor adjustment and standardisation further widened these differences for F10Z in relative terms (OR for mortality 0.26 for hospital A relative to B), leading to the standardised 12-month mortality rate in hospital A (1.36%) being approximately half that of hospital C and one quarter that of hospital B.

When evidence for DRG F10Z (with AMI) was combined with that for F15Z (without AMI with stent implantation), the differences in 12-month standardised mortality rates became even more stark with hospital A (0.66% standardised 12 month mortality rate) having one quarter the rate of C (2.66%) and approaching 1/6th that of B (3.74%), as shown in Table 9.5.

Further, this was not at the expense of additional readmissions or health systems costs. Hospital A had a lower readmission rate at 12 months than hospital C while marginally higher than B, although not if differences in survival rates are adjusted for. Readmission rates per survivor at 12 months for A is 0.5105/0.9934 = 0.514, while for B is 0.5065/0.9636 = 0.525.

In interpreting these findings for three South Australian hospitals, it should be noted that the 12-month mortality rates of hospitals B and C are either equivalent to or lower than those nationally observed in the ACACIA study (Chew et al. 2008), while hospital A has a significantly lower mortality rate. Hence, these results while strongly pointing to hospital A being the appropriate peer for hospitals B and C, also point to hospital A as a peer more generally. Why hospital A performs so well on 12-month outcomes warrants further research. Qualitatively this was suggested to be related to a more systematic handover and follow-up with patients and GPs in hospital A as part of monitoring the appropriateness of post acute care, rehabilitation and medication use, consistent with recommendations arising from the ACACIA study (Brieger et al. 2009).

## 9.4 Policy Implications of the NBCT Framework

In applying the NBCT, a three-step process is suggested to satisfy correspondence conditions (prevent cream skimming and cost-shifting incentives):

(i) Identify patient outcomes and predictive risk factors at admission as with decision analytic methods;

(ii) Measure costs and effects including those beyond discharge (data linkage or expected effects along clinical pathways given discharge state); and
(iii) Adjust quality of care or effect measures rates and costs per patient for patient population differences at admission (and for post separation effects, potentially also differences in environmental factors e.g. socioeconomic factors).

The NBCT method focuses current policy initiatives for data linkage and risk adjustment for patient characteristics at a clinical activity level. To the extent that many of these initiatives are already occurring, the correspondence theorem provides a systematic approach to combine these efforts in creating incentives for NB maximising quality of care in practice, but also prevent their unnecessary replication at different levels and across jurisdictions (e.g. State and National, hospital and MDC as well as DRG clinical activity level) and over time (reinventing the wheel).

Hence, the incremental policy cost is, at worst, marginal and likely cost saving, particularly in the long term given downstream health system cost impacts (considered in detail in Chap. 12).

In terms of impacts on internal hospital processes, where NBCT efficiency measures are robustly applied, with data linkage and standardised of effects and cost with risk factor adjustment across providers and effects incorporated consistent with maximising net benefit:

(i) Providers have their quality of care valued while becoming accountable for quality of care; and
(ii) Administrators can no longer act as accountants minimising cost per admission, need to consider trade-offs between the value and cost of quality.

Hence, the NBCT encourages joint cost and quality of care accountability and meaningful dialogue between administrators and clinicians, trading off cost and value of quality within hospital. This is in stark contrast to the perverse incentives and associated inefficiencies arising with internal warfare between administrators acting as accountants in cost minimising and clinicians hoarding resources for quality maximisation (Harris 1977).

The bottom line is that measuring performance consistent with maximising net benefit under the NBCT creates economic incentives for health system net benefit maximisation in practice, the objective underlying evidence-based medicine but more generally public policy expenditure. Hence, the NBCT applied to efficiency measures supports HTA in choice and use of available technology (allocative and technical efficiency) by clinicians and administrators in practice. Coverage and comparability conditions support risk adjustment and data linkage to prevent cream skimming and cost-shifting incentives.

## 9.5 Further Extensions

The net benefit correspondence theorem while illustrated for appropriately including quality of care in efficiency measurement across hospitals in practice in this chapter can more generally be applied wherever net benefit maximisation is

an appropriate economic objective for efficiency measurement. That is, the NBCT is a generalised theorem (Eckermann 2004; Eckermann Briggs and Willan 2008; Eckermann and Coelli 2013) that uniquely enables efficiency measures consistent with maximising net benefit, wherever that is appropriate. As illustrated for hospital comparisons, the method provides a robust framework to encourage risk factor adjustment and data linkage or modelling to prevent cost and effect shifting and cream skimming incentives and reflect system-wide appropriate incentives.

Hence, NBCT efficiency applications are suggested in comparison of health systems or care service providers' efficiency across other health-care and related settings such as aged care (see chapter 12). However, also more generally to create appropriate incentives for service quality of providers with efficiency measures in industries such as education, transport, corrective services, and also environmental systems. In each case, effects framed from a disutility perspective (e.g. course failure and graduate employment rates in education; stoppages, delays and/or missed connection rates in transport systems; recidivism rates of prisoners in prisons; levels of pollution, land degradation, etc., in environmental systems) simply need to be included as inputs in efficiency measures alongside traditional resource use or costs. Data linkage and/or modelling with risk factor adjustment enables coverage and comparability conditions to provide unqualified analysis in preventing cost and effect shifting and cream skimming incentives.

Chapter 10 highlights that the NBCT method and radial properties of comparison on the cost disutility plane highlighted in this chapter and net loss-based summary measures in Chap. 8 also naturally extend to allowing robust comparison of multiple domains of effect in cost disutility space. Such multiple domain summary measures were pointed to in Chap. 4 and are shown in Chap. 10 to be particularly important for palliative care settings following McCaffrey (2013) and McCaffrey et al. (2015).

The final extension of the NBCT considered as part of policy applications in Chap. 12 extends efficiency measure applications across providers such as that considered for the hospital example in this chapter to funding mechanisms (Eckermann 2004, 2009) in Sect. 12.6. In particular, a two-stage sequential funding mechanism where funding is relative to net benefit maximising peers and value of quality is gradually increased from current shadow value is shown to allow managed transition from cost per admission minimising quality of care with cost-based case-mix funding, to budget constrained health system net benefit maximising quality of care. Importantly this two-stage sequential funding mechanism unlike pay for performance-based measures with 'block payments' at a 'target level of quality' creates increasingly appropriate continuous incentives for quality of care across all hospitals. The robust framework underlying this funding mechanism provided by the NBCT systematically moves funding incentives towards net benefit maximising quality of care or more generally highest quality of care within any budget.

In terms of linking research, reimbursement and regulation in practice, we have also already seen that the NBCT and related comparison on the cost-disutility plane (Eckermann 2004) with consideration of net loss provide a robust framework that naturally leads to:

(i) Expected net loss curves and the expected net loss frontier – linking research and reimbursement in HTA as considered in Chap. 8 following Eckermann et al. (2008) and Eckermann and Willan (2011);
(ii) Support for joint nature of optimal research and reimbursement decisions using VOI methods as considered in Chaps. 5, 6 and 7; and
(iii) Performance (efficiency) measurement and funding consistent with net benefit maximisation in practice as we have considered in this chapter following Eckermann (2004) and Eckermann and Coelli (2013).

Consequently the NBCT in providing a robust framework for comparing multiple strategies and/or providers and/or outcomes, in HTA links to VOI methods and practice with risk adjustment and data linkage or modelling of post service impacts to satisfy comparability and coverage conditions. Therefore, it generalises not just to different settings but also provides a key link to jointly addressing optimal research, reimbursement and regulatory decisions.

## References

Arocena C, Garcia-Prado A. Accounting for quality in the measurement of hospital performance: evidence from Costa-Rica. Health Econ. 2007;16:667–85.

Arrow KJ. Uncertainty and the welfare economics of medical care. Am Econ Rev. 1963;53:941–73.

Australian Government. National Hospital Cost Data Collection 1998–99: final cost report (round 3). Canberra: Department of Health and Aged Care; 2000. Accessed 29 May 2011. http://www.health.gov.au/internet/wcms/publishing.nsf/Content/health-casemix-data-collections-NHCDCReports.

Brieger D, Kelly A, Aroney C, Tideman P, Freedman SB, Chew D, Ilton M, Carroll G, Jacobs I, Huang NP, on behalf of the National Heart Foundation ACS Implementation and Advocacy Working Group. Acute coronary syndromes: consensus recommendations for translating knowledge into action. Med J Aust. 2009;191(6):334–8.

Brook C. Casemix funding for acute hospital care in Victoria, Australia. Victorian state government. 2002. Available at casemix.health.vic.gov.au/about casemix updated 1 October 2002. 16 Sep 2003 date last accessed.

Chew DP, Amerena J, Coverdale S, Rankin J, Astley C, Soman A, Brieger D. Invasive management and late clinical outcomes in contemporary Australian management of acute coronary syndromes: observations from the ACACIA registry. Med J Aust. 2008;188:691–7.

Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. Health Econ. 1996;5(6):513–24.

Coelli TJ, Rao DSP, O'Donnell C, Battese GE. An introduction to efficiency and productivity analysis. 2nd ed. New York: Springer; 2005.

Dawson D et al. Developing new approaches to measuring NHS outputs and productivity – final report. York: Centre for Health Economics, University of York; 2005.

Donaldson C, Gerard K. Economics of health care financing: the visible hand. London: Macmillion; 1993.

Drummond MF, Sculpher M, Torrance G, O'Brien B, Stoddart G. Methods for the economic evaluation of health care programmes. 3rd ed. Oxford, UK: Oxford University Press; 2005.

Eckermann S. Measuring health system efficiency and funding for net benefit maximisation: the health economics of quality of care. SA Health Department working paper. 2009. ISBN 13 978 1 921402 050. Accessed 29 May 2011: http://clinicalchange.flinders.edu.au/files/workingpapers/wp_2009_08.pdf.

Eckermann S. Hospital productivity: a contradiction in terms? In: Harris A, editor. Economics and Health 1994: Proceedings of the Sixteenth Australian Conference of Health Economists, Australian studies in health administration series, vol. 78. Sydney: University of New South Wales; 1994. p. 167–81.

Eckermann S. Hospital performance including quality: creating economic incentives consistent with evidence-based medicine. Ph.D. Dissertation, University of New South Wales; 2004 http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/.

Eckermann S, Briggs A, Willan A. Health technology assessment in the cost- disutility plane. Med Decis Mak. 2008;28:172–81.

Eckermann S, Coelli T. Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice. Soc Sci Med. 2013;76:159–68. http://dx.doi.org/10.1016/j.socscimed.2012.10.020

Eckermann S, Coory M, Willan AR. Indirect comparison: relative risk fallacies and odds solution. J Clin Epidemiol. 2009;62:1031–6.

Eckermann S, Coory M, Willan A. Consistently estimating absolute risk difference when translating evidence to jurisdictions of interest. PharmacoEconomics. 2011;29(2):87–96.

Eckermann S, Karnon J, Willan A. The value of value of information: best informing research design and prioritization using current methods. PharmacoEconomics. 2010;28(9):699–709.

Eckermann S, Willan A. Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies. PharmacoEconomics. 2011;29(7):563–77.

Evans RG. Incomplete vertical integration: the distinctive structure of the health care industry. In: Van Der Gaag J, Perlman M, editors. Health, economics and health economics. Amsterdam: North Holland; 1981. p. 329–54.

Färe R, Grosskopf S, Lovell CAK, Pasurka C. Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. Rev Econ Stat. 1989;71:90–8.

Farrell MJ. The measurement of productive efficiency. J R Stat Soc Ser A (General). 1957;120:253–90.

Graham D. Public expenditure under uncertainty: the net- benefit criteria. Am Econ Rev. 1992;82:822–46.

Harris J. The Internal organization of hospitals: some economic implications. Bell J Econ. 1977;8:467–82.

McCaffrey N. Modelling joint cost and outcomes uncertainty on the cost-disutility plane – case studies in palliative care. PhD, Flinders University; 2013.

McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Better informing decision making with multiple outcomes cost-effectiveness analysis under uncertainty in cost-disutility space. PLoS One. 2015;10(3):e0115544.

McGuire A, Henderson J, Mooney G. The economics of health care: an introductory text. London: Routledge; 1988.

Newhouse J. Towards a theory of non-profit insitutions: an economic model of a hospital. Am Econ Rev. 1970;60(1):64–74.

Newhouse J. Frontier estimation: how useful a tool for health economics? J Health Econ. 1994;13:317–22.

Puig-Junoy J. Technical efficiency in the clinical management of critically ill patients. Health Econ. 1998;7:263–77.

Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Mak. 1998;18(Suppl):S68–80.

Tyteca D. On the measurement of the environmental performance of firms – a literature review and a productive efficiency perspective. J Environ Manag. 1996;46:281–308.

Willan AR, Briggs AH. The statistical analysis of cost-effectiveness data. Chichester: John Wiley & Sons; 2006.

Willan AR, Lin DY. Incremental net benefit in randomized clinical trials. Stat Med. 2001; 20:1563–74.

Zuckerman S, Hadley J, Lezzoni L. Measuring hospital efficiency with frontier cost functions. J Health Econ. 1994;13:255–80.

# Chapter 10
# Multiple Effects Cost-Effectiveness Analysis in Cost-Disutility Space

**Nicola McCaffrey and Simon Eckermann**

## 10.1 Introduction

Chapter 10 shows that the net benefit correspondence theorem methods, introduced and shown to have distinct advantages for multiple strategy comparisons in Chap. 8 and for multiple provider efficiency comparison consistent with maximising net benefit in Chap. 9, naturally extend such advantages to robust multiple domain comparisons under uncertainty. In Chap. 4 we highlighted that robust and generalisable methods to enable jointly considering costs and *multiple* effects under uncertainty are required to better inform funding decisions in complex clinical areas such as palliative care. While quality-adjusted life years (QALYs) enable integration of patient survival with morbidity, they are either unable, or struggle, to incorporate domains such as carer impacts, family distress, finalising personal and financial affairs and being in community of choice for place of palliative care and place of death. Consequently, without robust multiple domain methods of cost-effectiveness analysis, the use of conventional single outcome evaluation (QALY measures or otherwise) can misrepresent key palliative care preferences. Scarce resources and funds can easily end up supporting interventions, strategies or programmes with overall negative impacts and not supporting options that maximise palliative care outcomes from limited resources. In this chapter we show how cost-effectiveness analysis in cost-disutility (C-DU) space enables joint consideration of costs and multiple effects under uncertainty facilitating improved societal decision making. We outline and illustrate how the net benefit correspondence theorem (NBCT) and comparison on the C-DU plane introduced in Chap. 8 also facilitate robust multiple effect comparison under uncertainty with analogous multiple effect summary measures, illustrated with a palliative care modelled analysis. New summary measures identify across any set of threshold values for multiple domains of effect the strategies with lowest expected net loss (ENL) or highest net benefit with ENL planes and the potential value of undertaking further research for the optimal strategy as their lower bound, the ENL contour, as well as the probability of strategies having highest expected net benefit (CEA planes). We illustrate the approach with a palliative care modelled analysis.



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_10

Finally we discuss the relative merits of multiple outcome comparison in C-DU space to inform societal decision making under uncertainty compared with conventional analyses before drawing conclusions and highlighting implications.

## 10.2 Extending Cost-Effectiveness Analysis on the Cost-Disutility Plane

As described in Chaps. 8 and 9, there is a one-to-one correspondence between (i) maximising net benefit (NB) and (ii) minimising quality inclusive costs – costs plus the decision maker's value of effects framed from a utility-reducing (disutility) perspective (Eckermann et al. 2008; Eckermann and Willan 2011; Eckermann and Coelli 2013; Eckermann 2004). Reframing effects from a disutility perspective and comparing strategies on the C-DU plane allow costs and effects framed from a DU perspective (e.g. mortality, waiting time, iatrogenic events) to be equi-proportionally reduced to the origin with radial properties (see Figs. 9.2 and 9.3, Chap. 9). Note that performance (reduction in net loss or equivalently increase in net benefit) intuitively improves in contracting to the origin in C-DU space.

These radial properties enable the use of standard frontier estimation methods, such as data envelopment analysis (DEA), to compare relative performance with economic, technical and allocative efficiency measures (Eckermann 2004; Eckermann et al. 2008; Eckermann and Coelli 2013) as Chap. 9 highlighted in detail with efficiency measures in practice. DEA utilises a mathematical, non-parametric, linear programming technique to construct a technical efficiency frontier from observed data. Generally, employing input-orientated DEA, a piecewise, convex, inner boundary is constructed from observed data given multiple factor inputs and outputs. The frontier reflects combinations of *multiple* inputs that cannot be proportionally contracted without a reduction in a given output with given technology (Coelli 1996). It is the ability of DEA to simultaneously compare *multiple* variables in this manner that allows simple application of the NBCT to compare costs and multiple effects in C-DU space.

For a single effect, Eckermann et al. (Eckermann et al. 2008; Eckermann and Willan 2011) demonstrated in Chap. 8 that comparison on the C-DU plane naturally leads to consideration of expected net loss (ENL) curves and the ENL frontier as summary cost-effectiveness measures. In this chapter such comparison following McCaffrey et al. (2015) is further shown to facilitate explicit and robust consideration of the interaction of uncertainty between costs and multiple effects with:

(i) Threshold regions, mapping combinations of threshold values for joint outcomes where alternative strategies maximise NB;
(ii) Expected net loss planes which quantify differences in ENL across threshold values for multiple effects;
(iii) The expected net loss contour identifying the strategy that maximises ENB (minimises ENL) across bootstrapped replicates for combinations of threshold values across multiple effects; and

(iv) Cost-effectiveness acceptability planes representing the probability that each strategy maximises the ENB at given combinations of value of effects for multiple effects.

### *10.2.1 Technical Efficiency Frontier*

To compare multiple strategies with multiple outcomes in C-DU space, the technical efficiency frontier is simply constructed using DEA with a Farrell input-orientated measure technical efficiency model run under constant returns to scale (Eckermann et al. 2008; Eckermann and Coelli 2013; McCaffrey 2013; McCaffrey et al. 2015). Software to undertake this includes the freely available DEAP package developed by Coelli (http://www.uq.edu.au/economics/cepa/deap.php). Incremental cost relative to the cheapest strategy and incremental effects framed from a disutility perspective relative to the most effective strategy (Eckermann 2004) are included as strongly disposable inputs (Coelli 1996). For multiple effects, as with a single effect, the resulting frontier identifies the strategies that minimise cost for given effects framed from a disutility perspective or equivalently minimise disutility for given cost (Eckermann 2004; Eckermann and Coelli 2013). A technical efficiency score of one indicates the strategy forms part of the technical efficiency frontier and minimises costs for given effect rates at some combination of threshold values. Technical efficiency scores less than one indicate strategies are technically inefficient, i.e. both cost and effects measured from a disutility perspective can be equi-proportionally reduced relative to a given strategy or convex combinations of other strategies. Such strategies are dominated or extended dominated by the other strategies and do not form part of the frontier.

Theoretically, the technical efficiency frontier and the strategy maximising NB can be simply identified for any combination of threshold values for multiple effects in C-DU space using this approach, regardless of the number of dimensions. The optimal strategy in C-DU space at any given combination of threshold values of effects is identified at the point of tangency between the NB line closest to the origin in C-DU space and the technical efficiency frontier (Eckermann 2004; Eckermann et al. 2008; Eckermann and Coelli 2013).

### *10.2.2 Deterministic Analyses*

#### 10.2.2.1 Threshold Regions Across Effect Values where Strategies are Optimal

To determine threshold regions for values of effect domains where alternative strategies are preferred, net loss (NL) is first calculated from incremental analysis in C-DU space. Applying the NBCT (Eckermann 2004; Eckermann et al. 2008; Eckermann and Coelli 2013), the objective of maximising NB is equivalent to

minimising NL. The NL of any given strategy (i) is the loss in NB from choosing *i* rather than the optimal strategy (*) for threshold values represented for one effect by *k* (the monetary value assigned to one unit of effect) and can be found as follows:

$$NL_{*i} = \left(k\,DU_i + C_i\right) - \left(k\,DU_* + C_*\right) \tag{10.1}$$

where DU is the outcome framed from a disutility perspective and *C* represents costs.

This relationship can be extended to include multiple effects (McCaffrey et al. 2015). For example, for three effects with three associated threshold values ($k_1$, $k_2$, $k_3$) and three strategies, *i*, *j* and *m*, the NL is calculated as follows:

$$\begin{aligned} NL_{*i} = &\left(\left(k_1 \mathrm{DU}_{i1}\right) + \left(k_2 \mathrm{DU}_{i2}\right) + \left(k_3 \mathrm{DU}_{i3}\right) + C_i\right) \\ &- \left(\left(k_1 \mathrm{DU}_{*1}\right) + \left(k_2 \mathrm{DU}_{*2}\right) + \left(k_3 \mathrm{DU}_{*3}\right) + C_*\right) \end{aligned} \tag{10.2}$$

$$\begin{aligned} NL_{*j} = &\left(\left(k_1 \mathrm{DU}_{j1}\right) + \left(k_2 \mathrm{DU}_{j2}\right) + \left(k_3 \mathrm{DU}_{j3}\right) + C_i\right) \\ &- \left(\left(k_1 \mathrm{DU}_{*1}\right) + \left(k_2 \mathrm{DU}_{*2}\right) + \left(k_3 \mathrm{DU}_{*3}\right) + C_*\right) \end{aligned} \tag{10.3}$$

$$\begin{aligned} NL_{*j} = &\left(\left(k_1 \mathrm{DU}_{m1}\right) + \left(k_2 \mathrm{DU}_{m2}\right) + \left(k_3 \mathrm{DU}_{m3}\right) + C_i\right) \\ &- \left(\left(k_1 \mathrm{DU}_{*1}\right) + \left(k_2 \mathrm{DU}_{*2}\right) + \left(k_3 \mathrm{DU}_{*3}\right) + C_*\right) \end{aligned} \tag{10.4}$$

With deterministic analysis, the preferred option is the strategy which minimises mean NL (equivalently maximises NB) at any given combination of threshold values, i.e. $C + k_1\mathrm{DU}_1 + k_2\mathrm{DU}_2 + k_3\mathrm{DU}_3$ is minimised. For example, strategy *i* is preferred to strategy *j* when the mean NL of *i* is lower than the mean NL of *j*. To find the regions where alternative strategies are preferred, the boundary of the regions is first determined by equating the NL expressions for adjacent compared strategies on the frontier in C-DU space and solving for $k_1$ and $k_2$, e.g.

$$C_i + k_1\,\mathrm{DU}_{i1} + k_2\,\mathrm{DU}_{i2} = C_j + k_1\,\mathrm{DU}_{j1} + k_2\,\mathrm{DU}_{j2} \tag{10.5}$$

$$C_j + k_1\,\mathrm{DU}_{j1} + k_2\,\mathrm{DU}_{j2} = C_i + k_1\,\mathrm{DU}_{m1} + k_2\,\mathrm{DU}_{m2} \tag{10.6}$$

Values either side of the boundary readily identify the combinations of potential threshold values where each strategy is preferred (minimise mean NL).

### *10.2.3 Summary Measures Under Uncertainty: The Value of Accounting for Joint Uncertainty*

It is important that joint cost and multiple effect uncertainty is quantified to provide unbiased and rigorous assessment of the potential value of undertaking further research to inform joint research and reimbursement decisions (Eckermann and

Willan 2007; Koerkamp et al. 2007; Claxton 2008). Probabilistic sensitivity analysis (PSA) allows modelling of uncertainty across the joint distribution of incremental costs and multiple incremental effects. The following measures summarise the likely return on investment (differences in expected net loss, ENL), potential value of research (EVPI) and risk in return (indicated by the probability of minimising ENL or equivalently maximising ENB) at different relative and absolute threshold values for multiple outcomes given current uncertainty.

#### 10.2.3.1 Expected Net Loss (ENL) and ENL Planes

Expected net loss (ENL) for any strategy at any given combination of threshold values is simply calculated allowing for stochastic uncertainty across replicates (Eckermann et al. 2008). Choosing a strategy that does not maximise NB incurs a NL. The NL of that strategy relative to the NB maximising strategy is calculated in each replicate and the average taken across replicates to estimate the ENL for that strategy conditional on the threshold values for effects. Expected net loss arises in the proportion of replicates where the strategy does not maximise NB at specified threshold values, reflecting decision uncertainty given current evidence. With asymptotically risk-neutral preferences for societal decision making (Arrow and Lind 1970), the preferred strategy under uncertainty is the strategy which maximises ENB or equivalently minimises ENL.

ENL planes, as with ENL curves for a single effect (Eckermann et al. 2008), quantify differences in ENL (loss in ENB) across strategies for different combinations of values for effects. ENL planes are formed by varying the threshold values for the multiple effects and recalculating the average ENL across replicates for each strategy (McCaffrey et al. 2015). The distance between planes at any set of threshold values for effects represents the difference under uncertainty in ENB or ENL between strategies (Eckermann et al. 2008). This relationship arises given the common comparator across strategies in each replicate, i.e. comparing with the strategy minimising ENL across strategies at any given threshold value in each replicate (see Chap. 8).

### *10.2.4 Expected Net Loss Contour*

The contour is formed by the lower bound of the ENL planes across strategies, analogous to the ENL frontier as the lower bound of ENL curves (Eckermann et al. 2008; Eckermann et al. 2010). The ENL contour readily identifies the strategy that maximises ENB (minimises ENL) for any set of values for $k$'s (e.g. $k_1$ and $k_2$ for two effects). As with the ENL frontier where only one outcome is considered, the ENL contour also represents the expected value of perfect information (EVPI) per patient associated with choosing the strategy minimising ENL (Eckermann et al. 2008; Eckermann et al. 2010), but in the case of multiple effects as a function of potential threshold values across these multiple effects.

#### 10.2.4.1 Cost-Effectiveness Acceptability Planes

Cost-effectiveness acceptability planes (CEAPs) show the probability that each strategy maximises ENB (or equivalently minimises ENL) conditional on threshold values for multiple effects (McCaffrey et al. 2015). For each strategy, the CEAP is formed by determining the proportion of replicates where the strategy minimises net loss for different combinations of threshold values for effects.

We now illustrate multiple outcome comparison under uncertainty for a palliative care setting. Consideration of multiple domains in this context is particularly valuable given multiple domains arise in palliative care that cannot be integrated with survival. Quality-adjusted life years cannot integrate palliative patient-, family- and carer-valued domains with patient survival, in particular aspects of the process of death such as the ability to finalise personal and financial affairs (McCaffrey et al. 2014), patient location during palliative care, family and carer distress and other utility bearing impacts such as autonomy in decision making, grief and carer burden. Multiple outcome cost-effectiveness analysis in C-DU space enables joint consideration of such multiple effects under uncertainty critical for improved societal decision making in palliative care settings. The methods illustrated in the following sections use an example with modelled analysis of palliative care for patients with advanced cancer and anorexia and consider oedema and appetite domains that could potentially be integrated in QALY analysis. The same multiple domain methodological issues that arise in this example would be expected to arise if key palliative domains such as carer impacts, family distress, finalising affairs, etc. were available for analysis.

## 10.3 Multiple Domain Palliative Care Example

Anorexia-cachexia is a common syndrome at end of life, impairing quality of life (QOL) and contributing to morbidity and mortality (Berenstein and Ortiz 2005; Argiles et al. 2010; Tisdale 2010). Typical symptoms are loss of appetite, involuntary weight loss, tissue wasting and weakness (Inui 2002; Goebel 2010). Pharmacological treatments include corticosteroids, cytokine inhibitors and appetite stimulants (Inui 2002; Good et al. 2006; Dy and Apostol 2010; Tisdale 2010). However, there is little cost-effective evidence to guide treatment choices in advanced cancer populations.

### *10.3.1 Methods*

#### 10.3.1.1 Model Structure

A simple probabilistic decision tree model was developed to assess the cost-effectiveness of megestrol and dexamethasone for the palliative treatment of anorexia-cachexia related to cancer and to reflect the combined uncertainty in the model inputs when multiple effects are evaluated (Fig. 10.1).

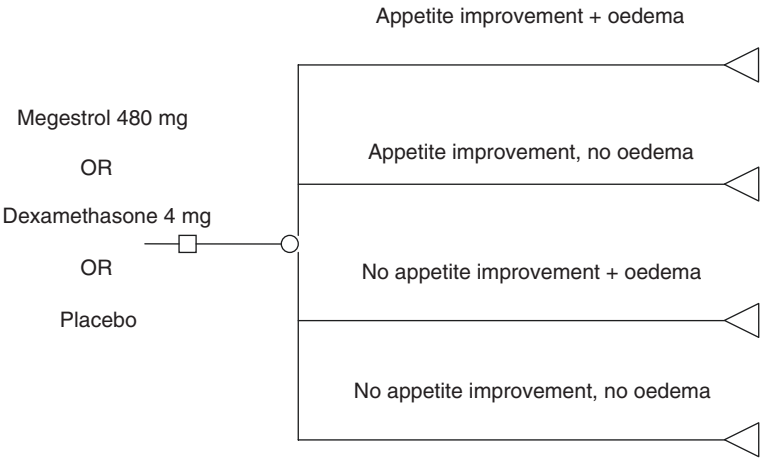


**Fig. 10.1** Model structure

Patients are treated with either megestrol 480 mg, dexamethasone 4 mg or placebo for 46 days (based on the median duration of included trials). The efficacy, effectiveness and safety of megestrol 480 mg and dexamethasone 4 mg are modelled to mirror results reported for cancer participants in a systematic review published by Berenstein and Ortiz (Berenstein and Ortiz 2005). Anorexia patients' appetite may or may not improve, and they may or may not develop oedema (toxicity). Oedema rates are the only adverse event rates included in the model as oedema was the only adverse event occurring with a statistically significant difference between megestrol and placebo in the Cochrane review (Berenstein and Ortiz 2005). It is assumed that if oedema occurs, then patients visit their general practitioner (GP) and have treatment with spironolactone 100 mg daily for 30 days (Black 2001). A health-care system perspective was taken for the analysis, reflecting Australian Pharmaceutical Benefits Advisory Committee (PBAC) guidelines (Pharmaceutical Benefits Advisory Committee 2013).

### 10.3.1.2 Parameters

**Effectiveness**

A search of the literature was conducted using the MEDLINE database from 1966 to 31st December 2008 to identify randomised controlled trials (RCTs) directly comparing megestrol, dexamethasone and placebo for the palliative treatment of anorexia associated with cancer. This search used the following inclusion criteria: English language, RCT, palliative care population, megestrol versus dexamethasone and placebo (three arms) and two or more reported outcomes. Due to the absence of RCTs directly comparing these interventions, the inclusion criteria were broadened to include meta-analyses and RCTs comparing, in a cancer population, (i) megestrol versus dexamethasone, (ii) megestrol versus placebo and (iii) dexamethasone versus placebo.

The literature search revealed Berenstein et al. (Berenstein and Ortiz 2005) had published the most recent and rigorous systematic review evaluating the efficacy, effectiveness and safety of megestrol in palliating anorexia-cachexia syndrome in patients with cancer, AIDS and other underlying pathologies. Briefly, Berenstein and Ortiz included RCTs indexed between database inception and 30th June 2006, involving participants with a clinical diagnosis of anorexia-cachexia related to cancer, AIDS or other pathologies, with less than 50% of participants lost to follow-up, and comparing megestrol (at any dose) versus placebo or other active treatments. Thirty-five trials were included in their Cochrane review: the megestrol dose ranged from 100 to 1600 mg per day and study duration from 10 to 126 days (median 56 days). Twenty-six of the 35 trials included participants with a diagnosis of cancer, with a combined total of 4148 participants diagnosed with cancer. The authors extracted outcome measurement data for appetite improvement (yes/no), weight gain (yes/no) and health-related quality of life (HRQOL) (improved/not improved) and analysed adverse events as the number of participants who suffer an event described as a side effect in the included studies.

The relative efficacy and safety of megestrol versus placebo and megestrol versus dexamethasone were estimated by directly pooling appetite improvement and oedema rates for participants with cancer from Berenstein and Ortiz (McCaffrey et al. 2011). Dichotomous variables were calculated using odds ratios (OR) with 95% confidence intervals (CI) for translating treatment effect to the local population (Eckermann et al. 2009). Odds ratios (OR) were applied to comparator arm odds to estimate treatment odds and converted back to probabilities to inform the model.

An adjusted indirect comparison (Bucher et al. 1997) was conducted to estimate the relative efficacy and safety of dexamethasone versus placebo in the absence of RCTs directly comparing dexamethasone and placebo for palliative treatment of anorexia unrelated to chemotherapy in patients with cancer.

The incremental proportion of patients with (without) appetite improvement and the incremental proportion of patients without (with) oedema were reported as effects of the model, alternatively framed from a utility/disutility bearing perspective. Health-related quality of life was not included as an effect measure for the modelled evaluation due to heterogeneity of scales and methods used in the trials to assess HRQOL. Appetite improvement was chosen over weight gain as an effectiveness measure for the anorexia domain as the former measure is suggested as more clinically relevant for a palliative care population diagnosed with a progressive life-limiting condition (López et al. 2005).

**Costs**

Cost estimates included intervention treatment costs, costs of GP visits and medication associated with treating oedema. Megestrol treatment costs were for a 480 mg dose per day, the mode dose reported in the trials used to inform the model. Consequently, the model efficacy parameters are also assumed to estimate those for megestrol 480 mg in the base-case analysis. Nevertheless, appetite improvement in an advanced cancer population has been reported for megestrol doses ranging from

160 to 800 mg/day (Berenstein and Ortiz 2005). Sensitivity analyses allowed the cost of megestrol to vary across the range from 160 to 800 mg, while invariant associated effects were modelled, given current evidence of no or minimal dose-effect relationship across this range (Inui 2002).

A dexamethasone 3 mg daily dose was administered in the trial used to inform the model (Loprinzi and Kugler 1999). However, dexamethasone treatment costs were based on a 4 mg daily dose in the model to reflect available Australian dosage forms and costing structure of dexamethasone, patient convenience and recommended daily dose (Black 2001; Fearon et al. 2011; Therapeutic Guidelines 2015). The net clinical benefit of dexamethasone 4 mg is not expected to be worse than 3 mg given the dose-response relationship for dexamethasone (Italian Group for Antiemetic Research 1998).

A proportion of patients were modelled to drop out or withdraw half way through the treatment period. Insufficient details were provided in Berenstein and Ortiz for withdrawal or dropout rates for participants with a cancer diagnosis. Therefore, data were extracted directly from trials included in the Cochrane review and pooled. Nine trials (56.25%) were excluded as attrition rates were not reported. Treatment costs were adjusted accordingly.

Only direct costs were included. All costs are reported in 2009 Australian dollars and were estimated from local prices (Australian Government Department of Health and Ageing 2009a, b). A discount rate was not applied as the model time horizon was less than a year.

#### 10.3.1.3 Analysis

The primary objective of the economic evaluation was to compare the incremental resource use, cost and consequences of megestrol, dexamethasone and placebo for the palliative treatment of anorexia-cachexia related to cancer. The evaluation provided estimates from the model of the incremental cost per additional patient with appetite improvement (effect 1), incremental cost per additional patient without oedema (effect 2), cost-effectiveness acceptability curves (CEACs) (van Hout et al. 1994; Briggs and Gray 1999; Löthgren and Zethraeus 2000), threshold regions (McCaffrey et al. 2015), expected net loss contour and cost-effectiveness acceptability planes (CEAPs) (McCaffrey et al. 2015). The probabilistic decision tree model was constructed in Microsoft® Office Excel 2003; statistical analyses were conducted in RevMan® 5.0.17 (Cochrane Collaboration 2008), Comprehensive Meta-analysis Version 2 (Biostat 2005) and Microsoft® Office Excel 2003; and the technical efficiency frontier was constructed in OnFront® (Version 2) using Farrell input-orientated DEA under constant returns to scale (reflecting per-patient estimates) (Coelli 1996).

Uncertainty was modelled probabilistically with Monte Carlo simulation conducted to form multivariate cost and effect distributions with 10,000 replicates. The 2.5 and 97.5 percentiles of the ordered replicates were taken to calculate 95% confidence intervals (Briggs et al. 1997). Probability distributions for each type of

parameter were modelled from relevant evidence and their natural functional form (McCaffrey et al. 2011), while sensitivity analysis was undertaken on key parameters.

## *10.3.2 Results*

### 10.3.2.1 Conventional Analyses

A summary of the modelled disaggregated and incremental costs and consequences framed from a utility perspective is presented in Table 10.1. The results of the PSA are presented as 95% CI. The estimates suggested dexamethasone has the greatest appetite improvement but also has the highest oedema rates versus megestrol and placebo. The mean per-patient costs associated with megestrol, dexamethasone and placebo therapy over follow-up of 46 days were $301, $34 and $9, respectively. This primarily reflects higher drug treatment costs per day for megestrol ($7.25 vs. $0.40 vs. $0, respectively).

**Table 10.1** Summary of modelled disaggregated incremental costs and consequences framed from a utility perspective at 46 days for deterministic analysis

| Comparison | Megestrol vs. placebo | | Dexamethasone vs. placebo | | Megestrol vs. dexamethasone | |
|---|---|---|---|---|---|---|
| Expected outcome | Placebo | Megestrol | Difference | Dexametha-sone | Difference | Difference |
| Drug treatment costs | $0 | $283.47 | $283.47 | $14.55 | $14.55 | $268.92 |
| Oedema costs | $9.12 | $17.12 | $8.00 | $19.78 | $10.65 | −$3.00 |
| Total costs | $9.12 | $300.59 | $291.46 | $34.32 | $25.20 | $266.26 |
| % with appetite improvement | 19.74 | 65.13 | 45.39 | 71.95 | 52.21 | –6.82 |
| % without oedema | 85.44 | 72.67 | −12.76 | 68.43 | −17.00 | 4.24 |
| ICER (95% CI) | Megestrol-placebo | | Dexamethasone-placebo | | Megestrol-dexamethasone | |
| Incremental cost per additional person with appetite improvement | $642 ($445–1583) | | $48 ($25–126) | | Dominated (dominated-$21,242) | |
| Incremental cost per additional person without oedema | Dominated (dominated) | | Dominated (dominated) | | $6281 (dominated-$39,935) | |

Considering appetite alone, the estimated incremental cost per additional patient with appetite improvement for megestrol versus placebo was $642 and $48 for dexamethasone versus placebo. Megestrol was dominated by dexamethasone and therefore did not sit on the technical efficiency frontier. However, when the incremental cost per additional patient without oedema was calculated, placebo dominated (had lower cost and oedema rate than) both megestrol and dexamethasone.

For appetite improvement alone, treatment with megestrol or dexamethasone was associated with increased benefits and increased cost versus placebo for all replicates in the bivariate distribution of incremental costs and effects (McCaffrey et al. 2011). If oedema is the sole effect considered, then placebo dominated megestrol (all but two replicates in the NW quadrant of the cost-effectiveness plane) and dexamethasone (98% of replicates in the NW quadrant). Hence it should be clear that the bivariate distribution of incremental costs and effects ranged considerably across effects and their threshold values when megestrol was compared with dexamethasone and placebo for each effect separately.

Figure 10.2 presents strategies CEACs for appetite improvement for the base-case analysis. The CEACs indicated that placebo has the highest probability of maximising NB (min. NL) up to a threshold value of $48 per patient with appetite improvement. Dexamethasone has highest probability beyond that threshold value,

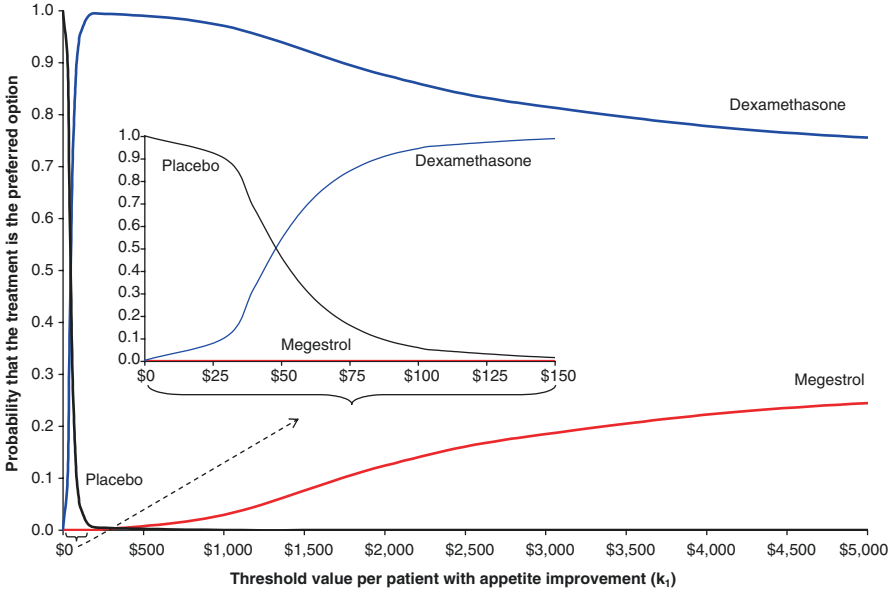


**Fig. 10.2** Cost-effectiveness acceptability curves for megestrol and dexamethasone versus placebo when appetite improvement is the only outcome considered

rising to 95% by $105 and 99% by $150 per patient with appetite improvement. If the decision maker threshold value were substantially higher, say $5000, then the likelihood that dexamethasone maximises NB (min. NL) is 76% while megestrol is 24%.

However, CEACs show very different preferences between strategies when oedema is the sole outcome considered. Placebo has more than a 99% chance of maximising NB (min. NL) when the decision maker has a threshold value per patient without oedema between $0 and $50,000 (see McCaffrey et al. 2011).

Hence with conventional analysis on the cost-effectiveness plane, treatment with megestrol or dexamethasone is associated with increased benefits and increased cost versus placebo for appetite improvement, while placebo dominates megestrol and dexamethasone for oedema. Megestrol does not feature on the technical efficiency frontier as an optimal strategy for some set of effect values in either of these partial analyses, where each effect is considered independently.

#### 10.3.2.2 Comparison in Cost-Disutility Space

A summary of the modelled incremental costs and consequences framed from a disutility perspective is presented in Table 10.1. Mean incremental costs of strategies relative to the cheapest treatment and mean incremental effects framed from a disutility perspective relative to the most effective treatment are calculated for each strategy. For example, when considering the additional number of patients with oedema over 46 days, the mean incremental effect for dexamethasone is zero because dexamethasone is the most effective treatment (14.6–14.6%). Similarly, as placebo is the cheapest strategy, the mean incremental cost for placebo is zero ($9.12 – $9.12).

Technical Efficiency Frontier

In contrast to separate and partial analysis of appetite improvement and oedema on the cost-effectiveness plane, all three therapies sit on the technical efficiency frontier in C-DU space where costs and multiple outcomes are jointly considered (no appetite improvement and oedema; Fig. 10.3). That is, all three strategies minimise mean NL for some set of threshold values for appetite improvement and reduced oedema rates. The frontier in C-DU space represents the inner bound of linear (convex) combinations of strategies closest to the origin, i.e. those minimising mean NL at different combinations of threshold values for effects. For example, placebo forms part of the technical efficiency frontier with an additional 52.2% of patients without appetite improvement relative to the most effective strategy (dexamethasone), 0% of patients with oedema relative to the most effective strategy (placebo) and $0 incremental cost relative to the cheapest strategy (placebo).

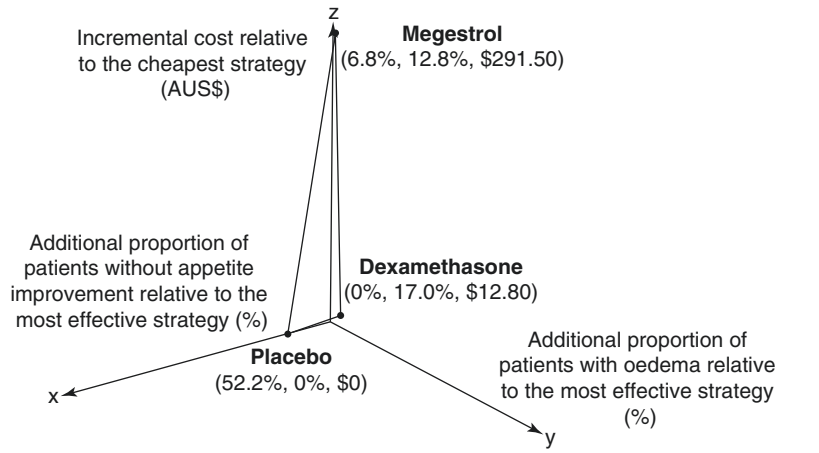


**Fig. 10.3** Technical efficiency frontier in three-dimensional cost-disutility space

**Table 10.2** Summary of modelled incremental analysis framed from a disutility perspective at 46 days for megestrol, dexamethasone and placebo

| | MG | DX | Placebo |
|---|---|---|---|
| Incremental cost per patient over 46 days ($)[a] | $291.46 | $25.20 | $0 |
| Additional proportion of patients without appetite improvement over 46 days[b] | 6.8% | 0% | 52.2% |
| Additional number of patients with oedema over 46 days[b] | 12.8% | 17.0% | 0% |

*DX d*examethasone, *MG m*egestrol
[a]Relative to the cheapest strategy (Argiles et al. 2010)
[b]Relative to the most effective strategy

### 10.3.2.3 Deterministic Analyses

Threshold Regions

Following the incremental analysis in C-DU space (Table 10.2), the mean NL for each strategy can be expressed as follows:

(i) Placebo: $\$0 + 0.52k_1 + 0k_2 = 0.52k_1$
(ii) Dexamethasone: $\$25.20 + 0k_1 + 0.17k_2 = 0.17k_2 + \$25.20$
(iii) Megestrol: $0.07k_1 + 0.13k_2 + \$291.46$

where $k_1$ = threshold value per patient with appetite improvement and $k_2$ = threshold value per additional person who does not experience oedema.

The regions where any given strategy or intervention is preferred can be found by considering where their mean NL is minimised relative to other strategies. The boundaries between such regions are where mean NL are equal between potentially optimal strategies on the deterministic frontier as per multiple strategy, one effect deterministic analysis on the C-DU plane (Eckermann et al. 2008). In a three-strategy comparison with multiple effects, such as our example, this requires comparison of NL for adjacent strategies on the technical efficiency frontier (Fig. 10.3).

Megestrol is preferred to placebo when the mean NL for megestrol is less than that for placebo and hence,

$0.07k_1 + 0.13k_2 + \$291.46 < 0.52k_1$ or rearranging, equivalently.

$$k_2 < 0.45k_1 / 0.13 - 291.46 / 0.13$$

$$k_2 < 3.46k_1 - 2242.0$$

The boundary where mean NL is equal between megesterol and placebo strategies takes the form of the line $k_2 = 3.46k_1 - 2242.0$ as shown in Fig. 10.4.

Megestrol is preferred to dexamethasone when the mean NL for megestrol is less than that for dexamethasone and hence,

$0.07k_1 + 0.13k_2 + \$291.46 < 0.17k_2 + \$25.20$ or rearranging, equivalently.

$$k_2 > 0.07k_1 / 0.04 + 266.26 / 0.04$$

$$k_2 < 1.75k_1 + 6656.5$$

The boundary where mean NL is equal between megestrol and dexamethasone strategies takes the form of the line $k_2 = 1.75k_1 + 6656.5$ as shown in Fig. 10.4.

Dexamethasone is preferred to placebo when the mean NL of dexamethasone is less than that for placebo and hence,

$0.17k_2 + \$25.2 < 0.52k_1$ or rearranging, equivalently.

$$k_2 < 0.52k_1 / 0.17 - 25.2 / 0.17$$

$$k_2 < 3.06k_1 - 148.2$$

The boundary where mean NL is equal between dexamethasone and placebo strategies takes the form of the line $k_2 = 3.06k_1 - 148.2$ as shown in Fig. 10.4.

The regions where each strategy is preferred and the boundary conditions between strategies are presented in Fig. 10.4. The mutually exclusive and exhaustive regions where each strategy is preferred (minimises mean NL) are easily identified directly from the diagram. For example, when $k_1$ = \$3000 and $k_2$ = \$10,000, placebo is the preferred option. The regions are separated by threshold lines where two strategies maximise mean NB and an indifference point where mean NB is equalised across all the strategies (the intersection of the threshold lines; point A in

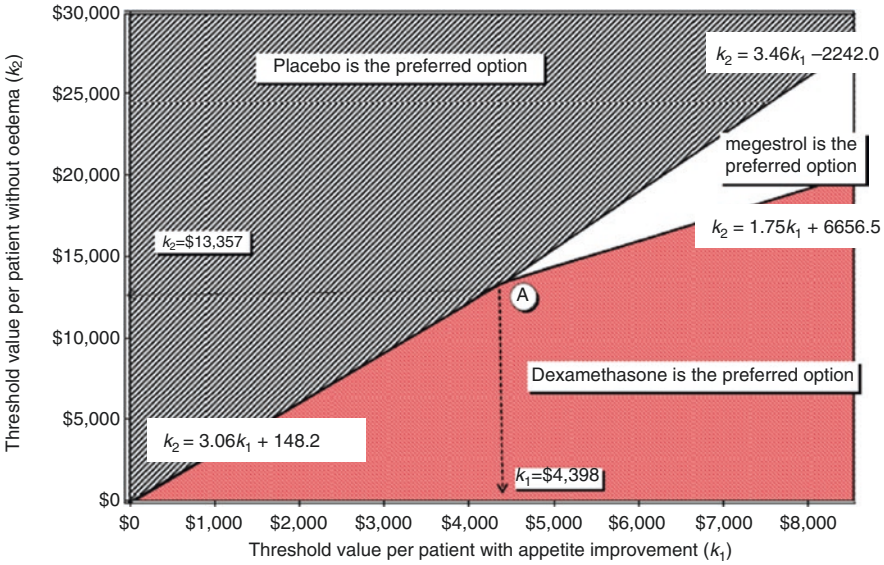


**Fig. 10.4** Regions over which each strategy is the preferred treatment option

Fig. 10.4). Furthermore, Fig. 10.4 illustrates when $k_1 > \$4398$ and $k_2 < \$13{,}357$ dexamethasone maximises mean NB and when $k_1 < \$4398$ and $k_2 > \$13{,}357$ placebo maximises mean NB.

#### 10.3.2.4 Stochastic Analysis

Expected Net Loss

At any given combination of value for effects, the ENL is estimated by averaging NL across all replicates where the NL for each strategy is measured relative to the NB maximising approach in each replicate. For example, when $k_1 = \$300$ and $k_2 = \$200$, the ENL of adopting dexamethasone is \$1, placebo is \$92, and megestrol is \$277. At these threshold values, expected net losses reflect that dexamethasone minimises NL in 9686 replicates, placebo minimises NL in 314 replicates, and megestrol does not minimise NL in any replicates.

Expected Net Loss Planes

Expected net loss planes for each strategy extend the concept of expected net loss curves for one outcome to map ENL across potential threshold values for multiple outcomes.

Figure 10.5 shows the ENL plane for megestrol. The ENL plane is formed by varying the threshold values for appetite improvement and oedema and recalculating

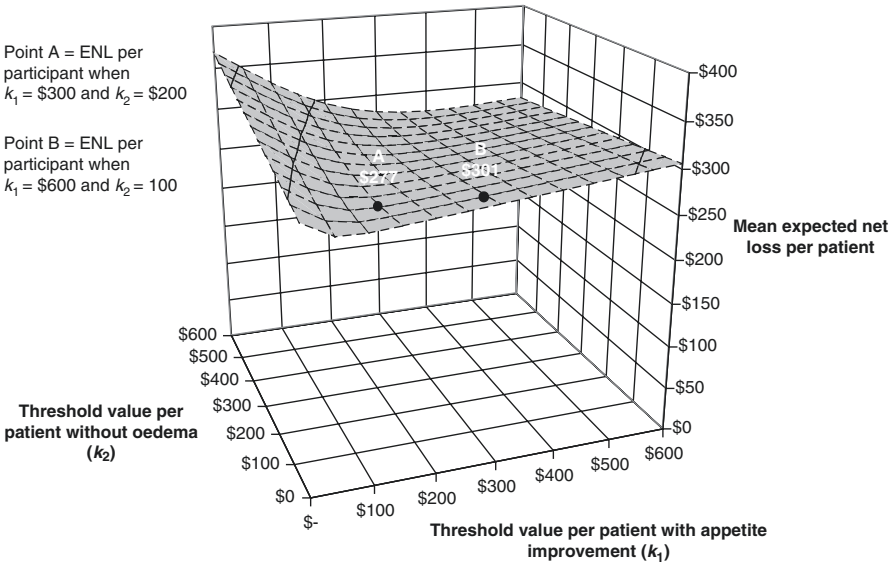


**Fig. 10.5** Expected net loss plane for megestrol

the average ENL across replicates for megestrol. The ENL at different combinations of values for effects can be readily identified from the ENL plane. Using the previous example, when $k_1$ = \$300 and $k_2$ = \$200, the ENL is \$277 if megestrol is adopted (point A, Fig. 10.5), whereas when $k_1$ = \$600 and $k_2$ = \$100, the ENL increases to \$301 (point B, Fig. 10.5).

## Expected Net Loss Contour

The ENL contour is formed by the lower bound of the ENL planes across the alternatives (Fig. 10.6) akin to the ENL frontier formed by the lower bound of ENL curves (Eckermann et al. 2008). The ENL contour readily identifies the strategy that minimises ENL (maximises ENB) at combinations of threshold values to inform risk-neutral and somewhat risk-averse decision making (Eckermann and Willan 2011). The shaded area indicates the combinations of $k_1$ and $k_2$, where placebo minimises the ENL. The unshaded area illustrates combinations of $k_1$ and $k_2$ where dexamethasone minimises the ENL. Megestrol does not minimise the ENL at the threshold value combinations illustrated in Fig. 10.6, as in the deterministic analysis in Fig. 10.4. Megestrol does not minimise NL (max. NB) until threshold values for appetite improvements are greater than \$4158 and threshold values for avoiding oedema are greater than \$11938.

When $k_1$ = \$300 and $k_2$ = \$200, dexamethasone minimises the ENL with an average loss in NB of \$1 per participant, as can be seen in the unshaded region in Fig. 10.6. This loss of \$1 per participant from choosing dexamethasone reflects that

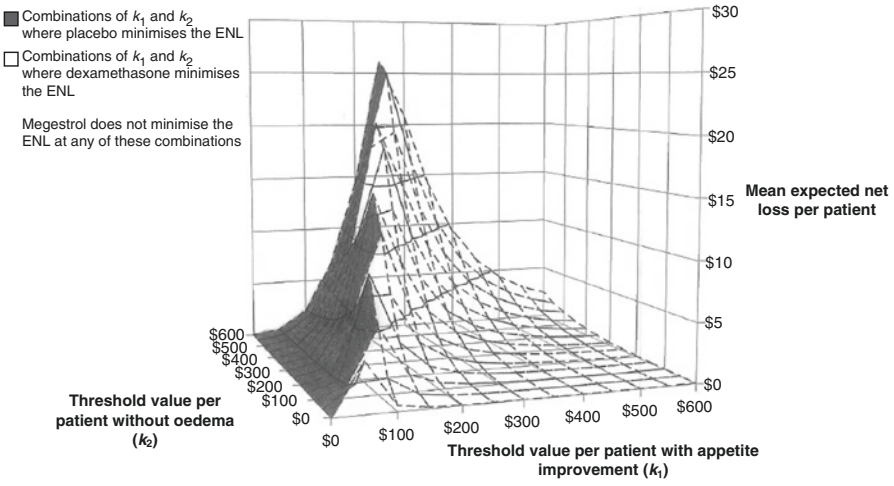


**Fig. 10.6** Expected net loss contour

placebo minimises ENL in 314/10,000 replicates and would be avoided with perfect information where the decision maker could pick the treatment minimising ENL in each realisation.

### Cost-Effectiveness Acceptability Planes

Cost-effectiveness acceptability planes (CEAP) show the probability of minimising NL (max. NB) at given combinations of value of effects for multiple outcomes.

Figure 10.7 shows the megestrol CEAP illustrating that at lower threshold values megestrol has little chance of minimising NL, whereas when $k_1$ and $k_2$ rise to $5000 each, megestrol has a 36% chance of being the preferred option given megestrol minimises ENL in 3600/10,000 replicates.

## 10.4 Discussion

When assessing the cost effectiveness of multiple strategies with multiple effects, analysis in C-DU space and the use of ENL planes and contours have been illustrated to overcome limitations arising with both conventional cost-effectiveness and cost-consequences analyses. When improved appetite was the sole effect measure considered, conventional partial analysis on the cost-effectiveness plane suggests only dexamethasone (with highest appetite improvement) and placebo (with lowest cost) are potentially optimal. If the oedema rate alone was considered then placebo was optimal in dominating (having lower cost and lower odema than)

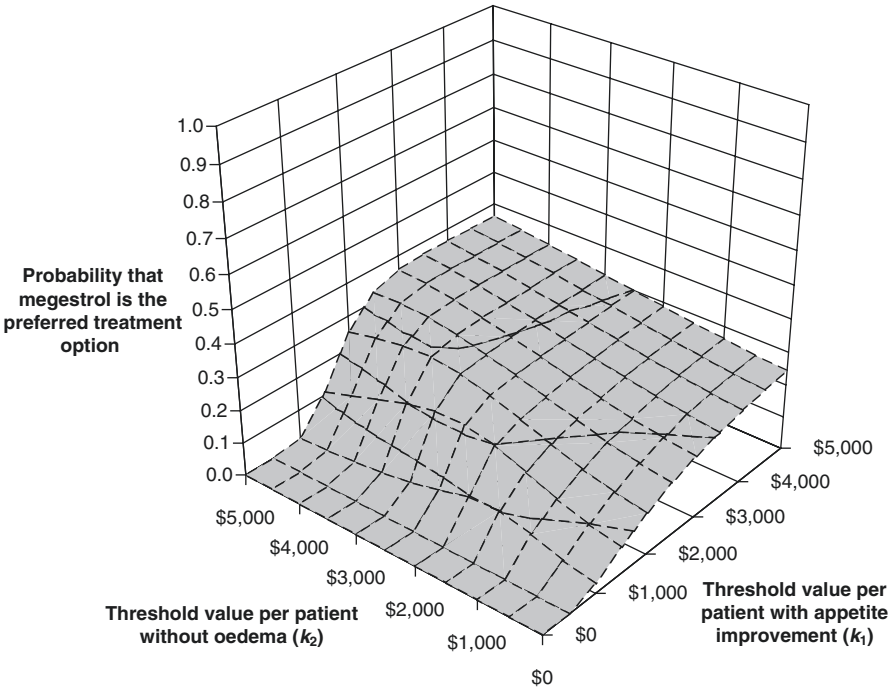


**Fig. 10.7** Cost-effectiveness acceptability plane for megestrol

magestrol or dexamethasone. Comparison in C-DU space and associated summary measures facilitates multiple outcome and multiple strategy inference, showing that no strategy dominates when incremental cost and multiple outcomes are jointly considered. That is, megestrol is still potentially optimal if threshold values for both avoiding oedema and appetite improvement are high enough.

The three strategies are clearly contrasted in C-DU space, facilitating identification, presentation and exploration of trade-offs between joint appetite improvement and oedema rates in optimising decision making. Collectively the new measures ensure appropriate joint coverage under uncertainty of multiple effects in estimating net clinical benefit and that their preferences and value across feasible threshold values are appropriately included in cost-effectiveness (ENL) assessment, summary measures and related decision making. Threshold regions display the combinations of values over which alternative strategies minimise mean NL across multiple effects and strategies. ENL planes present differences in ENL between alternatives across any set of joint threshold values. The ENL contour simultaneously establishes at any set of joint threshold values: (i) the intervention that minimises ENL and (ii) the EVPI with current evidence. Finally, CEAPs estimate the probability that alternate strategies maximise NB (min. NL) across multiple effects. This can be relevant where the strategy with the highest NB does not have the greatest probability of maximising NB. Note however that CEAPs, like CEACs, do not directly

identify the strategy that minimises ENL (Eckermann and Willan 2011), while ENL planes and contours do. To the extent that societal decision making may be somewhat risk averse (while asymptotically risk neutral across many decisions and populations under the Arrow-Lind theorem Arrow and Lind (1970)), the summary measures presented in this chapter allow decision makers to explicitly trade off return on investment (represented by ENL planes) with the risk in return (indicated by the probability of minimising NL in CEAPs) at different relative and absolute threshold values for multiple outcomes.

The methodology presented can be applied to any number and type of effect measures. However, while graphical comparison is feasible in two- and three-dimensional space for multiple outcome comparison, further dimensions cannot be easily represented in this manner. Despite this spatial limitation graphically, the general ENL formulae with radial properties in C-DU space can be mathematically extended to compare efficiency of strategies (or providers as in Chap. 9) across any number of dimensions with methods such as DEA. Further, the methods inherently promote transferability of the method across different populations and jurisdictions given different values can be imputed for each outcome depending on local perspective, context and absolute and relative values. Analysis in C-DU space can also provide an alternative framework for consideration of the multiple objectives of health-care services for decision making under uncertainty at a local and national level, such as health outcomes, the equitable distribution of health and access to services (Nord et al. 1995; Mooney et al. 2007; McKie et al. 2009).

Previously there have been only a few attempts to develop methods to evaluate and present costs and multiple outcomes under uncertainty in an explicit manner. Bjorner and Keiding (2004) proposed a relative cost-effectiveness measure. The Bjorner and Keiding approach compares the performance of each intervention relative to the worst performing intervention in a set of interventions using DEA (Bjorner and Keiding 2004). However, their relative cost-effectiveness measure does not inform decisions between non-dominated interventions, nor allow consideration of stochastic uncertainty, or allow estimation of the probability of maximising ENB. Negrin and Vazquez-Polo (2006) presented an alternative Bayesian cost-effectiveness framework. Their Bayesian methodology graphically represents the intervention most likely to maximise NB at various combinations of threshold values for multiple outcomes using a cost-effectiveness acceptability frontier. Geometrically, this is similar to the CEAP derived from comparison in C-DU space. However, neither of these previous approaches provides summary measures for quantifying differences in ENB between strategies under uncertainty, which is the critical comparison required to inform risk-neutral or somewhat risk-averse decision making (Arrow and Lind 1970; Zivin 2001; Eckermann and Willan 2011).

Other contemporary economic evaluation methods such as cost-consequences analysis (CCA) are limited when comparing interventions with multiple outcomes. Although CCA explicitly considers multiple effects and costs, cost and effect uncertainty are considered independently rather than jointly. Also, this approach does not allow for differences in ENB. Multiple outcome comparison in C-DU space retains advantages of CCA in enabling comaprison of multiple effects, while additionally

allowing rigorous assessment of joint uncertainty across multiple outcomes and estimation of differences in ENL (equivalently to differences in expected NB under the net benefit correspondence theorem). Furthermore, the ability of the methods developed to allow for different relative preferences (Hanson and Winzelberg 2013) is also valuable for informing individual patient as well as population-level decisions under uncertainty.

Multiple domain comparison in C-DU space also allows explicit exploration of relative and absolute values for multiple effects on preferred strategies. Further, this approach permits robust analysis of the impact of uncertainty around threshold values with ENL contours.

Conventional cost-effectiveness analysis, as illustrated in this chapter, is limited by a unidimensional outcome measure where choice of outcome can give rise to outcome selection bias. Cost-utility analysis is limited by narrowly defined health-related QOL domains in the commonly applied generic MAUIs which exclude important attributes in complex disease areas such as palliative care, e.g. utility from preparing for death (McCaffrey et al. 2014), and non-health outcomes such as place of palliative care (Agar et al. 2008). The novel methodology for multiple domain of effect comparisons presented in this chapter enables simultaneous and robust evaluation of such diverse domains of effect.

There are some pertinent issues to consider in interpreting the results of the modelled example and generalising the methods presented to other palliative and wider settings. One often-cited limitation of multiple outcome methodologies is that trade-offs are left to the decision maker. However, the proposed framework explicitly considers and presents trade-offs between alternative outcomes at potential combinations of decision maker threshold values in C-DU space with ENL planes, ENL contours and CEAPs.

The model used assumed that patient appetite and oedema outcomes are statistically independent, i.e. orthogonal. However, covariant relationships could be explicitly included with estimation of appropriate joint distributions if identified. For example, outcomes could have included both oedema and appetite, only oedema or appetite or neither appetite nor oedema, and each of these distinguishable combinations have threshold values that apply to their respective effects.

The cost effectiveness of megestrol, dexamethasone and placebo for the palliative treatment of anorexia-cachexia related to cancer has been evaluated with multiple effect comparison in C-DU space using two effects appetite improvement and oedema rates. Joint consideration across multiple effects enables consideration of trade-offs between expected costs and benefits under uncertainty to better inform societal decision making in relation to joint research and reimbursement decisions.

Alternatively, in the illustrated case, given the chosen outcomes (appetite improvement and oedema rates), a disease-specific MAUI instrument could have been used to incorporate both outcomes, and a cost-utility analysis conducted. However, such a substantial piece of work was beyond the resources of the research, while multiple outcome cost-effectiveness analysis in C-DU space was considered a more pragmatic, versatile and readily accessible approach. Computationally,

construction of the ENL planes, ENL contour and CEAPs are no more burdensome than creating NB curves and CEACs.

More generally, application of the proposed methodology in palliative and other settings where multiple effect domains and their joint consideration under uncertainty are key to preferences and decision making is suggested to be a valuable area for future research. Similarly, applying the values and preferences of individuals, different patient populations and societal decision makers across jurisdictions for relevant multiple outcomes should be considered in future research to further benefit from the flexibility of the methods illustrated. Importantly, the highly flexible nature of comparison on the cost-disutility plane under the net benefit correspondence theorem, in allowing robust comparison of multiple effects, strategies or providers in practice provides a powerful and robust research tool to satisfy coverage and comparability principles.

## 10.5 Conclusion

A novel approach to the analysis of the relative cost effectiveness with multiple outcomes using C-DU space and associated NL-based summary measures has been illustrated with a palliative care example. This readily accessible methodology has been shown to better inform societal decision making particularly in complex disease areas like palliative care where comparisons involve multiple health domains and non-health outcomes missing from generic MAUIs. Expected net loss planes and the ENL contour provide summary measures to fully inform risk-neutral societal decision making under the Arrow-Lind theorem (Arrow and Lind 1970). Cost-effectiveness acceptability planes additionally inform somewhat risk-averse societal decision making across such multiple domains. In summary, multiple outcome cost-effectiveness analysis with presentation in C-DU space and associated NL summary measures provides a systematic way of combining multiple outcomes in net benefit assessment under uncertainty, unlike conventional cost-effectiveness methods or cost-consequences analysis.

The methods enable a more robust consideration of trade-offs between costs and benefits, the consequences of funding decisions and the need for future research (Eckermann and Willan 2007; Koerkamp et al. 2007; Claxton et al. 2008; Eckermann et al. 2010).

In palliative care, analogous to the use of oedema and appetite domains in the illustration, key domains such as finalising affairs, family distress and carer effects urgently need to be incorporated into cost-effectiveness analyses to reflect community palliative care preferences and improve decision making related to investing in research, adoption of interventions and resource allocation. As highlighted in policy analysis of options for successful ageing in Chap. 12 until this occurs, assessment of palliative interventions runs the risk of supporting interventions, strategies, treatments or programmes which don't reflect palliative preferences or optimise outcomes from scarce resources on primary key domains for palliative care and undermining strategies and interventions that do.

## References

Agar M, Currow DC, Shelby-James TM, Plummer J, Sanderson C, Abernethy AP. Preference for place of care and place of death in palliative care: are these different questions? Palliat Med. 2008;22(7):787–95.

Argiles JM, Olivan M, Busquets S, Lopez-Soriano FJ. Optimal management of cancer anorexia-cachexia syndrome. Cancer Manag Res. 2010;2:27–38.

Arrow KJ, Lind RC. Uncertainty and the evaluation of public investment decisions. Am Econ Rev. 1970;60(3):364–78.

Australian Government Department of Health and Ageing. Medicare benefits schedule. Retrieved 11th March. 2009a. From http://www.health.gov.au/internet/mbsonline/publishing.nsf/Content/Medicare-Benefits-Schedule-MBS-1.

Australian Government Department of Health and Ageing. Pharmaceutical benefits schedule. Retrieved 11th March. 2009b. From http://www.pbs.gov.au/html/healthpro/home.

Berenstein EG, Ortiz Z. Megestrol acetate for the treatment of anorexia-cachexia syndrome. Cochrane Database Syst Rev. 2005;2:CD004310.

Biostat. Comprehensive meta analysis Version 2 [Computer program]. Englewood: Biostat Inc; 2005.

Bjorner J, Keiding H. Cost-effectiveness with multiple outcomes. Health Econ. 2004;13(12):1181–90.

Black I. Palliative medicine handbook. Cardiff: BPM Books; 2001.

Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of healthcare interventions. Health Technol Assess. 1999;3(2):1–134.

Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. Health Econ. 1997;6(4):327–40.

Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol. 1997;50(6):683–91.

Claxton K. Exploring uncertainty in cost-effectiveness analysis. Pharmaco Econ. 2008;26(9):781–98.

Claxton K, Briggs A, Buxton MJ, Culyer AJ, McCabe C, Walker S, Sculpher MJ. Value based pricing for NHS drugs: an opportunity not to be missed? BMJ. 2008;336(7638):251–4.

Coelli T. A guide to DEAP Version 2.1: A data envelopment analysis (computer) program. CEPA Working Papers No. 8/96. 1996.

Dy SM, Apostol CC. Evidence-based approaches to other symptoms in advanced cancer. Cancer J. 2010;16(5):507.

Eckermann S. Hospital performance including quality: creating economic incentives consistent with evidence-based medicine. PhD Dissertation, UNSW, Sydney; 2004.

Eckermann S, Coelli T. Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice. Soc Sci Med. 2013;76(1):159–68.

Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16(2):195–209.

Eckermann S, Willan A. Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies. Pharmaco Economics. 2011;29(7):563–77.

Eckermann S, Briggs A, Willan A. Health technology assessment in the cost-disutility plane. Med Decis Mak. 2008;28(2):172–81.

Eckermann S, Coory M, Willan A. Indirect comparison: relative risk fallacies and odds solution. J Clin Epidemiol. 2009;62(10):1031–6.

Eckermann S, Karnon J, Willan AR. The value of value of information: best informing research design and prioritization using current methods. Pharmaco Economics. 2010;28(9):699–709.

Fearon K, Strasser F, Anker SD, Bosaeus I, Bruera E, Fainsinger RL, Jatoi A, Loprinzi C, MacDonald N, Mantovani G, Davis M, Muscaritoli M, Ottery F, Radbruch L, Ravasco P, Walsh

D, Wilcock A, Kaasa S, Baracos VE. Definition and classification of cancer cachexia: an international consensus. Lancet Oncol. 2011;12(5):489–95.
Goebel M. Anorexia-cachexia syndrome in advanced cancer. J Palliat Med. 2010;13(5):627–8.
Good PD, Cavenagh JD, Currow DC, Woods DA, Tuffin PH, Ravenscroft PJ. What are the essential medications in pallative care? - a survey of Australian palliative care doctors. Aust Fam Physician. 2006;35(4):261–4.
Hanson LC, Winzelberg G. Research priorities for geriatric palliative care: goals, values, and preferences. J Palliat Med. 2013;16(10):1175–9.
Inui A. Cancer anorexia-cachexia syndrome: current issues in research and management. CA Cancer J Clin. 2002;52(2):72–91.
Italian Group for Antiemetic Research. Double-blind, dose-finding study of four intravenous doses of dexamethasone in the prevention of cisplatin-induced acute emesis: Italian Group for Antiemetic Research. J Clin Oncol. 1998;16:2937–42.
Koerkamp BG, Hunink MGM, Stijnen T, Hammitt JK, Kuntz KM, Weinstein MC. Limitations of acceptability curves for presenting uncertainty in cost-effectiveness analysis. Med Decis Mak. 2007;27(2):101(111).
López AP, Figuls MR, Cuchi GU, Berenstein EG, Pasies BA, Alegre MB, Herdman M. Megestrol acetate - probably less effective than has been reported! Authors' Response. J Pain Symptom Manag. 2005;30(1):5–6.
Loprinzi C, Kugler J. Randomized comparison of megestrol acetate versus dexamethasone versus fluoxymesterone for the treatment of cancer anorexia/cachexia. J Clin Oncol. 1999;17:3299–306.
Löthgren M, Zethraeus N. Definition, interpretation and calculation of cost-effectiveness acceptability curves. Health Econ. 2000;9(7):623–30.
McCaffrey N. Modelling joint cost and outcomes uncertainty on the cost-disutility plane – case studies in palliative care. PhD, Flinders University; 2013.
McCaffrey N, Karnon J, Currow D, Eckermann S. There's life in the old dog yet: modelling joint uncertainty in cost-consequences analysis. Flinders Centre for Clinical Change and Health Care Research: Adelaide/ South Australia; 2011.
McCaffrey N, Skuza P, Breaden K, Eckermann S, Hardy J, Oaten S, Briffa M, Currow D. Preliminary development and validation of a new end-of-life patient-reported outcome measure assessing the ability of patients to finalise their affairs at the end of life. PLoS One. 2014;9(4):e94316.
McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Better informing decision making with multiple outcomes cost-effectiveness analysis under uncertainty in cost-disutility space. PLoS One. 2015;10(3):e0115544.
McKie J, Shrimpton B, Richardson J, Hurworth R. Treatment costs and priority setting in health care: a qualitative study. Aust New Zealand Health Policy. 2009;6:11.
Mooney G, Coast J, Jan S, Ryan M, Wiseman V. Searching for a threshold – not so NICE. J Health Ser Res Policy. 2007;12(3):190–1.
Negrin MA, Vazquez-Polo FJ. Bayesian cost-effectiveness analysis with two measures of effectiveness: the cost-effectiveness acceptability plane. Health Econ. 2006;15(4):363–72.
Nord E, Richardson J, Street A, Kuhse H, Singer P. Maximizing health benefits vs egalitarianism: an Australian survey of health issues. Soc Sci Med. 1995;41(10):1429–37.
Pharmaceutical Benefits Advisory Committee. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.4). Department of Health, Commonwealth of Australia: Canberra; 2013.
The Cochrane Collaboration. Review Manager (RevMan) [Computer program]. Version 5.0. Copenhagen: The Nordic Cochrane Centre; 2008.
Tisdale MJ. Cancer cachexia. Curr Opin Gastroenterol. 2010;26(2):146–51.
van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ. 1994;3(5):309–19.
Zivin JG. Cost-effectiveness analysis with risk aversion. Health Econ. 2001;10(6):499–508.

# Part IV
# The Health Shadow Price and Other Key Political Economy and Policy Issues: Appropriate Threshold Pricing and Policy Application of Methods for Optimising Community Net Benefit with Budget Constraints

Part IV addresses key political economy issues arising in budget-constrained health systems attempting to optimise decision making in better using and integrating existing and new technology starting from characteristic market failure and allocative and displacement efficiency conditions. The health shadow price of Pekarsky (2012, 2015) is shown in Chap. 11 to provide a pathway to allocative efficiency by establishing an economically meaningful threshold value for effects in net benefit assessment, central to all decision making across joint research, reimbursement and regulatory decisions in practice in Fig. p4.1.

The health shadow price in doing so addresses research and reimbursement (adoption and displacement) biases against better use of existing technology and towards new technology that underlie many current health-system inefficiencies. Chapter 12 addresses policy options for meeting the twenty-first-century challenge of successful ageing of the baby boomer population. In particular highlighting promising approaches to age and dementia friendly communities, age care facility design and lower cost while more effective and palliative preferences appropriate alternatives for palliative care.

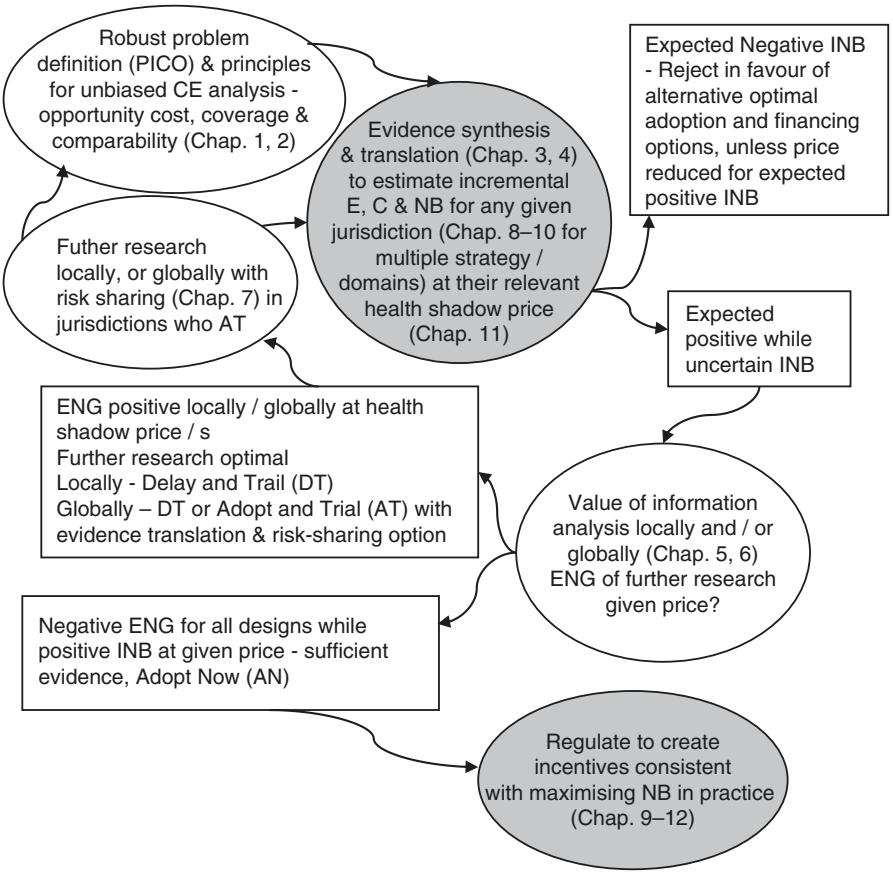


**Fig. p4.1** Optimal decision making cycles for joint research, reimbursement, and regulatory processes locally and globally

## References

Pekarsky B. Trusts, constraints and the counterfactual: reframing the political economy of new drugs. PhD thesis University of Adelaide; 2012.

Pekarsky BAK. The new drug reimbursement game: a regulator's guide to playing and wining. London: Springer; 2015.

# Chapter 11
# The Health Shadow Price and Economically Meaningful Threshold Values

## 11.1 Overview

In this chapter we make clear the importance of an economically meaningful threshold value reflecting opportunity costs (best alternative actions) for optimising budget-constrained societal decision making across research, reimbursement and regulatory processes. Drawing on the research of Pekarsky (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014), historical attempts to assert or arrive at economically meaningful threshold values or rules, to compare new technologies against for the usual case of interest where new technologies have expected additional incremental cost and effects, are critiqued. These historical attempts include notions of a critical ratio, willingness to pay thresholds and various forms of displaced services thresholds. However, in each case, these thresholds are shown to fail to reflect opportunity cost of joint adoption and financing actions with constrained and fixed budgets in reimbursing such new technologies. These thresholds more generally are shown to not provide a pathway towards allocative efficiency in practice where manufacturers have market power and price up to a threshold, and health systems start from a position of allocative and displacement inefficiency in provision of current services and programmes. Allocative and displacement inefficiency are characteristic of health-care systems in the presence of market failure for evidence on the cost-effectiveness of existing and non-patented or patentable services and programmes and their best expansion and contraction (Arrow 1963, Pekarsky 2012).

Critically, the health shadow price for reimbursement derived by Pekarsky (Pekarsky 2012, 2015) is shown to enable an economically meaningful threshold value for net benefit (Eckermann and Pekarsky 2014) that reflects opportunity costs of joint adoption and displacement actions for general health system conditions of allocative and displacement efficiency, to overcome deficiencies of historical thresholds.

Pekarsky (2012, 2015) demonstrates that with a fixed budget investing in new technologies with higher expected net costs to a health system should be determined by identifying the opportunity cost or best alternative action joint adoption and



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_11

displacement actions that budget-constrained reimbursement involves. The highest value alternative to investing in such technologies is the most cost-effective expansion of current services (ICER = $n$) funded by displacement of least cost-effective current programmes and services (ICER = $m$). Hence, the highest value alternative is optimal unless the ICER of new technology with net costs requiring financing is equal to or below the health shadow price, $\beta c = \left(\frac{1}{n} + \frac{1}{d} - \frac{1}{m}\right)^{-1}$. This health shadow price is derived as shown in Sect. 11.6 from equating investment return from new technology (ICER in adoption βc) financed by displacement of services (ICER $d$) with best expansion in adoption (ICER = $n$) and contraction in financing (ICER = $m$). Importantly this health shadow price reflects allocative inefficiency ($n < m$) and displacement inefficiency ($d < m$) characteristic of health systems and appropriately values and creates incentives for addressing market failure in provision of research and evidence for displaced services and non-patented technology required to avoid allocative and displacement inefficiency (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014).

The health shadow price differs qualitatively as well as quantitatively where new technologies lie on the SW quadrant and hence are expected to save net costs while having lower net effect.

In the case of new technologies on the SW quadrant with lower net costs and effects Eckermann (2015) clarifies that the health shadow price differs qualitatively as well as quantitatively. In such cases if the health budget is free to contract, then the health shadow price above which funds generated outweigh health losses is the best alternative action, contraction of least cost-effective activities with ICER m.

For cases where the budget is strictly fixed in contraction as well as expansion, such that the funds generated by a cost saving new technology have to be spent on adoption, a new health shadow price result is derived on the SW quadrant. In that case the health shadow price is derived in finding the threshold ICER equating returns of generating funding F with the cost saving technology used to finance adoption (ICER a) with that of the best alternative fund generating and adoption actions – contraction with ICER m to fund expansion with ICER $n$. This leads to a health shadow price with a strictly fixed budget in contraction for cost saving technologies above which they represent the best fund generating option of $\beta f = \left(1/a + 1/m - 1/n\right)^{-1}$.

Strictly the health shadow price of budget contraction to generate funds (on the SW quadrant) only coincides with the health shadow price $\beta c = \left(\frac{1}{n} + \frac{1}{d} - \frac{1}{m}\right)^{-1}$ for reimbursing new technology with additional costs (on the NE quadrant) at the single point of perfect allocative and displacement efficiency, where $n = m = d = a$. Differences between shadow prices for new technologies that have additional cost and require displacement of existing programmes and those where new technologies are cost saving are reflected in a kink in the threshold line on the CE plane where the threshold on the SW quadrant is greater than that on the NE quadrant (Eckermann 2015). The extent of the kink is dependent on the extent of health system allocative ($n < m$), displacement ($d < m$) and adoption ($a > n$) inefficiency. These health shadow prices together support best actions in expansion and contraction consistent with programme budgeting and marginal analysis (PBMA) principles.

## 11.2 Why Are Economically Meaningful Threshold Values Critical

In Chaps. 2 to 10, we have shown that accepting the incremental cost ratio (ICER) of a strategy, intervention or new technology at a threshold value for effects of care, $\lambda$, is equivalent to maximising net benefit (NB) or equivalently minimising net loss (NL). That is, if we consider the usual case of interest where there is positive expected incremental effects and costs, ($\Delta C/\Delta E < \lambda$ where $\Delta E > 0$) or the less studied case of the ICER being above a threshold value for a negative effect ($\lambda < \Delta C/\Delta E$ where $\Delta E < 0$) (Willan and Briggs 2006), then:

(i) INB = $\lambda\Delta E - \Delta C > 0$ in any two-strategy comparison as in chapters 2-7; and more generally;
(ii) Max NB = $\lambda E - C \leftrightarrow$ min NL = $\lambda$DU + $C$ across multiple strategy and/or multiple outcome comparisons under the net benefit correspondence theorem, as shown in Chaps. 8 and 10 (Eckermann 2004, Eckermann et al. 2008; Eckermann and Willan 2011; McCaffrey et al. 2015).

Importantly NB and NL metrics unlike the ICER have robust statistical properties in ordering strategies in two strategy (Willan and Briggs 2006), multiple strategy or multiple effect cost effectiveness comparisons, while each making explicit the decision rule for adopting conditional on threshold values for effects, or $\lambda's$(Eckermann et al. 2008; Eckermann and Willan 2011; McCaffrey et al. 2015). More generally, maximising NB and minimizing NL have been shown to enable best summary measures (INB curves, ENL curves and frontiers) for cost effectiveness societal decision making and underlie optimal decision making under uncertainty across joint investment decisions for research reimbursement and regulation.

INB assessment applying coverage and comparability principles with joint consideration of cost and effects conditional on $\lambda$ was highlighted as the basis for informing unbiased cost-effectiveness analysis. This was the case for robust analysis whether trial based (Chap. 2), model based evidence synthesis, translation and extrapolation (Chap. 3) or both in jointly satisfying coverage and comparability principles. Similarly, jointly satisfying coverage and comparability conditions was extended in Chap. 4 to community-based health promotion programmes with the use of multiplier and network methods and to palliative care settings with relevant multiple domains of effect considered in greater detail in Chap. 10.

In allowing for whether further research should be required given decision uncertainty, Chaps. 5, 6 and 7 showed joint research and reimbursement decisions under uncertainty can be optimised in maximising expected value to cost or return on investment allowing for key decision contexts. In each case optimal trials or research designs are explicitly conditional on INB distributions relevant to conditions in each jurisdiction. This is the case both in Chap. 5 for locally optimal (Eckermann and Willan 2007, 2008a, 2008b) and Chap. 6 and 7 globally optimal trial design (Eckermann and Willan 2009, 2013), with translatable evidence providing feasible and valuable options for adopting and trialling and risk sharing, and extended to

optimal implementation and pricing (Willan and Eckermann 2010, 2012). Finally, Chaps. 8, 9 and 10 highlighted the NL metric and the net benefit correspondence theorem (Eckermann 2004; Eckermann et al. 2008; Eckermann and Willan 2011; Eckermann and Coelli 2013; McCaffrey et al. 2015). They provide the basis for robustly and optimally informing societal decision making with multiple strategy, multiple provider (efficiency) or multiple outcome comparison of cost-effectiveness and efficiency analysis in practice. Importantly, robust analysis is conditional on economically meaningful threshold value for each of these metrics and all summary measures in performance comparison except for technical efficiency and industry shadow price values in Chap. 9.

Consequently, in each case optimising across research, reimbursement and regulatory decisions such as that considered in Chaps. 2 to 10 are conditional on having economically meaningful threshold values for health effects, or $\lambda s$ in NB and NL assessment. Maximising budget-constrained population health and other objectives from a given budget whether in reimbursement (adoption and displacement), research or pricing decisions in any jurisdiction requires threshold values that reflect best alternative actions given decision contexts (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014). That is, threshold values for NB (Graham 1981, 1992) and NL (Eckermann et al. 2008) need to reflect opportunity cost (Eckermann and Pekarsky 2014; Eckermann 2015) in order to optimise from constrained budgets. Further, economically meaningful threshold values are also required to create appropriate incentives for allocation of resources across joint research, reimbursement and pricing decisions. Also, in regulation of their allocative and net benefit efficiency performance allowing for cost and quality of care – in optimising budget-constrained performance in practice allowing for existing and new technology and its pricing (Eckermann 2004; Eckermann et al. 2010; Pekarsky 2012; Eckermann and Coelli 2013).

To address the clear need for economically meaningful threshold values for net benefit we begin following Eckermann and Pekarsky (2014), by first critiquing the basis of various threshold values proposed historically, before deriving the health shadow price solution. The health shadow price is first derived for the usual case of interest of reimbursing a new strategy or intervention with additional costs within a constrained budget and given current health system inefficiency. The less considered case on the SW quadrant of new strategies or interventions with net expected cost savings while lower effects is later tackled following (Eckermann 2015) and extended to derive a new health shadow price result in this quadrant.

## 11.3 Historical Threshold Values and Opportunity Costs

The critical nature of the threshold value for cost-effectiveness decisions was first highlighted by Weinstein and Zeckerhaus (1973) suggesting it represents a critical ratio when contemplating allocative efficiency in health care. They noted total health would be maximised from a budget if all services were ranked by their cost-effectiveness and funding allocated according to this ranking up to the budget. Based on such ordered funding, the threshold value at this point is consistent with

the cost per effect of the last service financed and the least cost-effective programme of services financed. However, note that ordered ranking of services up to a budget counterfactually assumes both complete allocative efficiency and that services or programmes are discrete (i.e. cannot be expanded or contracted). In reality programmes can be expanded and contracted, while health systems do not have competitive market characteristics to promote allocative efficiency (Arrow 1963). Further, note that consideration was not given to whether the use of this threshold as a decision rule provides a pathway to allocative efficiency starting from a point of inefficiency, nor considers the impact of strategic behaviour such as pricing up to a threshold (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014).

Later Weinstein and Stason (1977) in their foundations of cost-effectiveness analysis paper adopted the same logic as Weinstein and Zeckerhaus (1973) but recognise pressure of medical decision makers and consider societal willingness to pay (SWTP) as a threshold where the budget expands to allow this. In the 1990s league tables emerged for ICERs of what was currently funded to compare ICERs for new projects against on grounds of 'consistency and fairness', reflecting notions of societal willingness to pay (SWTP). For example, in 1993 the first international health technology assessment process employing cost-effectiveness analysis began with the Pharmaceutical Benefits Advisory Committee (PBAC) in Australia employing such league tables for comparable funded technologies or programmes in evaluating new technologies (PBAC 2013). During this period Birch and Gafni (1992) highlighted and stressed that $\lambda$ should be the opportunity cost, defined as the cost per effect of the most cost-effective service that would otherwise be financed with money invested in new technology, not the least cost-effective of current services to contract. Johannesson and Weinstein (1993) responded to Birch and Gafni (1992) by suggesting that while $\lambda$ should theoretically reflect this opportunity cost, the use of SWTP is applied in practice in the absence of evidence for the most cost-effective alternatives. SWTP as a threshold does not consider nor appropriately change with alternative options or budget constraints. After 1993 the SWTP was used extensively, while Birch and Gafni are continuing to argue against it, including papers with apt titles such as 'ICERs: the silence of the lambda' (Gafni and Birch 2006), noting the sub-optimality and health and resource cost of any approach that does not reflect budget-constrained opportunity costs. This also resonated with the budget constrained appropriate threshold for NB assessment suggested in Graham (1981, 1992).

Truemen et al. (2001) admit budget impact should be considered. These sentiments are intensified over the period to 2007 with crises emerging with NICE decisions in the UK priced at SWTP such as Herceptin, where mandated use had significant impact in displacing other health-care services (Barrett et al. 2006). In large part in response to this, various UK-based authors (Culyer et al. 2007; McCabe et al. 2008; Claxton et al. 2008, 2013; Griffin et al. 2008) suggest shifting to displaced services (ICER = $d$) as a threshold to compare new interventions against.

Following Eckermann and Pekarsky (2014), we consider the case for these various displaced service thresholds and whether they reflect opportunity costs or provide an economically meaningful threshold value that provides a pathway towards allocative efficiency.

## 11.4 Considering Displaced Services as a Threshold: The Straw Man Outside the Room

Across displaced service proponents, an attempt at a theoretical basis for employing displaced services as the threshold to compare new technologies against is most clearly stated in Griffin et al. (2008: 24) with a two-part argument referencing the shadow price of the budget constraint.

The two parts of the argument suggest that:

(i) The incremental cost and health outcomes of marginal services that would be displaced determine the shadow price (incremental QALY/ additional unit cost) of the budget constraint.
(ii) The threshold for the ICER (incremental cost per QALY) of new treatments should in principle represent the inverse of the shadow price of the budget constraint.

The two parts of this argument are later combined by Griffin et al. (2008:24) to assert that new treatment should be reimbursed where incremental health offered by the new treatment option exceeds the health foregone with displacement of marginal programmes (Eckermann and Pekarsky 2014).

Other proponents of displaced service thresholds either explicitly or implicitly assert various forms of displaced services (least cost-effective current services that should be displaced or actual current or historical displaced services) as threshold ICER and/or opportunity cost of new investment. For example, Sculpher and Claxton (2012:133) assert that health outcomes foregone due to the displacement of existing services in funding additional cost of new programmes and technologies reflect the opportunity costs of new programmes and that the threshold should reflect this.

Strictly, as considered in Eckermann and Pekarsky (2014), there are four distinct types of threshold values related to displaced services presented by proponents:

(i) The least cost-effective current programme, assuming that this is the programme that is actually displaced to finance the additional costs of the new technology (Culyer et al. 2007; Griffin et al. 2008)
(ii) The least cost-effective programme, regardless of whether displaced or not (McCabe et al. 2008)
(iii) The ICER of the services actually displaced to finance that technology regardless of the ICER of that displaced service relative to other services (Claxton et al. 2008; Sculpher and Claxton 2012)
(iv) The average ICER of National Health Services (NHS) services displaced historically (Claxton et al. 2013)

If one restricts consideration of displacement to optimal displacement where displaced services are the least cost-effective of current services (programme or technology), then these four arguments for threshold values coincide. That is, they would each result in a threshold value comparing new technologies with the least cost-effective services.

If displacement were optimal (i.e. least cost-effective technologies were displaced), then key related questions that naturally arise for the common displacement threshold definitions under this assumed conditions include whether holding new technology accountable to a threshold value of least cost-effective services displaced:

(i) Reflect the opportunity cost of adopting those new technologies; and
(ii) Provide a pathway towards allocative efficiency and optimal allocation of budgets.

To illustrate and practically consider such questions, consider a hypothetical typical situation that might arise where the least cost-effective service to be displaced has an ICER of $200,000 per QALY and displacing this service would finance a $100 million investment per year required for the new technology. In that case holding the new technology accountable to the ICER of displaced services, the new technology could have an ICER up to $200,000 per QALY, say $199,000 per QALY, and claim to have net benefit from adoption of that technology and displacement of the least cost-effective technology.

Indeed, displaced threshold proponents in that case would claim an annual net health gain of 2.5 QALYs per year ($100 million/199,000 – $100 million/200,000 = 502.5–500) and that the opportunity cost of investing the $100 million per year in the new technology is the 500 QALYs per year of displaced services.

However, the best alternative adoption action or opportunity cost to investing in the new technology in that case is investing the $100 million from displacing the least cost-effective service in the best expansion of existing services. For the sake of argument, assume the best expansion of existing services in a $100 million per year expansion of existing services has an ICER of $10,000 per QALY. In that case the best alternative action and hence opportunity cost of investing $100 million per year in the new technology is $100 million/10,000 = 10,000 QALYs per year, not the 500 QALYs suggested by the use of displaced services.

Consequently, if displacement of services were efficient, with the least cost-effective services displaced, then the use of the ICER for displaced services as the threshold rather than leading to a 2.5 QALY gain, would result in a loss of 9497.5 QALYs relative to the highest value alternative of best expansion of current services.

This hypothetical empirical example simply illustrates problems of displaced services not reflecting opportunity cost but also raises the theoretical question as to why the logic of the attempted two-part argument for displaced services as reflecting opportunity presented in Griffin et al. (2008) fails. The two-part argument fails theoretically as it conflates shadow prices in expansion and contraction to misrepresent opportunity cost as the lowest value alternative to be displaced in contraction, rather than the highest value alternative to be expanded. The lack of a relationship between the two parts of their argument for displaced services represent opportunity costs would have been made clear if the first part had clarified this was for the shadow price in contraction and the second part for opportunity cost of investment the shadow price in expansion. That is, it becomes clear that the two parts do

not represent the same shadow price and generally can't be combined, if the text is added as per that italicised in (i) and (ii) below:

(i) The incremental cost and health outcomes of marginal services that would be displaced determine the shadow price (incremental QALY/ additional unit cost) of the budget constraint *in contraction*.
(ii) The threshold for the ICER (incremental cost per QALY) of new treatments should in principle represent the inverse of the shadow price of the budget constraint *in expansion*.

Each part has merit when considered independently and clarified in relation to which shadow prices they reflect. However, when considered together without clarifying one shadow price is in expansion and the other in contraction, they conflate these shadow prices to misrepresent opportunity cost as the lowest rather than highest value alternative. That is, appropriately adding 'in expansion' to the end of Part (i) and 'in contraction' to Part (ii), it becomes clear that the health shadow prices referred to in Parts (i) and (ii) generally differ with characteristic allocative inefficiency conditions of health systems. That is, the health shadow price in contraction to determine programs to get rid of can be many orders of magnitude greater than that in expansion below which to invest in. This is particularly the case across health systems given allocative efficiency is theoretically expected and in practice observed with characteristic lack of conditions for competitive markets and displacement efficiency with market failure in providing CE evidence for displaced services (Arrow 1963; Pekarsky 2012, 2015; Eckermann and Pekarsky 2014).

Hence, the shadow prices considered in Parts (i) and (ii) can substantially differ quantitatively with allocative inefficiency characteristic of health systems, as well as being diametrically opposed qualitatively given adopting below a threshold for cost per unit effect with expansion and above a theshold cost saving per unit effect reduction with contraction. That is, where the ICER for the most cost-effective expansion of current programmes ($n$) is less than the ICER of least cost-effective services in contraction ($m$) (Pekarsky 2012; Eckermann and Pekarsky 2014). This conflation is consequently critical both qualitatively and quantitatively in health care in misrepresenting opportunity cost of investing in new technology with additional costs as the lowest value alternative action (estimated by the health shadow price of the budget in contraction) rather than the highest value action (estimated by the shadow price of budget in expansion).

Indeed, such clarification would have made clear that the two-part argument presented in attempting to justify displaced services in contraction as the opportunity cost of investing in new technology and the ICER of displaced services and the threshold value relies on already being at a point of complete allocative efficiency. That is, assuming that the health system is already at the singular point where the ICER of the most cost-effective programme or service in expansion coincides with that of the least cost-effective programme or service in contraction ($n = m$).

## 11.5 Distinct Dangers of Using Displaced Service Thresholds Over Time, Whether Assumed or Actual and Applied Inconsistently or Consistently

Displaced service thresholds can easily lead to reductions in budget-constrained health effects from processes of displacement to finance adoption of new technologies. This is expected to arise directly in the case of a least cost-effective service displacement definition, simply noting that displacement is characteristically not efficient. For example, inefficient displacement characteristically arises where services that are displaced to fund patented new technologies, rather than being least cost-effective in contraction or displacement are often not patented or non-patentable services (Pekarsky 2012). That is, services such as palliative and rehabilitative care services, health promotion and prevention programmes that do not have the active vested interest of patented services to obtain cost-effectiveness evidence for expansion or contraction. More generally market failure in displacement evidence arises with lack of a vested interest to research least cost-effective services in practice to be displaced.

In the absence of societal decision-maker processes or institutions for undertaking cost-effectiveness research for non-patented services from trials and/or evaluation of non-patented or patented programmes in practice, it cannot be objectively determined whether existing programmes should be expanded or contracted, let alone displaced. Hence alongside trial evidence, processes such as programme budgeting and marginal analysis (PBMA) are key to undertaking objective expansion, contraction or displacement (Ruta et al. 1996, 2005). Otherwise unpatented services can be treated as though they are least cost-effective services and as in the UK displaced to finance reimbursement of patented services (Barrett et al. 2006). Indeed in the UK, it was the mandated use of a series of patented technologies after NICE approval by health districts that had significant impact in displacing such non-patented health-care services and programmes.

In light of such inefficient displacement processes characteristically arising in practice with market failure for displacement evidence, the assumption of efficient displacement ($d = m$), initially considered to allow consistent displaced services definitions across 1–4, clearly needs to be relaxed. That is, to allow for the realistic, more general and characteristic case in health care where $d < m$ and the displacement of services to finance adoption is not efficient.

Where $d < m$ but is treated as though $d = m$ (McCabe et al. 2008 explicitly and Culyer et al. 2007; Griffin et al. 2008 by assumption), then it should be clear that with manufacturers expected to price up to a threshold of $m > d$, then processes of adopting and financing unequivocally lead to health losses from reimbursement. For example, consider where m, the ICER of least cost-effective current services (in contraction), were £30,001 per QALY and *d,* the ICER of services actually displaced were £15,000 per QALY and the most cost-effective service (in expansion) were £5000 per QALY. Then consider adopting a £30 million per year new technology (drug or medical device) when priced up to just below the threshold of £30,001

per QALY, at say £30,000 per QALY. The process of reimbursement, adopting and financing, the new technology would then lead to a net loss to the health system of 1000 QALYs per year given the new technology is expected to lead to 1000 QALYS (£30 million/£30,000 per QALY) per year, but services and programmes displaced reduce an expected 2000 QALYs (£30 million/£15,000 per QALY).

However, note that the opportunity cost of reimbursing the new technology with a least cost effective displaced threshold is significantly greater than that, given £30 million invested in adopting the best expansion of existing technology would have led to an expected 6000 QALYs (£30 million/£5000 per QALY), while optimal financing with displacement (or contraction) of least cost-effective services would have offset this by less than 1000 QALYs (£30 million pounds/£30,001 per QALY). Hence, rather than a 1000 QALY net loss to the health system optimal actions would have lead to a more than 5000 QALY gain. Consequently the net opportunity loss of applying a least cost effective displaced threshold relative to optimal reimbursement (adoption and displace meant actions) is more than 6000 QALYs, 5000 due to suboptimal adoption and more than 1000 due to suboptimal displacement.

Alternatively if the threshold value for new technology were set at an observable ICER for displaced services of $d < m$ (of £15,000 per QALY), that still does not reflect the opportunity cost (highest value alternative action) of the adoption decision with most cost-effective expansion of existing programmes and technology (of £5000 per QALY), with reimbursement actions in relation to adoption and financing both being able to be significantly improved.

Indeed, where new technology (pharmaceutical medication, device, etc.) manufacturers are expected to strategically price almost or actually up to a threshold, then respectively a minute little or no net immediate gain to the health-care system is expected from joint actions with a fixed budget of displacing or contracting any current service or programme to finance adoption of any new technology. Hence, using the ICER of observable displaced services as the threshold, rather than having a more than 5000 QALY per year gain from optimal adoption and displacement actions reflected in the health shadow price would be little or no immediate QALY gain immediately expected from reimbursing (adopting and financing) the new technology. Further, given pricing of new technologies up to such displaced services thresholds if this threshold rule were consistently applied across patented and unpatented technologies and programmes, then new technologies would be in line to be displaced in the next or future cycles. Consequently new technologies adopted with pricing up to the threshold would be expected to be cycled through in the next or subsequent cycles. Hence reversal cost for new technologies such as unamortised training and capital costs and reversal of public health messages, considered in Chap. 7, would additionally be faced if displaced services decision rules were consistently applied.

Critically from a health system perspective, such costs of reversal with expected pricing up to any (least cost effective or actual) displaced service threshold would be expected to lead over time away from, rather than provide a pathway to, budget-constrained allocative efficiency. That is the health system would be expected to go backwards rather than forwards with any potential minute immediate gains from pricing up to a displaced service threshold more than wiped out with costs of reversal from cycling through such investments. Hence, very real dangers generally

arise with the use of displaced services threshold values for new technology over time. They don't just fail to optimise budget-constrained population health of decisions in not reflecting the opportunity cost of best alternative investment and disinvestment, but can be expected to actively reduce budget-constrained population health over time.

This is the case with expected manufacturer pricing up to threshold levels for a displaced service threshold whether:

(i) Applied inconsistently across patented and non-patented technologies and programmes in displacing non-patented services treated as though least cost-effective in the absence of evidence of their ICER (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014); or
(ii) Applied consistently across patented and unpatented services cycling through new technologies priced up to threshold values resulting in addition costs of reversal not accounted for in cost-effectiveness estimates and ICERs (Eckermann and Pekarsky 2014).

In general then distinct problems and dangers emerge from employing displaced services as a threshold decision rule over time whether applied consistently or not.

The central problem of displaced services thresholds failing to represent opportunity cost or allow a pathway to allocative efficiency in theory and practice also applies to least cost-effective current service thresholds suggested as far back as Weinstein and Zeckhauser (1973). Having established central problems with such historically posed threshold values, the natural question which we now address is what threshold value does represent opportunity costs of reimbursing new technology and provide an appropriate pathway towards allocative and displacement efficiency over time?

## 11.6 The Health Shadow Price for Reimbursement (Adoption and Financing)

Pekarsky (2012, 2015) derived the health shadow price, $\beta c$, in assessing whether to reimburse a new technology or programme with incremental net cost in a budget-constrained health system for the general case of displacement inefficiency ($d < m$) and allocative inefficiency ($n < m$). This derivation identifies that as reimbursement involves both adoption and financing actions, the highest value alternative to reimbursing (adopting and financing) such a new technology with net investment of I is to adopt the most cost-effective expansion of existing services (ICER = $n$) funded by the displacement of the least cost-effective services (ICER = $m$). Hence, the health shadow price is derived equating the outcomes of adoption (ICER of $\beta$c, the health shadow price threshold) and financing (ICER of displaced services = $d$) actions in reimbursing a new technology with that from optimal adoption (ICER = $n$) and financing (ICER = $m$) of the same investment amount ($I$). That is, the health shadow price or threshold ICER for effects is derived equating the health gain from actual adoption and displacement for an investment amount I

with that of optimal adoption and displacement for the same investment amount I and hence:

$$\frac{I}{\beta c} - \frac{I}{d} = \frac{I}{n} - \frac{I}{m}$$

Dividing through by $I$ and rearranging, we arrive at the health shadow price of Pekarsky (2012, 2015):

$$\frac{1}{\beta c} = \frac{1}{n} + \frac{1}{d} - \frac{1}{m}$$

$$\beta c = \left( \frac{1}{n} + \frac{1}{d} - \frac{1}{m} \right)^{-1}$$

where $d$ is the ICER of displaced services,
$n$ is the ICER of most cost-effective service expansion and
$m$ is the ICER of least CE services in contraction.

Importantly $\beta c$ appropriately allows for allocative but also displacement inefficiency characteristic of health systems. Allocative and displacement inefficiency is characteristic of health systems with imperfect information, uncertainty, lack of competitive market conditions and market failure in providing CE evidence for displaced services (Arrow 1963; Pekarsky 2012, 2015; Eckermann and Pekarsky 2014). Health systems in practice consequently only have the ability to move resources with contraction and expansions at the margins and require time, information and consideration of critical decision contexts to optimize joint research and reimbursement decisions.

If there is allocative inefficiency ($n < m$), while displacement is efficient ($d = m$), then

$$\beta c = \left( \frac{1}{n} + \frac{1}{d} - \frac{1}{m} \right)^{-1} = \left( \frac{1}{n} + \frac{1}{m} - \frac{1}{m} \right)^{-1} = n$$

However, more generally for the usual case in health systems where there is allocative ($n < m$) and displacement inefficiency ($d < m$),[1] then

$$\beta c = \left( \frac{1}{n} + \frac{1}{d} - \frac{1}{m} \right)^{-1} < n$$

[1] Proofs for all combinations of economic conditions are presented in Pekarsky (2012) Chap. 7, which also includes discussion of the implication of alternative choices of decision thresholds for the economic loss associated with decision to adopt a new drug at a range of potential thresholds.

The threshold value falling below $n$ under these conditions reflects the opportunity to improve displacement as well as undertake the highest value adoption action.

Within a budget-constrained health system, this points to the need to address inefficient displacement of services (Pekarsky 2012) as well as invest in the most cost-effective expansion of existing services (Birch and Gafni 1992), which the health shadow price allows for simultaneously.

As a result, $\beta c$ creates appropriate incentives for allocative and displacement efficiency in expanding the use of the most cost-effective services and technologies and contracting the least cost-effective services and technologies. This in turn provides a robust pathway towards allocative and displacement efficiency with appropriate incentives as well as providing appropriate value to collecting evidence for $n$, $m$ and $d$ in order to avoiding adoption and displacement inefficiency.

The health shadow price also makes clear that using $d$ as a threshold value generally denies the true opportunity cost of reimbursement (adopting and funding) actions with new health technologies – the most cost-effective expansion of existing health system interventions as well as displacement of least cost-effective interventions required to create a pathway to allocative and displacement efficiency. Hence, the ICER for displaced services coinciding with that for the most cost efficient expansion and contraction of existing services ($d = n = m$) at the single point of complete allocative and displacement efficiency ($n = m$ and $d = m$) should not be confused with a threshold of $d$ providing a pathway to reach allocative efficiency (Eckermann and Pekarsky 2014: 322).

The health shadow price $\beta c$ in reflecting the budget constrained opportunity cost or highest value alternative joint adoption and financing actions does provide this pathway, and represents an ICER less than $d$ with any form of allocative inefficiency ($m > n$). This arises given that in the case of allocative inefficiency either:

(i) The health shadow price of reimbursement reflecting opportunity cost of joint adoption and financing actions is less than $n$ when there is displacement inefficiency ($d < m$), and hence the health shadow price must be less than $d$ even in the extreme case where displacement was so inefficient that $d = n$ (i.e. the most cost-effective programmes to expand were contracted or displaced), then

$$\beta c = \left(\frac{1}{n} + \frac{1}{d} - \frac{1}{m}\right)^{-1} < n \leq d < m; \text{or}$$

(ii) If displacement is efficient ($d = m$), then the health shadow price equals the ICER of the most cost-effective expansion ($n$) which implies that with allocative inefficiency ($n < m$) that the health shadow price will be less than that of the least cost-effective service displaced, then $\beta c = \left(\frac{1}{n} + \frac{1}{d} - \frac{1}{m}\right)^{-1} = n < d = m$.

Hence with allocative and displacement inefficiency characteristic of health systems, $d$ is always greater than the health shadow price $\beta c$, or equivalently opportunity costs of reimbursement (adoption and financing actions) are always greater than that suggested by $d$ and do not provide a pathway to allocative efficiency.

More generally there are distinct losses from using displaced services related thresholds rather than the health shadow price in reimbursing new technology with expected net costs under budget constraints in relation to allocative efficiency arising from:

(i) Not undertaking the most cost-effective expansion of existing technology – reflecting the true opportunity cost of adoption;
(ii) A process of accepting new interventions that either worsens or entrenches allocative inefficiency where pricing is up to a threshold of least cost-effective services assumed displaced or actual services displaced; and
(iii) Risking health losses from financing processes either directly where not consistently applied across patented and non-patented technologies or over time where costs of reversal are faced in cycling through new technologies where consistently applied.

The health shadow price in contrast to the ICER of displaced services provides a direct pathway to allocative and displacement efficiency across new or existing technology in reimbursing new technology in comparison with adoption of most cost-effective expansion of current programmes financed by contraction of least cost-effective programmes. This allows benefits from optimising the use of existing and new technologies to flow through to maximising health gains from any given budget. It also prevents dangers arising over time of new technologies priced up to the threshold being cycled through and costs of reversal faced. New technologies with net expected additional costs reimbursed against the health shadow price of Pekarsky (2012, 2015) are compared and priced relative to the most cost-effective expansion of current technology where displacement or contraction is efficient rather than least cost-effective or actual services displaced to finance adoption and hence are not expected to be cycled through in the next cycle or over time at any point with consistent application of this threshold.

More generally the health shadow price $\beta c$ points to the need for and value of research on the most cost-effective expansion of current services ($n$) as well as the least cost-effective contraction of services ($m$) with existing technology alongside evidence of what is actually displaced ($d$). Given allocative and displacement efficiency characteristic of health systems, there are large costs to not addressing market failure in relation to cost-effectiveness evidence for unpatented services and technologies and existing technologies and services to expand and contract. Without such evidence the key opportunities to optimally expand and contract are missed, and new technologies are not appropriately evaluated or priced relative to opportunity cost of their reimbursement. Associated losses arise from allocative and displacement inefficiency and suboptimal adoption and displacement actions in reimbursement decisions but also appropriate threshold values and incentives in research and regulation processes are not created.

The cost faced by societal decision makers relative to optimal actions in informing and using the health shadow price across reimbursement, research and regulatory decisions points to the multifaceted value of undertaking research on best expansion and contraction of existing programmes and technologies. Indeed, they point to an imperative to do so if health systems wish to move towards allocative and displacement efficiency.

### *11.6.1 The Health Shadow Price and PBMA as a Pathway to Allocative Efficiency*

What historically allows such evidence of best expansion and contraction and far better aligns with the health shadow price as a pathway to allocative and displacement efficiency is programme budgeting and marginal analysis (PBMA) (Ruta et al. 1996, 2005). PBMA is a process that considers the marginal gains and losses from expanding and contracting programme budgets in service use. This provides a pathway to technical and allocative efficiency with existing services by expanding services with greatest marginal gains and contracting those with lowest marginal losses until, under diminishing marginal returns, no further gains can be made. The health shadow price $\beta c$ aligns with such PBMA principles in creating active incentives for best expansion and contraction of existing services to move towards allocative efficiency, but additionally allows for new as well as existing technologies and for their appropriate pricing, adoption and reimbursement (Eckermann and Pekarsky 2014: 323).

Methods to aid PBMA processes and the health shadow price integrate existing evidence with new technologies point to value of information methods or more generally appropriate evaluation methods to aid efficient evaluation of cost-effectiveness of existing technologies. Globally optimal societal decision-maker trials designed using VOI methods presented in Chaps. 6 and 7 following Eckermann and Willan (2009, 2013) ensure the ability to feasibly adopt and trial in robustly evaluating the cost-effectiveness of existing as well as new technologies and optimize global value relative to cost of such trials. That is, where trials are undertaken in jurisdictions yet to adopt such technologies and translate evidence to jurisdictions who have adopted in evaluating whether to expand, contract or keep services as they are, with trial funding shared according to expected benefits across jurisdictions. Such global trial-based approaches are likely to be particularly valuable for non-patentable programmes and services with individual-based interventions such as rehabilitative care services.

Robust evidence for community-based health promotion and disease prevention programmes would be best served by applying multiplier or network evaluation methods highlighted in Chap. 4 following the research of Hawe and Shiell (Hawe and Shiell 2000; Hawe et al.2009; Shiell and Hawe 1995; Shiell et al. 2008) as illustrated with evaluation of the Stephanie Alexander Kitchen Garden program in school communities (Eckermann et al. 2014). Such multiplier and network methods quantify community effects, local ownership and sustainability over time. In triangulation with qualitative evidence of programme acceptance and short-term individual attitude or behavioural impact, they allow to identify whether community programmes with potential for costless expansion of community effects over time are expected in the long term to be successful, sustainable and cost-effective (De Salazar et al. 2007; Eckermann et al. 2014). Chapter 4 also highlighted the need for multiple domain comparison to enable appropriate coverage and comparability in evaluation of palliative care, which were addressed with robust multiple domain evaluation applying the net benefit correspondence theorem (NBCT) in Chap. 10 (McCaffrey 2013; McCaffrey et al. 2015). These enable key palliative care process of death domains not able to be integrated with survival such as finalizing personal and financial affairs, family and

carer distress and place of palliation and death to be appropriately incorporated into evaluation.

Such methods facilitate informing the health shadow price and PBMA processes where, as Mooney et al. (2008) strongly argued, objectives other than health – i.e. equity, process utility, etc. should be included to reflect societal values and objectives. More generally the NBCT methods for robust comparison of multiple providers as well as multiple strategies or outcomes in practice highlighted in Chaps. 8, 9 and 10 also provide more direct approaches to identify existing services and programmes for expansion and contraction.

## 11.7 Health Shadow Prices for Cost Saving Investment Options

Thus far consideration of the health shadow price and threshold values has been for new technologies, strategies or programmes in the usual case of interest where new technologies are expected to have higher incremental effects and incremental costs and hence lie on the NE quadrant. We now turn to the less considered SW quadrant where new technologies, programmes or strategies save costs while trading off potentially lower effects.

The flipside of efficient displacement ($d = m$) on the NE quadrant is considering investment in options on the SW quadrant that are cost saving (and hence generate funds for financing) relative to current practice at a threshold value that minimises the health loss arising from generating cost savings for financing (Eckermann 2015). We consider the health shadow price on the SW quadrant first for the simple case where the budget can be contracted following Eckermann (2015) and then derive the shadow price for a strictly fixed budget where funds generated are required to be used in financing further adoption. In both cases we consider implications for a kink between the NE and SW quadrant and illustrate empirically what that looks like with UK data.

Eckermann (2015) shows that the health shadow price and threshold value appropriate to the SW quadrant on the CE plane consequently differs from that on the NE quadrant both:

(i) Qualitatively, as a threshold value that strategies need to be greater than on the SW quadrant (compared with less than on the NE quadrant) in being acceptable with optimal decision making; and
(ii) Quantitatively in the presence of allocative inefficiency, with the SW quadrant health shadow price and appropriate threshold value where the budget can be freely contracted relating to the least health-reducing way of generating financing (ICER = $m$).

Consequently, in comparison with $\beta c$ for new technologies with greater expected net cost on the NE quadrant, the appropriate threshold value on the SW quadrant is greater up until the point of complete allocative ($n = m$) and displacement ($d = m$) efficiency. They coincide strictly only at the point of complete allocative and dis-

placement efficiency, i.e. only where $n = m = d$ does $\beta c = n = m$, and then only because they happen to coincide at that one point.

However, even at the hypothetical (and highly counterfactual relative to efficiency empirical evidence) point of complete allocative ($n = m$) and displacement ($d = m$) efficiency, the acceptance region for new technologies on the NE quadrant (positive incremental cost and effects) should be an ICER less than or equal to $n$ (less than if to improve fixed budget outcomes) while greater than or equal to $m$ on the SW quadrant. That is, the threshold is generally qualitatively different in making optimal decisions on the SW from the NE quadrant as the ICER needs to be greater than a threshold value to enable expected cost savings to have greater value than any expected losses in incremental effect, as considered in Willan and Briggs (2006).

The quantitatively greater threshold value in the SW quadrant ($m$) relative to NE quadrant $\beta$c for all points other than perfect allocate and displacement efficiency produces a kink in the appropriate health shadow price and hence threshold line about the origin. The extent of the kink reflects the degree of a health system allocative and displacement inefficiency. In the case of the UK, evidence of health system inefficiency from PBMA processes summarized in Claxton et al. (2013) points to current NICE decision-maker thresholds of £20,000–£30,000 per QALY (Devlin and Parkin 2004; NICE 2008) being orders of magnitude too high for the NE quadrant in identifying best expansion in optimising adoption and orders of magnitude too low in the SW quadrant in identifying best contraction in optimising displacement. Indeed, using the same PBMA evidence that was employed by Claxton et al. (2013) to estimate the average ICER for displaced services in the UK as £12,976 per QALY, the most cost-effective expansion across 23 MDCs estimates $n$ at £2000 per QALY in expanding services for respiratory problems. However, m, the ICER of the least cost-effective programme for contraction, is £2.73 million per QALY for neonates. Hence the health shadow price for more expensive new technology relative to best expansion and contraction in financing of existing programmes is $\beta c$ = £1734/QALY

$$\left( \beta c = \left( \frac{1}{n} + \frac{1}{d} - \frac{1}{m} \right)^{-1} = \left( \frac{1}{2000} + \frac{1}{12,976} - \frac{1}{2,730,000} \right)^{-1} = 1734 \right)$$

while that for cost saving approaches which may reduce health outcomes on the SW quadrant is £2.73 million per QALY. This UK evidence leads to a kink in the threshold value under relevant conditions shown on the CE plane in Fig. 11.1.

The kink represents the higher threshold value in the SW quadrant than the NE quadrant on the CE plane in turn reflecting allocative and displacement inefficiency and differences in the ICER of best expansion available (NE quadrant) relative to best contraction available (SW quadrant) in joint reimbursement processes of adoption and financing. Indeed a much higher threshold in the case of the UK (at £2.73 million vs. £1734 per QALY), indicating that there is very considerable allocative, adoption and displacement inefficiency in practice currently.

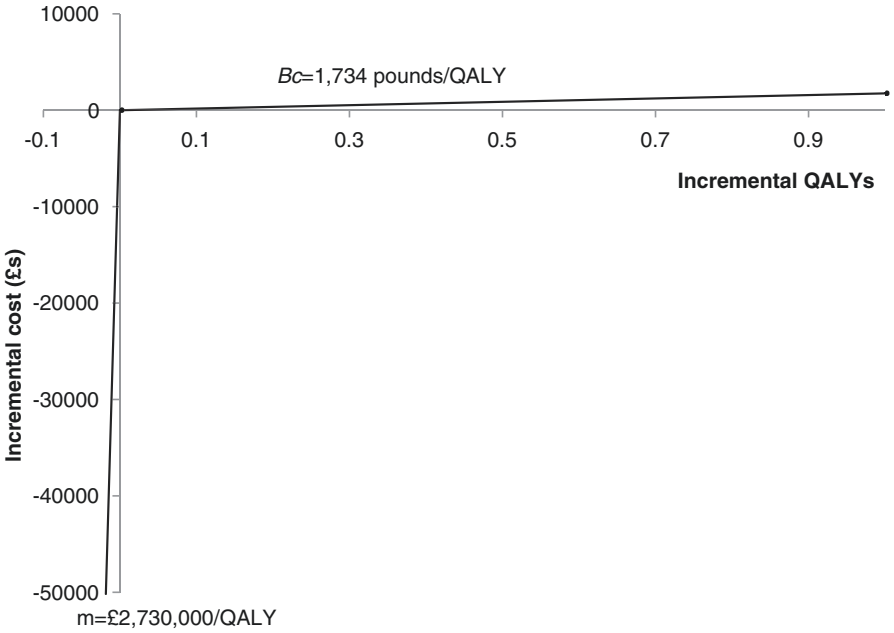


**Fig. 11.1** UK kinked threshold for health shadow prices on the CE plane with *n* £2000/QALY, *m* £2,730,000/QALY and *d* £12,976/QALY (Adapted from Eckermann (2015))

Note that, strictly, the above analysis assumes that in generating funds from adopting a new technology or option in the south-west quadrant or more generally where additional funds are generated in adoption (i.e. also the south-east quadrant where an option dominates), there is not a fixed or binding budget constraint – those funds do not have to be spent and the health budget can freely contract. Under those conditions, which reflect a fixed budget for expansion while the ability to contract without a budget constraint to spend money saved in contraction on adoption, the opportunity cost only needs to consider the best alternative raising of funds, and hence the health shadow price is m. This is analogous to the health shadow price on the NE quadrant if the budget was able to be expanded rather than being fixed, in which case displacement to finance would not be required to undertake reimbursement, and the opportunity cost purely relates to the best alternative adoption action, and hence $\beta c = n$.

Note that a budget strictly fixed in relation to not being able to be expanded with adoption on the NE quadrant does not necessarily imply the budget cannot be contracted without associated adoption on the SW quadrant. However, it is a moot point whether for any given health system budget, whether contraction is free from associated adoption or if in practice the budget is strictly fixed and cost savings are required to be spent on adoption.

If there were symmetry in the budget being strictly fixed with contraction on the SW quadrant as well as expansion on the NE quadrant, then funds raised from a cost

saving intervention on the SW quadrant have to be spent, and both funding and adoption actions should be considered. This is the analogous case to adoption and financing actions needing to be considered on the NE quadrant where a strict budget constraint with a fixed budget applies given the need to finance adoption decisions with additional costs.

As we shall now see for cases on the SW quadrant where additional funds raised need to be spent on adoption (the budget is strictly fixed in relation to contraction as well as expansion) then the shadow price is $m$ only if adoption is efficient, that is, funding made available (F) by contraction is spent in adoption on best expansion (ICER = $n$). More generally, with a strictly fixed budget, the health shadow price (Bf) on the SW or SE quadrant in generating financing should be derived by equating the expected return from actual adoption (ICER = $a$) from the funds generated by contraction F, with that from best displacement in generating funds F and those funds used for best adoption (ICER =$n$), that is

$$\frac{F}{a} - \frac{F}{\beta f} = \frac{F}{n} - \frac{F}{m}$$

$$\Rightarrow \frac{F}{\beta f} = \frac{F}{a} + \frac{F}{m} - \frac{F}{n}$$

$$\beta f = \left( \frac{1}{a} + \frac{1}{m} - \frac{1}{n} \right)^{-1}$$

Where adoption is efficient ($a = n$), this simplifies to $\beta f = m$. For example in the UK if there were a strictly fixed budget in contraction as well as expansion then provided adoption was efficient and hence the most cost effective expansion was employed (a = n = £2000 /QALY), then the SW quadrant health shadow price would be £2.73 million per QALY given the estimate for $m$ in the UK. However, where adoption is inefficient ($a > n$), then with a strictly fixed budget adoption actions can be improved alongside needing to compare with the best fund-generating action for financing adoption.

In the case of the UK with a NICE decision threshold of £20,000 to £30,000 per QALY or the proposed £12,976 per QALY where adoption decisions are based on actual displaced services (Claxton et al. 2013), then assuming pricing up to the threshold in adoption and with $n$ = £2000 per QALY, adoption inefficiency would need to be allowed for. Hence if there were a strictly fixed budget with funds generated needing to be spent, then Bf becomes

$$\beta f = \left( \frac{1}{a} + \frac{1}{m} - \frac{1}{n} \right)^{-1} = \left( \frac{1}{a} + \frac{1}{2,730,000} - \frac{1}{2000} \right)^{-1}$$

Regardless of whether values of the ICER for adoption ($a$) are £12,976, £20,000 or £30,000 per QALY, the shadow price is negative. This indicates that with a strictly

fixed budget (where funds raised have to be spent on adoption) the degree of inefficiency in adoption is such that a cost saving therapy would have to dominate an existing therapy to allow a better outcome than generating equivalent funding F with contraction of least cost-effective services (ICER = $m$) to fund best expansion (ICER = $n$). Indeed the shadow prices indicate that even in the case of a = £12,976 per QALY, the inefficiency of having to adopt at this threshold rather than undertake best expansion of current programmes (n = £2,000 per QALY) with a strictly fixed budget is so great that the cost saving technology would also need to have an improvement in effects to provide a better option (and a larger improvement if greater than this, indeed very large by £30,000 per QALY). More importantly, the threshold value for the ICER of adoption beyond which cost saving technologies would need to dominate can be calculated to occur at the point where Bf becomes undefined and hence where:

$$a = \left(\frac{1}{n} - \frac{1}{m}\right)^{-1}$$

In the UK case, the threshold value beyond which a cost saving technologies would need to dominate is $\left(\frac{1}{2000} - \frac{1}{2,730,000}\right)^{-1}$ = £2001.47 per QALY.

Interpreting this threshold, the ICER for best contraction of exiting services (m = £2.73 million per QALY) is so high that for almost any inefficiency in adoption the cost saving technology would have to dominate to be preferred to the best alternative optimal action of contraction of least cost-effective programmes (ICER = $m$) to generate funds F to fund optimal expansion (ICER = $n$).

In general shadow prices on the NE and SW quadrants both reinforce the need to identify best options for expansion (adoption) and contraction (displacement) of budgets, and in particular including research low cost or factor price alternatives in the case of new, non-patentable or non-patented existing technologies, and does not preclude them from being more effective while also cost saving (dominating) in comparison to current practice in any given jurisdiction. For example, in considering where policy reform should be heading for health and aged care policies with the challenge posed by baby boomer populations, Chap. 12 highlights highly effective prevention and factor price interventions for meeting community needs with dementia, chronic pain, aged care and palliative care. These options are considered as part of more generally allowing for appropriate incorporation of new and existing technologies in addressing challenges of successful ageing with constrained health and aged care budgets in countries such as Australia and include:

(i) Policies for dementia- and aged-friendly cities and communities in allowing successful ageing while preventing need for expensive aged care (Kalache 2013; Phillipson et al. 2016);
(ii) Better designing age care facilities for dementia patient needs, community and meaning while reducing the use of pharmaceutical constraints as part of

increasing their physical and mental wellbeing (Fleming and Purandere 2010; Zeisel et al. 2003); and

(iii) Alternatives in meeting palliative patient primary preferences for being able to finalise their personal and financial affairs in their community of choice (usually at home) while minimising family and carer distress (McCaffrey et al. 2015, 2016).

In relation to (iii) note that the key need for alternative pain therapies in palliative populations with clear dangers of current therapies such as ketamine and opioids (Hardy et al. 2012). Promising options in such palliative populations include optimised use of terpene, CBD and THC rich medicinal cannabis therapies to better meet patient pain relief needs without adverse side effects (Bachuber et al. 2014; Bradford and Bradford 2016; Johnson et al. 2010; Gallily et al. 2015) and enable finalising affairs while staying in their community setting of choice, as considered in detail in Sect. 12.5.

In each case the preventative programmes and factor price options with existing approaches and technology considered in Chap. 12 are suggested to dominate (be less expensive while more effective than) current services and point to allocative inefficiency in historical approaches for integrating new and existing technology in service provision which the health shadow price approaches of Pekarsky (2012, 2015) correct for. Pekarsky (2012, 2015) analogously highlights such potential for factor pricing rather than value-based pricing with the shadow price in considering dung beetles as a low cost preventative alternative to the use of chemicals or fly-screens for controlling fly populations in Canberra.

## 11.8 Conclusion

The health shadow price $\beta c = \left(1/n + 1/d - 1/m\right)^{-1}$ of Pekarsky (2012, 2015) provides an appropriate threshold value below which the ICER of new technologies need to be optimal for the usual case of interest with a fixed budget and new technologies that cost more (NE quadrant) and hence require displacement to finance. In particular it provides a pathway to allocative efficiency in adoption and displacement decisions. That is, it appropriately allows for opportunity costs of adoption and financing actions in related reimbursement decisions and price negotiation processes between societal decision makers and manufacturers. Where new technologies are cost saving while potentially less effective (SW quadrant), the health shadow price above which the ICER needs to be optimal has been shown to be $m$ for a budget that can contract, while price $\beta f = \left(1/a + 1/m - 1/n\right)^{-1}$) where the budget is strictly fixed in relation to contraction as well as expansion and funding generated is spent on adoption (ICER = $a$).

In general health shadow prices in the NE and SW quadrants provide appropriate coverage of the scope of appropriate alternative options considered for adoption and displacement and point to research required in relation to best expansion and contraction of existing programmes and a pathway for optimal decision making.

These considerations in turn highlight the importance of recognising the political economy and strategic behaviour in optimising underlying objectives for societal decision making consistent with community values. New technology manufacturers (e.g. pharmaceutical companies) do not have a monopoly on health improvement. Communities or societal decision makers charged with representing community interests in health systems can invest in or buy health improvements in many other ways. These can include but are by no means restricted to expanding the use of the most cost-effective current programmes or strategies, better implementation of current health-care strategies and technologies or undertaking research on promising non-patented or patentable programmes or strategies whether in prevention, health-care treatment or palliative services (see Chap. 12). Hence, scarce public funding for new health-care investment, research or initiatives in new technologies in optimising budget-constrained outcomes should be compared with the best alternatives for investing in current technologies, service implementation or related research when determining threshold values for effects, negotiating prices and making decisions. Decision theory should be at the basis of optimising decision making but needs to allow for best alternatives (opportunity cost) in satisfying coverage conditions and needs to be game theoretic in informing optimal decision making where outcomes depend on acting strategically (e.g. negotiating prices for new technologies). Chapter 12 considers some of those options in current policy analysis and reforms facing the challenges of an ageing baby boomer population in Australia and internationally.

## References

Arrow KJ. Uncertainty and the welfare economics of medical care. Am Econ Rev. 1963;LIII(5):941–73.

Bachhuber M et al. Medical cannabis laws and opioid analgesic overdose mortality in the United States, 1999-2010. JAMA Intern Med. 2014;174(10):1668–73.

Barrett A, Riques T, Small M, Smith R. How much will Herceptin really cost? Br Med J. 2006;333:1118–20.

Birch S, Gafni A. Cost effectiveness/utility analyses: Do current decision rules lead us to where we want to be? J Health Econ. 1992;11:279–96.

Birch S, Gafni A. Changing the problem to fit the solution: Johannesson and Weinstein's (mis) application of economics to real world problems. J Health Econ. 1993;12(4):469–76.

Bradford AC, Bradford WD. Medical marijuana laws reduce prescription medication use in medicare part D. Health Aff (Millwood). 2016;35:1230–6. doi:10.1377/hlthaff.2015. 1661. pmid: 27385238

Claxton K, Buxton M, Culyer A, Walker S, Sculpher M. Value based pricing for NHS drugs: an opportunity not to be missed? BMJ. 2008;336:251–4.

Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, Devlin N, Smith P, Sculpher M. Methods for the estimation of the NICE cost effectiveness threshold. Final Report. York: Working Paper 82 Centre for Health Economics, University of York; 2013.

Culyer AJ, McCabe C, Briggs A, Claxton K, Buxton M, Akehurst R, Sculpher M, Brazier J. Searching for a threshold not setting one: the role of the national Institute for health and Clinical Excellence. J Health Ser Res Policy. 2007;12(1):56–8.

de Salazar L, Jackson S, Shiell A, Rice M. Guide to economic evaluation in health promotion. Washington: Pan American Health Organization, Regional Office of the WHO; 2007.

Devlin N, Parkin D. Does NICE have a cost effectiveness threshold and what other factors influence its decisions? A binary choice analysis. Health Econ. 2004;13(5):437–52.
Eckermann S. Hospital performance including quality: creating economic incentives consistent with evidence-based medicine. Ph.D. Dissertation, University of New South Wales, Sydney; 2004. http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/.
Eckermann S. Kinky thresholds revisited: opportunity costs differ in the NE and SW quadrants. App Health Econ Health Policy. 2015;13:7–13. doi:10.1007/s40258-014-0136-3.
Eckermann S, Coelli T. Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice. Soc Sci Med. 2013;76(1):159–68.
Eckermann S, Pekarsky B. Can the real opportunity cost stand up: displaced services the straw man outside the room. Pharmaco Economics. 2014;32(4):319–25.
Eckermann S, Briggs A, Willan A. Health technology assessment in the cost-disutility plane. Med Decis Mak. 2008;28:172–81.Eckermann S, Karnon J, Willan A. The value of value of information: best informing research design and prioritization using current methods. Pharmacoeconomics. 2010;28(9):699–709.
Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16:195–209.
Eckermann S, Willan AR. The option value of delay in health technology assessment. Med Decis Making. 2008a;28:300–5.
Eckermann S, Willan AR. Time and EVSI wait for no patient. Value Health. 2008b;11:522–6.
Eckermann S, Willan AR. Globally optimal trial design for local decision making. Health Econ. 2009;18:203–16.
Eckermann S, Willan A. Optimal global VOI trials: better aligning manufacturer and decision maker interest and enabling feasible risk sharing. PharmacoEconomics. 2013;31:393–401.
Eckermann S, Willan A. Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies. Pharmaco Econ. 2011;29(7):563–77.
Eckermann S, Dawber J, Yeatman H, Quinsey K, Morris D. Evaluating return on investment in a school based health promotion and prevention program: the investment multiplier for the Stephanie Alexander Kitchen Garden National Program. Soc Sci Med. 2014;114:103–12.
Fleming R, Purandere N. Long-term care for people with dementia: environmental design guidelines. Int Psychogeriatr. 2010;22(7):1084–96.
Gafni A, Birch S. Incremental cost-effectiveness ratios (ICERs): the silence of the lambda. Soc Sci Med. 2006;62(9):2091–100. Epub 2005 Dec 1
Gallily R, Yekhtin Z, Ondřej HL. Overcoming the bell-shaped dose-response of cannabidiol by using cannabis extract enriched in cannabidiol. Pharmacol Pharm. 2015;6:75–85.
Graham DA. Cost-benefit analysis under uncertainty. Am Econ Rev. 1981;71:715–25.
Graham D. Public expenditure under uncertainty: the net-benefit criteria. Am Econ Rev. 1992;82:822–46.
Griffin S, Claxton K, Sculpher M. Decision analysis for resource allocation in health care. J Health Ser Res Policy. 2008;13(Suppl 3):23–30.
Hardy J, Quinn S, Fazekas B, Plummer J, Eckermann S, Agar M, Spruyt O, Rowett D, Currow D. 2012. A randomised, double-blind placebo controlled study to assess the efficacy and toxicity of subcutaneous ketamine in the management of cancer pain. J Clin Oncol. 2012;30(29):3611–7.
Hawe P, Shiell A. Social capital and health promotion: a review. Soc Sci Med. 2000;51(6):871e885.
Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. Am J Community Psychol. 2009;43(3e4):267–76.
Johannesson M, Weinstein MC. On the decision rules of cost effectiveness analysis. J Health Econ. 1993;12:459–67.
Johnson JR, Burnell-Nugent M, Lossignol D, Ganae-Motan ED, Potts R, Fallon MT. Multicenter, double-blind, randomized, placebo-controlled, parallel-group study of the efficacy, safety, and

tolerability of THC:CBD extract and THC extract in patients with intractable cancer-related pain. J Pain Symptom Manag. 2010;39(2):167–79. doi:10.1016/j.jpainsymman.2009.06.008. Epub 2009 Nov 5

Kalache A. The longevity revolution: creating a society for all ages. Adelaide thinker in residence 2012–2013. Government of South Australia: Adelaide; 2013. Available at: http://www.flinders.edu.au/sabs/fcas-files/Publications/The%20Longevity%20Revolution.pdf

McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: What it is and what that means. Pharmaco Economics. 2008;26(9):733–44.

McCaffrey N. Modelling joint cost and outcomes uncertainty on the cost-disutility plane - case studies in palliative care. PhD, Flinders University, Adelaide; 2013.

McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Better informing decision making with multiple outcomes cost-effectiveness analysis under uncertainty in cost-disutility space. PLoS One. 2015;10(3):e0115544.

Mooney G, Coast J, Jan S, McIntyre D, Ryan M, Wiseman V. Not so nice. J Health Ser Res Policy. 2008;12(3):190.

National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. London: NICE; 2008.

Pekarsky B. Trusts, constraints and the counterfactual: reframing the political economy of new drugs. Adelaide: PhD Thesis University of Adelaide; 2012.

Pekarsky BAK. The new drug reimbursement game: a regulator's guide to playing and winning. London: Springer; 2015.

Pharmaceutical Benefits Advisory Committee. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.4). Canberra: Department of Health, Commonwealth of Australia; 2013.

Phillipson L, Hall, D, Cridland L, Fleming R, Brennan-Horley C. Dementia-Friendly Kiama. Final Evaluation Report. October, 2016. University of Wollongong. 2016.

Ruta D, Donaldson C, Gilray L. Economics, public health and health care purchasing: the Tayside experience of programme budgeting and marginal analysis. J Health Serv Res Policy. 1996;1:185–93.

Ruta D, Mitton C, Bate A, Donaldosn C. Programme Budgeting and marginal analysis: bridging the divide between doctors and managers. BMJ. 2005;330(25):1501–3.

Sculpher M, Claxton K. Real economics needs to reflect real decisions. Pharmaco Econ. 2012;30(2):133–6.

Shiell A, Hawe P. Health promotion community development and the tyranny of individualism. Health Econ. 1995;5(3):241e247.

Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. Br Med J. 2008;336(7656):1281e1283.

Trueman P, Drummond M, Hutton J. Developing guidance for budget · impact analysis. Pharmacoeconomics 2001;19:609–21.

Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. N Engl J Med. 1977;296:716–21.

Weinstein M, Zeckerhauser R. Critical ratios and efficient allocation. J Public Econ. 1973;2:147–57.

Willan AR, Eckermann S. Optimal clinical trial design using value of information methods with imperfect implementation. Health Econ. 2010;19(5):549–61.

Willan AR, Eckermann S. Expected value of information and pricing new health care interventions. PharmacoEconomics. 2012;30(6):447–59.

Willan A, Briggs A. The statistical analysis of cost-effectiveness data. Chichester: Wiley and Sons; 2006.

Zeisel J, Silverston N, Hyde J, Levkoff S, Lawton MP, Holmes W. Environmental correlates to behavioural health outcomes in Alzheimer's special care units. Gerontologist. 2003;43(5):697–711. doi:10.1093/geront/43.5.697.

# Chapter 12
# Policy Implications and Applications Across Health and Aged Care Reform with Baby Boomer Ageing - from Age and Dementia Friendly Communities to Palliative Care

## 12.1 Introduction

This chapter highlights policy application of health economic principles and methods introduced and developed in this book to current health and aged care issues and areas of policy reform challenges internationally, illustrated predominantly with Australian policy-related examples. In particular reform challenges posed for successful ageing of the baby boomer population within increasingly constrained budgets and consequent imperatives for being more efficient in use of existing technology and pricing and integration of new technology and related research, reimbursement and regulatory decisions.

In meeting health and aged care system reform challenges of successful ageing within budget constraints, the clear empirical and theoretical advantages of universal access public systems are initially considered, before nevertheless turning to the need for related reform in the face of an ageing baby boomer cohort for:

(i) Age- and dementia-friendly community policies to promote active ageing while delaying or preventing the need for aged care and nursing home facilities;
(ii) Dementia-friendly aged care and nursing home design;
(iii) Inexpensive palliative care options that reflect palliative care primary preferences for having the ability to finalise affairs in their community of choice while minimising family and carer distress; and
(iv) Funding mechanisms that provide active incentives for budget-constrained health and aged care system optimisation of quality of care rather than for minimum cost per service quality of care, cost shifting and cream skimming, with current case-mix funding methods.

Successful ageing without breaking the budget requires such policy reform is undertaken and involves research and practice comparisons assessing best use of existing technologies at factor prices and better integration and more appropriate pricing across existing and new technology.



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_12

## 12.2 Health-Care Policy for Successful Ageing: Where Should Health and Aged Care Reform Be Heading (The Importance of Dementia and Age-Friendly Community Environments)

In considering successful ageing, Alexandre Kalache (previous director of the WHO global programme on ageing and current President International longevity Centre Brazil (ICLB 2015 & Global) highlights that the baby boomer generation (born 1946–1964), who established adolescence for teenagers in the 1960s, want and will increasingly demand ageing to be a 'gerentolescence' (Kalache 2013). That is, where baby boomers as they age remain actively engaged socially and physically in their communities and at work and play.

Key areas or pillars of capital investment for successful ageing and a gerentolescence are highlighted by Kalache (2013) as:

(i) Education capital (lifelong learning);
(ii) Health capital;
(iii) Participation or social capital (community/friends/family); and
(iv) Security – financial capital and/or resilience.

To this end the WHO under the direction of Alexandre Kalache and Louise Plouffe developed the Global Age-Friendly Cities project, a global network of what in 2016 is 1800 'age-friendly' cities or communities – committed to creating inclusive and accessible urban environments to benefit population ageing and an age-friendly city guide (WHO 2007). As the guide indicates WHO (2007) 'An age-friendly city adapts its structures and services to be accessible to and inclusive of older people with varying needs and capacities'.

The guide offers policy and practical advice and checklists for whole of government and community support of age-friendly communities from top-down, but also bottom-up, perspectives gathered from aged populations in 33 pilot cities, across eight domains of urban living: outdoor spaces and buildings, transportation, housing, social participation, respect and social inclusion, civic participation and employment, communication and information and community support and health services.

Kalache (2013) notes that given we all age, such social capital investment in successful ageing of the baby boomer generation also benefits all those to follow in providing a pathway to successful ageing. Additionally, this strengthens communities and their voice and preferences in decision making, respecting of rights and addressing discrimination. However, such investment in successful ageing is naturally budget constrained and needs to consider related research, reimbursement and regulatory decisions in enabling optimal outcomes within budget constraints.

### *12.2.1 How Can Health Economics Help: More Than Cost-Effectiveness Analysis*

Health economics principles and methods developed in this book enable optimisation of joint-related research, reimbursement and regulatory decisions across health and care systems (e.g. aged care) in addressing community net benefit. The political economy is also a key consideration in making sure community preferences rather than vested interests are represented as the underlying community net benefit objective in such decisions and that associated appropriate incentives are created for agents – providers and institutions.

Unlike 'market economics' health system decisions are made in the context of bounded rationality and characteristic market failure (Simons 1957; Arrow 1963) where:

(i) Asymmetry of information between patient and provider and bounded rationality combine to create perfect conditions for supplier-induced demand (overservicing); and
(ii) The limiting of access to necessary services and health care causes social system inefficiency as well as inequality – downstream health and cost of care impacts.

Consequently, health economics advice needs to take a system-wide 'big picture' perspective of health policy impacts, and the direction reform should take across joint research, reimbursement and regulatory decision making. System-wide policy impacts are most obvious in contrasting the health, equity, expenditure and efficiency outcomes of the US health-care system (without universal access) and other OECD countries like Australia with universal access and public provision of necessary care, in Australia's case via Medicare (Deeble 1999).

For example, consider Commonwealth Fund (Davis et al. 2014) comparisons across 11 comparable OECD countries including the USA and Australia. The USA despite spending significantly more on health than any other country, at US$8508 per capita, is ranked worst (11th/11) for each of equity, healthy lives or efficiency, as well as overall. Australia in comparison had health expenditure of $3800 per capita and is ranked fourth for healthy lives, efficiency and overall.

More generally across the whole OECD, the USA spends double the percentage of GDP on health as the other 33 OECD countries (18% vs. 9% – OECD 2013: 157) in 2011 yet had worse outcomes than the OECD average in terms of life expectancy at birth or life expectancy increases in over 65 populations over the previous 50 years (OECD 2013: 25, 173). Indeed, life expectancy and improvement over 50 years were lower in the USA than all countries with health expenditure greater than $2000 per capita (OECD 2013: 25). The bottom line then for reform across OECD countries is that they should not move towards a US style health system without universal access and associated equity, efficiency and health outcome problems.

While the empirical evidence of higher cost and worse health outcomes and hence inefficiency is clear, a natural theoretical question is why are health systems like the USA without universal access so inefficient? Two key-related problems arising for systems without universal access to primary care patient fees for such services are that:

(i) Those without appropriate access to necessary health care are underserviced and have expected worse health outcomes and downstream health and social system care needs; and
(ii) Those with access are overserviced by providers with active fee for service incentives to induce demand and also have worse health outcomes from over-testing and treatment (such as unnecessary surgery, polypharmacy and testing for rare diseases or those without population net clinical benefit from treatment) compared with appropriate care.

These expected downstream impacts in both underserviced and overserviced populations highlight decision-analytic principles of adequate coverage and comparability as key in considering the effects and cost of alternate policy pathways for robust policy analysis, just as they are for robust health technology assessment, research design and regulation in practice. Adequate coverage requires sufficient length of follow-up and scope of resource use and health outcomes to capture and translate evidence of downstream cost and effect impacts, such as impacts of delaying necessary care but also side effects and long terms impacts of over-treatment (e.g. polypharmacy with over-medicating), and overtesting given false positives and their treatment with over testing for rare diseases etc. Comparability requires considering expected relative impacts of policies on health effects/resource use (cost) across appropriate options or comparator/s. As Chaps. 1 and 2 in particular highlighted these coverage and comparability principles are key to robust consideration of expected costs, effects and net benefit (value effects less costs) of policy alternatives and before assessing whether further evidence is required in the presence of decision uncertainty, using VOI methods as per Chaps. 5–7 or otherwise.

For example, consider Australian Federal Government 2014 budget proposals for mandatory patient co-payments of $7 for primary care services (Parliment of Australia 2014), proposed as a measure to ‘support Medicare’ and undertake the ‘heavy lifting’ for the health system.

Evidence from the USA shows increasing primary care co-payments for elderly (Medicare over 65) populations by $7 (from $7 to $14) and in specialist care by $9.50 (from $12.50 to $22) lead to (Trivedi et al. 2010):

(i) 20 fewer outpatient visits per 100 population
(ii) 2.2 more hospital admissions and 13.4 greater inpatient days

While this lead to outpatient costs reducing on average by $71 per patient, inpatient costs increased by $240 per patient and total costs by $169 per patient. Consequently, for every dollar ‘saved’ in primary care, $3.38 was spent on downstream hospital services. This ratio of greater additional costs spent in downstream services than ‘saved’ in primary care increased for populations with chronic disease such as hypertension, diabetes and MI. Similarly in terms of overall impact of introducing primary care charges, Helms et al. (1978) in a RAND study found that the

$1 Medicaid co-payment in the 1970s in California lead to a net increase in overall Medicaid costs allowing for downstream alongside direct costs of 3–8%.

In Australia's case 2014 budget-proposals for a $7 mandatory patient co-payment of $7 to access GP primary care would be expected to lead to reduced access by low socioeconomic and reticent patient population. That is, those most cost effective to receive preventative services and be treated early in primary care (Eckermann 2014a, b). However, this impact would have been in large part hidden in overall GP service use by supplier inducement of demand of GPs in response to fill holes in their patient lists and incomes, in populations remaining able to afford services (Richardson and Peacock 2006; Peacock and Richardson 2007; van Dijk 2013). That is, shortfalls arising in GP lists and incomes from reduced use of lower socioeconomic and reticent populations would be expected to lead to inducement of demand and discretionary overservicing of those remaining (Eckermann 2014a, b; Eckermann et al. 2016; Eckermann and Seridan 2016).

Critically, both underserviced and overserviced populations would be expected to cost the health system more over time (McKay et al. 2014; Eckermann 2014b). Underserviced populations in undermining prevention services and delaying necessary care, leading to worse health outcomes and much more complex and expensive downstream treatment in settings such as hospitals, specialist care and aged care. Overservicing by GPs of those remaining able to afford access with higher costs of supplier induced demand directly and in dealing with downstream health impacts of unnecessary treatment such as that associated with over-testing and associated false positives and their treatment. Hence, far from doing the heavy lifting, such mandatory GP co-payments for patients undermine universal access and give the health system a hernia, creating US style health system problems of reduced access, higher cost, worse outcomes and unequivocally lower net benefit.

To efficiently and equitably address supplier-induced demand in primary care requires consideration of UK style capitation-based funding models where GPs are paid for the population they serve over time rather than fee for service (Al-Zaidy 2015). As in the UK this would also provide potential for adjusting payment for population-level outcomes in creating appropriate incentives for quality of primary care across the patient population. In this regard to create appropriate incentives for net benefit maximising quality improvement, the net benefit correspondence theorem could be used to create budget-constrained continuous incentives for population-level quality of care analogous to their use with hospital-based efficiency and funding mechanisms (see Chapt. 9 and Sect. 12.6). The use of such capitation funding mechanisms would also quickly rationalise current distributions of GPs and use of their services across Australian communities. Currently, the greatest concentration of working GPs is in urban communities (228/100,000 in 2011) and even higher in those at least need, such as the eastern suburbs of Sydney. Rural and remote populations with high primary care service needs in comparison have lowest access to working GPs; 145/100,000 in regional and 113/100,000 in remote Australia, respectively (ABS 2013a, b).

Universal access public health systems such as Medicare in Australia and the NHS in the UK have been shown theoretically and empirically to be relatively effective, efficient and equitable in relation to the alternative (Arrow 1963; OECD 2013; Davis et al. 2014). However, such health systems still face challenges and particularly with ageing of the baby boomer cohort. The baby boomer ageing revolution in Australia is reflected in the percentage of population over 65 projected to increase

from 14% in 2012 to 24% in 2045 and the life expectancy of Australians aged 65 in 2013–2015 being 19.5 years for males and 22.3 years for females (ABS 2016).

Given the first of the baby boomers (born 1946–1964) are currently 71 and the tail end of the baby boomers reach 65 in 2029, it is from 2011 to 2050 that we face rapidly increasing challenges of the baby boomer ageing revolution to enable successful ageing within resource constraints. In meeting this challenge in Australia, better solutions are required than recent reform suggestions for GP co-payments, the continual freezing of payments to GPs for bulk-billed services provided free of charge to patient (Britt et al. 2013; Harrison et al. 2015) or 'business as usual' in relation to pricing of new technologies or aged, dementia and palliative care. To see why, it is important to arm oneself with evidence of both what has driven health expenditure growth (where as we will see until recently ageing impacts on expenditure have been protected by increasing life expectancy), and what is expected if we don't reform as the baby boomer cohort ages – the myths and challenges.

### *12.2.2 Ageing Expenditure Catastrophe: Prior Myths and Future Challenges*

Common ageing claims are that ageing populations cost the health system more due to:

(i) Higher per-patient cost in older age groups
(ii) High cost of complex new technology used in treating aged populations

However, in relation to (i), health expenditure better relates to proximity to death than chronological age (Fuchs 1984) where, for example, Lubitz and Riley (1993) found within over 65 age groups sevenfold Medicare expenditure in the last year of life, and 2.3-fold in second last year, compared with those who survived 2 years. Hence, the first claim overstates impacts of ageing on health expenditure to the extent that where life expectancy increases the proportion of the population dying at any age or in any age groups generally falls, reducing projected health expenditure given the relationship of health expenditure to proximity to death.

That is, increasing life expectancy lowers the proportion dying and associated death related health costs in each age cohort. For example, the ageing impact on Australian health spending projections for over 65 populations from 1990 to 2020 (Goss et al. 1992) reduced by 30% allowing for health expenditure by proximity to death, given ABS expectations of increased life expectancy in 1992. In the UK expected increase in life expectancy halved annual ageing attributable growth from 0.8 to 0.4% in projections over the period 2002–2026 (Seshamani and Gray 2004).

However, in relation to the projections of health expenditure in Australia from 1990 to 2020, note that actual Australian mortality rates reduced (LE increased) much faster than projected. For example, between 1970 and 2002 mortality rates more than halved in all male and female age groups aged 44–77 (Productivity Commission 2006). Indeed, they fell so fast that despite the over 65 population

increasing from 8 to 13% over that period, the crude overall population death rate in Australia actually fell significantly between 1970 and 2002 from 9.0 to 6.7 per 1000. Allowing for actual mortality reduction (largely attributable to reduced smoking, improved sanitation, vaccination, safety and other preventative health measures), Productivity Commission (2006) estimates of ageing impacts on health expenditure fell from a 0.5–0.6% annual rate to a 0.18% annual rate from 1970 to 2002. The bottom line then is that ageing represented only about one twentieth of the 3–4% annual per capita real health expenditure growth.

### *12.2.3 What Has Driven Real Health Expenditure Growth Rather Than Ageing?*

The increase in Australian real (inflation adjusted) health expenditure and as a proportion of GDP from 6.4% in 1989–1990 to 9.7% in 2013–2014 (AIHW 2016) is not attributable to an ageing population, but rather related to:

(i) Increased use of expensive new technology across age groups, particularly medications such as those for pain, reflux, depression, cancer, CHD and AIDS; scans and pathology; and ICU, dialysis, CHD procedures and cardiac devices such as implantable defibrilators and drug eluting stents.

(ii) The private health insurance rebate while having grown to now more than A$6 billion annually, far from meeting a stated objective of taking pressure off public hospital budgets, has been shown to have had little or no impact on insurance rates via reducing the threat of lifetime cover (Ellis and Savage 2008). Indeed, these combined policies since 2000 have actually placed net pressure on the public system with wage pressure resulting from increased private sector activity for unnecessary care (Butler 2002; Ellis and Savage 2008; Harris 2013; Eckermann 2014a; Eckermann et al. 2016, PHIAC 2014).

(iii) Increasing non-government expenditure on privately funded and predominantly privately provided treatment, particularly after introduction of lifetime cover in 2000. Indeed, the rate of private funding increase mirrored increasing wealth (AIHW 2016), while predominantly related to non-necessary elective procedures, tests and treatment (unnecessary imaging tests, cosmetic procedures, unnecessary surgery – knee arthroscopy, back surgery, high caesarean rates, etc.) without health benefit or actually risking long-term health issues and treatment needs as in the USA (Colombo and Tapay 2004; Ellis and Savage 2008; Harris 2013; Eckermann 2014a, b).

While ageing only explained about 5% of increased Australian health expenditure from 1970 to 2002, the reduction in crude death rate (associated with increasing life expectancy) that has been protective of the ageing impact on health expenditure in Australia was projected to bottom out during the 2005–2015 (Productivity Commission 2006) period. Beyond 2015, with the baby boomers starting to enter their eighth decade, the crude mortality rate is expected to increase rapidly from 6.7% to 10% over the period 2015–2045. Any further life expectancy gains will be

swamped by the baby boomer populations ageing beyond eighty, where mortality rates are significantly higher in absolute terms and have not fallen.

To recap, decline in age-specific and indeed overall mortality rates with increasing life expectancy has decreased mortality-related costs and postponed such costs to older ages. This has been important in protecting health expenditure from ageing impacts but won't be protective as the baby boomers enter old age. Mortality and health expenditure associated with proximity to death will not be delayed forever and hence what has been protective of ageing impacts on health expenditure – falling absolute mortality rates with increasing life expectancy will be faced with the baby boomer cohort increasingly entering their eighth decade and beyond.

In Australia this ageing effect combined with new technology cost estimates lead to the Productivity Commission (2006) projection of a rapid rise in health expenditure from 8% of GDP in 2005 to 12% GDP by 2045, with half of that increase attributable to ageing. The extent to which such increases arise in practice depends on whether new technology costs and proximity to death costs either:

(i) Continue as currently as the Productivity Commission modelling assumes; or
(ii) Are addressed as part of the ageing revolution.

Productivity Commission projections assume continuing high end-of-life costs and high prices and costs of integrating new technology. However, costs of end-of-life care prices of new technology are endogenous, not exogenous to system decisions and policy choices, as highlighted in Chaps. 10 and 11, respectively.

Baby boomer ageing can be successful and not break the health budget if some current key health and aged care system inefficiencies are addressed with policy reforms such as those now considered in Sects. 12.3 to 12.6 for:

(i) Age- and dementia-friendly cities and communities to maintain active aged populations in the community and minimise need for aged care (Kalache 2013; WHO 2007; ADI 2012; Phillipson et al. 2016);
(ii) Dementia-friendly architecture for populations in aged care (Fleming and Purandere 2010; Zeisel et al. 2003);
(iii) Better providing for end-of-life and palliative population preferences and models of care in community (McCaffrey 2013; McCaffrey et al. 2014, 2015);
(iv) Better use of existing technologies at factor prices to improve budget-constrained outcomes directly but also via better pricing of new technologies relative to best alternative investment actions in support of the health shadow price (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014); and
(v) Improvement of quality of care in practice with accountability for downstream impacts of care quality in settings such as hospitals (Eckermann 2004; Eckermann and Coelli 2013).

Such initiatives are informed by and support the WHO active ageing policy framework (WHO 2002a, b; Kalache 2013). The WHO framework for active ageing aims via individuals and communities to 'realise their potential for physical, social and mental wellbeing throughout the life course and to participate in society according to their needs, desires and capacities, while providing them with adequate protection, security and care when they require assistance'. Importantly, these reforms

directly face the joint challenges of successful ageing and address the inefficiency and lack of sustainability of existing health and aged care paradigms and practices in the context of budget constraints and an ageing baby boomer cohort.

For example, populations with dementia worldwide are expected to double by 2030 and triple by 2050 from that in 2010 (World Health Organization and Alzheimer's Disease International 2012). Similarly, in Australia the population with dementia is projected to rapidly increase from 2010 to 2040 with ageing of the baby boomer cohort, to more than 600,000 in 2040, more than two and a half times that in 2010 (Access Economics 2009).

In the context of such projections, the Productivity Commission report 'Caring for older Australians' was released in 2011 (Productivity Commission 2011) recommending a simplified gateway into aged care and a separation of accommodation and care aspects of aged care. In April 2012 the Federal Government 'Living longer living better' age care reform package was released in response. Residential age care places increased from 40,000 to 100,000, with mean testing of individuals' contribution (family home exempted) alongside a principle focus on aiding transition care from hospital into aged care and addressing hospital bed blocking. In addition $1.2 billion was allocated over 5 years to tackle workforce shortages and modest increases in community care most notably a $880 million for additional home care packages and $75 million for co-ordination of care. Overall, the recommendations and policy response focused on volume of places in age care facilities and represented business as usual rather than reform.

Key issues that were not addressed included:

(i) Community and self-care to minimise the need for aged and dementia residential care places – age- and dementia-friendly community policies;
(ii) The appropriateness of design of age care facilities for dementia care; and
(iii) Palliative care options reflecting palliative population preferences and domains.

As Kalache (May 2012; Kalache 2013) noted huge amounts went to institutions, some was paid to professionals, and a little to the community, especially the most women as carers behind the scenes holding things together; women untrained, unsupported and unrewarded while expected to give their best. Virtually nothing is given to self-care where people are in control of their health and process of ageing. Health education and promotion in the community were almost completely absent, with health education only funded for palliative care and accessing of primary health facilities by people in residential accommodation, with no mention of those still at home.

## 12.3 Ageing Reform Options in the Community

Kalache emphasises that more needs to be done to facilitate self-care and informal care in the community to enable successful ageing and preventing the need for expensive aged care, while in Australia 'the emphasis on residential and institutional care and the funding allocated to them is misguided' (May 2012). He suggested the distribution of funds were allocated the wrong way around given only 6%

of the over 65 population were in some form of residential care (7% of those aged 75–84 and 31% of those over 85 ABS 2007), while the vast majority live at home with informal family, friend, neighbourhood and local community care. Seventy-one percent of over 65 populations in home were living with other people in 2012 (ABS 2013a, b), while 2.7 million informal carers were estimated in aged, disability and long-term health condition care settings. More recently, during the 2013–2014 financial year, 7.8% of the over 65 population were in some form of residential care during the year (AIHW 2015).

In considering community programmes, recent Australian policy has shifted from funding services to funding individuals (DSS 2015). While having potential for more flexible care options over time, this can lead, as in the UK experience, to devolution of community capacity, agency cream skimming and reduced local sector collaboration in response to local needs (Glendinning 2012). Removing services and capping community care and respite funds (both the number of packages available and funds to support community care for older people) can also be problematic when government restricts supply or where market providers are unwilling or unable to meet these needs – with potential system unintended impacts and associated costs of older people being prematurely admitted to residential aged care (Phillipson 2016).

Alongside the ageing of the population, we are witnessing a corresponding increase in the number of people living with dementia (World Health Organisation and Alzheimer's Disease International 2012). Many of the issues faced by older people are likely to be felt most acutely by particular groups, such as those living with the cognitive impairment. However, the issues faced by those with dementia have been largely ignored within age-friendly initiatives, especially within urban design (Buffel et al. 2012).

To address this gap, we have seen a growing call for the support and creation of 'Dementia-Friendly Communities' (ADI 2016; DOH 2015). A dementia-friendly community has been defined as 'a place or culture in which people with dementia and their carers are empowered, supported and included in society, understand their rights and recognise their full potential' (ADI 2016). Examples of dementia-friendly initiatives recently documented by ADI (2016) in 35 countries throughout the world include Alzheimer Cafés in the Netherlands; 'Dementia Friends' programmes in Japan and the United Kingdom; the 'Together for a Dementia-friendly Bruges' campaign in Belgium; in Uji in the prefecture of Kyoto, Japan; and 'Dementia-friendly Kiama' in NSW, Australia, which has been recognised by the WHO as a model healthy community (Phillipson et al. 2016).

To enable successful ageing in the community, the health system needs to robustly evaluate and invest in programmes for the elderly that build health, education and social capital. That is, programmes to make communities and services age and dementia friendly. This includes age- and dementia-friendly public transport systems (trains, busses), taxis and taxi drivers, street signage, public libraries, shops, walking and community garden groups, walking paths, gardens, parks and facilities, communal eating areas and mechanisms for civic participation (ADI 2016). The success of such policies in practice in keeping aged and dementia populations active in the community are highlighted in whole community approaches internationally such as 'Age Friendly Cities' (WHO 2007) internationally and the WHO recognised 'Dementia Friendly Kiama' in Australia (Phillipson et al. 2016).

Such successful community-based health promotion and chronic disease prevention programmes where they have community ownership have potential for positive network and multiplier effects across populations and over time. As Chap. 4 highlighted, this creates potential for a broadening of benefits that does not arise with individual interventions. The extent of multiplier effects over time on community activities from programme investment flowing across networks provides a robust quantitative indicator of community ownership; engagement with, and building of, social networks and capital; and sustainability of programmes over time (Hawe et al. 2009; Hawe and Shiell 2000; Shiell and Hawe 1995; Shiell et al. 2008, Eckermann et al. 2014; Phillipson et al. 2016).

Triangulated with qualitative evidence and typically short-term effects on individuals observed during evaluation periods, multiplier impacts on activity across community networks enable informed extrapolation of whether findings translate into successful and sustainable programmes with long-term impacts across communities. This was illustrated in the evaluation of the SAKGNP in Chap. 4 where triangulated policy relevant assessment of SAKGNP capital investment found (Eckermann et al. 2014; Yeatman et al. 2013, 2014):

(i) Improved food choices ($p = 0.02$) and kitchen lifestyle behaviour ($p = 0.02$) of individuals in case control as well as improvement in pre-post analysis;
(ii) A multiplier on Commonwealth investment of 5.07-fold ($226,737/$44,758) at 2 years; 1.60 attributable to school and 2.47 to wider community activity; and
(iii) SAKGNP classes scaling up by an average of 17% beyond 2 years and local adaptation in schools – including full curriculum integration indicating strong long-term community ownership.

In light of this evaluation, the Australian government committed further funding of $5.4 million – providing opportunities for 400 new schools to undertake Stephanie Alexander Kitchen Garden National Programs, bringing the total number of schools to 650.

Similarly in communities for aged care, evaluation of community network multiplier effects over time triangulated with qualitative evaluation of community ownership of health promotion programs or initiatives and program effects on participant attitudes and/or behaviours is key to assessing long-term effectiveness and cost-effectiveness of health promotion and primary prevention programmes in these complex community settings.

Such assessment is consequently also key to fair and unbiased comparisons in identifying best options for expansion and contraction across community-based health promotion programmes and individual focused treatment interventions. Effective health promotion programmes with community ownership and network multiplier impacts can costlessly widen benefits over time and across community networks and hence significantly improve their cost-effectiveness. This compares with individual-targeted interventions such as pharmaceuticals. Alongside issues of adverse interactions with polypharmacy (and particularly in older populations) for each medication independently; compliance issues, side effects, time profile of resistance and intolerance, etc. often reduce long-term effects, increase cost and reduce cost-effectiveness in practice, from that in trial settings.

Other low cost while potentially highly effective strategies supporting active ageing in the community, and particularly where they have community ownership and build community networks and social capital include flexible housing and carer living arrangements and more generally modified homes and communal living arrangements. There are in turn many communal living arrangement options including co-housing, granny flats, laneway housing, naturally occurring retirement communities and virtual retirement villages (Newton 2015) that can be adapted to be age- and dementia-friendly environments. Key to evaluating them for any given case are interrelated architectural, social and care factors in relation to supportive living arrangements, the accessibility, safety, functional adaptability and meaningful nature of environments to aged residents (Zeisel 2003, 2006; Fleming 2013; Fleming et al. 2008, 2010, 2015). Factors which we now turn to consider in detail in relation to age- and dementia-friendly nursing home or age care facility design (Sect. 12.4).

## 12.4 Dementia-Friendly Aged Care and Nursing Home Design

In providing dementia-friendly aged or nursing home care, the National Institute on Aging (Zeisel et al. 2003) supports an environmental approach balanced with medication and behavioural supports as necessary as the most effective and cost-effective treatment of dementia symptoms. In Australia, Hammond Care under the leadership of Richard Fleming designed dementia-friendly facilities in line with such principles in the early 1990s. These designs included a circular communal eating and kitchen area radiating out to corridors with clear line of site to this central communal community focus and unobtrusive safety features (Fleming and Bennet 2015; Fleming 2013; Fleming and Purandare 2010). Similarly, in the US 'Memory Care' facilities have been designed under the leadership of Zeisel. Synthesis of evidence supporting such nursing home and community designs from Zeisel et al. (2003) includes:

(i) Common spaces with non-institutional character are associated with reduced social withdrawal (Gotestam and Melin 1987).
(ii) Residential character is associated with reduced social withdrawal, greater independence, improved sleep and more family visits (Minde et al. 1990).
(iii) Increased safety leads to greater independence (Sloane et al. 1991), which in turn is associated with fewer falls (Capezuti et al. 1998).
(iv) Sensory comprehension reduces verbal agitation (Burgio et al. 1996; Cohen-Mansfield and Werner 1998).
(v) Privacy reduces aggression and agitation and improves sleep (Morgan and Stewart 1998).
(vi) Camouflaged exits reduce elopement attempts (Dickinson and McLain-Kark 1998).
(vii) Therapeutic garden access reduces elopement attempts and improves sleep (Stewart 1995).
(viii) Walking paths with multisensory activity nodes decrease exit seeking, improve mood and engage family members (Cohen-Mansfield and Werner 1998).

Such evidence is also reflected in dementia-friendly design principles (Fleming and Purandere 2010; Fleming 2013):

(i) Unobtrusively reduce risks
(ii) Provide a human scale
(iii) Allow people to see and be seen
(iv) Reduce unhelpful stimulation
(v) Optimise helpful stimulation
(vi) Support movement and engagement
(vii) Create a familiar space
(viii) Provide opportunities to be alone or with others
(ix) Provide links to the community
(x) Respond to a vision for a way of life

These principles are applied and reflected in HammondCare dementia-friendly facilities which maintain private areas for residents with rooms connected to a central common area – usually a country style kitchen and eating area with corridors that radiate out to their rooms. This ensures that residents leaving their rooms have immediate line of sight and indeed total visual access to a purposeful communal area. The central indoor communal area also typically has direct line of sight to an easily accessible circular designed garden for residents to access outdoor areas and walk around. Memory pathways with objects or pictures a resident is familiar with lead them back to their rooms. The purposeful while safe and secure private and communal environment encourages residents to remain active and maintain purpose and meaning in using their abilities. HammondCare dementia-friendly facilities designed since the 1990s on these principles enable active communities of dementia patients, observed to have improved physical and mental health outcomes and reduced medication use and more generally provide highly functional and effective while inexpensive care. Places at these facilities, some of which are now 25 years old, are highly sought after.

These types of architectural environmental solutions enable nursing home residents and particularly those with dementia to successfully remain active and interact as a community, avoid common symptoms and move beyond restraints and overmedication. Historically, decline in the use of physical restraints had the unintended consequence of increasing use of such pharmacological restraints (Sloane et al. 1991) with common agitation, aggression, psychotic, depression, and social withdrawal symptoms treated with medications that have multiple side effects.

Like HammondCare, Memory Care facilities in the USA reflect National Institute on Aging key recommendations and finding (Zeisel et al. 2003), that environmental factors are at least as important as medication and behavioural approaches, and support a holistic approach for residents with dementia. The key approach is to look for a non-pharmacological environmental solution before considering a medical solution. In these Memory Care facilities, gardens are seen as crucial in helping dementia care residents feel less trapped and become more attuned to the natural rhythms of day and night. In all facilities an easily accessible garden comprises a simple circular path feature with simple unilateral paths that prevent residents from feeling lost, the feeling that generally leads to wandering. Memory Care facilities for

dementia residents designed under the guidance of Zeisel have made sure they look like homes, not institutions (A place for mom 2013). Nursing stations are absent. The staff do not wear uniforms, and every room is at a residential scale and encourages natural congregation. Sensory elements decorate hallways – pictures cohesive with destination and era of the residents. Beside rooms, 'memory boxes' contain personal memorabilia, and residents don't have to remember their room number or location rather they can recognise iconic images from their past.

More generally, the senses of residents are triangulated to their location in a setting – so that the social cues and what people see, hear, touch and smell, all give them the same information about the environment. Social hubs such as a country kitchen setting should feel, sound, look and smell like a country kitchen, not an institutional setting, while gardens should be visible and highly accessible via an unlocked and easily located door to enable and encourage frequent use. These Memory Care facilities designed following the National Institute on Aging recommendations (Zeisel et al. 2003) in practice have shown that dementia-friendly environments can reduce anxiety, depression, social withdrawal, hallucinations, agitation injuries, sleep disturbances and wandering and require less medication (A place for mom 2013).

Nevertheless, arguably the most compelling evidence in practice of the impact of age- and dementia-friendly care environmental design on people living with dementia comes from the Hogewey village in the Netherlands for people with advanced dementia. The village opened in 2009 and has 23 lifestyle and memory customised residences (Godwin 2015). The residences are on a residential scale and include internal gardens while more generally enabling free and safe movement within the village in having the central village enclosed by the residences around it. Each residence has 6–7 residents (152 residents in all) and one of seven lifestyles reflecting residents backgrounds and interests. These lifestyles include 'Gooise' or aristocratic Dutch; the 'ambachtelijke' or working class; the 'Indische' or those of Indonesian origin who migrated to Holland from the former colony; the 'Huiselijke' or homemakers; the 'culturele' who enjoy art, music and theatre; the urban sophisticates who relish city life; and the 'Christelijke', for whom religion is paramount – whether Christianity or another faith. Dementia residents in this familiar environment are given reminiscence therapy by carers, given prompts throughout the residential environment to aid recall (Godwin 2015, Fernandes 2012). The aim of reminiscence therapy is to maximise independence and autonomy, supported by more general interaction in the central communal area to the village which has a hairdresser, cinema, grocery store/supermarket, restaurant/café/pub and doctor and physiotherapist, each run by carers. Carers support residents' environmental options, recall and choices, where for example as Godwin (2015: 28–29) notes dementia residents eating in the supermarket are not accused of shoplifting while staff simply later return food where a resident takes home many times more food than they need. Like the US memory centres, staff have no uniforms and are trained to where possible include a resident in their activities, reinforced by rules such as staff being required to be accompanied by a resident to be served in the supermarket.

In general the village design enables free movement and is communal, purposeful and active and promotes fun. It is also a safe environment, with the 23 residences

closing the central communal area to traffic, while if residents wander to the unobtrusive village exit, the receptionist tells residents it is currently not possible to access. Hogewey had building costs equivalent to that of other advanced dementia patient facilities (19.1 m Euros in 2009), and has the same staff running costs as Dutch nursing homes with equivalent advanced dementia patients. However, Hogewey residents have been shown to require fewer medications and are observed to have more fun and joy – regarded as the most important thing at Hogewey (Godwin 2015; Tagliabue 2012, Tinker 2013). Hence, for those with advanced dementia, the Hogewey village is both effective and cost effective relative to alternatives and indeed represents care which Goodwin (2015) describes as that which in her experience comes closest to engendering Kitwood's (1998: 23) 'trustful serenity'.

The dementia village model with residential and communal memory environments created for reminiscence therapy and active dementia community interaction freedom of movement and much lower use of medications has generated intense international interest. It has been described as an example of a successful model of 'liberty in an adapted environment' (Kremer 2013) or 'a prosthetic environment compensating for certain disabilities' (Godwin 2004) where residents live until they die, and death is handled sensitively (Godwin and Walters 2009). The general model of Hogewey is being adapted to other cultural backgrounds in designing similar facilities for people with dementia in other places such as Rome, Italy, while tailoring adapted environmental settings to resident's lifestyles as part of optimising reminiscence therapy (Dementia Village Advisers 2016).

The only real criticism at a social level has been an essentially ethical concern raised by Charter (2012), suggesting that the reminiscence therapy and village environment may be too insular and removed from other parts of society and a deception or illusion that hoodwinks residents. That is, the village is suggested to be too safe in creating an environment for people with dementia perceptions of what is socially acceptable, that is in contrast with the 'real world'. In the context of people with late-stage dementia living in Hogewey until they die and alternative environments available for late stage dementia care, Godwin (2015) rejects outright Charter's (2012) suggestion that this is an ethical problem. A view no doubt also supported by the residents' positive interactions as a community in the environment created and the long waiting list to gain entry to Hogewey. Godwin (2015) does nevertheless offer some practical suggestions for improvement following principles and guidelines such as those of Zeisel (Zeisel et al. 2003; Zeisel 2006) and Fleming (Fleming et al. 2008; Fleming and Purande 2010; Fleming 2013). These included suggestions in relation to use of better street signage and direction to toilets with words and symbols, contrasting colours to distinguishing toilet seats and walls from the toilet and use of raised garden beds to allow greater access of residents to gardening.

Hence, analogous to age- and dementia-friendly cities or communities enabling active successful ageing, various forms of aged care facilities and nursing homes can be designed to have indoor and outdoor spaces allowing active dementia-friendly communities, while reducing use of physical and pharmaceutical constraints (Sloane et al. 1991; Zeisel et al. 2003; Godwin and Waters 2009; Fleming and Purandere 2010; Fleming and Bennett 2015; Godwin 2015). The principles for designing and

making aged care facilities dementia friendly are essentially the same as for communities and cities – providing safe, functionally active, purposeful environments for dementia populations to remain active in their community and have meaning.

## 12.5 Palliative Care Reforms – Optimising Potential of Some Promising Low-Cost and Palliative Domain Supportive Options

Similar to community and aged care, in palliative care, it is key that the community environment in which palliation takes place enables active participation and meaning for palliative populations to both meet palliative preferences as well as reduce institutional care. Palliative populations consistently indicate primary preferences for finalising personal and financial affairs, while communicating and spend time with their family and friends in the place where they want to be – usually at home or in their community of choice and associated minimising of family and carer distress in the process of dying (McCaffrey et al. 2014, 2015, 2016).

These palliative patient preferences do not generally support hospitalisation or expensive medical therapies with side effects that interfere with the ability of patients to finalise affairs, often distress families and carers and usually remove them from where they want to be for palliative processes and place of death. Rather, palliative preferences usually support at home care with greater community (family/friends) involvement. Importantly, this implies that if our society gets palliative care options and policies right for patient preferences, expenditure in last year of life can be reduced while improving palliative patient primary domains of interest – the ability to finalise affairs with family and friends in place they want to be, usually their home and community.

As considered in Chap. 4, to enable moving beyond palliative inappropriate survival-focused measures, requires explicitly allowing for multiple key domains that can't be integrated with patient survival – finalising affairs, family and carer distress, place of palliation and place of death. As Chap. 10 highlighted and illustrated, evaluation methods to support palliative evaluation have been developed which extend advantages of robust multiple strategy comparisons (Eckermann et al. 2008, 2011) to enable such robust multiple domain comparison (McCaffrey et al. 2015). We now consider some promising low-cost therapeutic options to enable palliative domains to be better supported in prevalent key palliative care indications of intractable pain management, delerium and cancer care, and associated policy issues in optimising their net clinical and economic benefit.

Palliative populations with intractable pain would prefer therapies which better optimised their needs for pain relief and reduced side effects relative to opioid therapies (Lyapustina and Alexandre 2015) or ketamine (Hardy et al. 2012) and more generally to enable them to finalise their personal and financial affairs with family and friends in their place of choice, usually at home. Medicinal cannabis offers a promising low-cost palliative care option in each of these respects and par-

ticularly where policies and programs are optimised to palliative population and individual patient daily needs. In particular this points to research, reimbursement and regulatory policy issues around optimising net clinical and economic benefit of terpene, CDB and THC rich medicinal cannabis strain cultivation, production and provision, as we now consider in detail.

### *12.5.1 Optimising Medicinal Cannabis as an Effective, Low-Cost and Palliative Domain Supportive Programme Option*

When informing public policy decisions in public health systems, community interests, values and preferences should predominate over those of vested interests such as medical device and pharmaceutical manufacturers (Mooney 2012). Within this public policy context, the underlying objective for health economic analysis in informing joint research, reimbursement and regulatory decisions and can be best represented as budget-constrained optimisation of community incremental net benefit. That is, the community value of incremental effects less incremental cost derived from alternative actions across health system and wider social systems such as aged care and palliative care (Eckermann et al. 2010). Hence, to best inform policy options and regulatory decisions requires synthesis of international evidence of expected individual patient and population effects and costs of medicinal cannabis strain cultivation, production and provision in practice relevant to the local decision context.
In optimising medicinal cannabis regulatory options for cultivation and programme provision, synthesis of international evidence should consider the joint health outcome, cost and research implications across options and key local decision contexts such as jurisdictional cultivation conditions. In the case of Australia, these considerations are particularly important, given legislation passed in February 2016 to enable cultivation and manufacturing for compassionate access programmes and research (Lee 2016). Details for associated regulation of cultivation, manufacturing and distribution followed more recently in October 2016 (Australian Government Department of Health Office of Drug Control 2016; Cooper 2016 ).

International scientific, trial and practice evidence and particularly that from the most advanced medicinal cannabis programme internationally in Israel (Kapalos 2016, Tikun Olam 2016) align in pointing to maximising net clinical and economic benefit from medicinal cannabis with the lowest cost cultivation and production method of the highest quality varieties required for optimising patient individual needs and tolerance. That is, cultivating whole plant (i.e. full spectrum) varieties rich in terpenes and cannabinoids (CBD- and THC-rich cultivars) using good agricultural practice (GAP) outdoors where climatic conditions permit. This then naturally extends to manufacturing and distributing appropriately cultivated varieties in whole plant or extract forms (crude plant, infused oils, tinctures, etc.) to Medicinal Grade Cannabis. The appropriate available palette of terpene- and cannabinoid-rich varieties enable provider optimising of patient outcomes with precision medicine given their needs and tolerance.

Outdoor cultivation of terpene-, cannabidiol- (CBD) and tetrahydrocannabinol (THC)-rich varieties in countries with ideal outdoor growing conditions such as Australia optimises:

(i) The range and quality of cannabis varieties and subsequent extracts required to enable net clinical benefit maximisation in compassionate access and particularly palliative populations for the most prevalent intractable pain conditions treated;
(ii) Cultivation and health system cost and environmental impacts from related energy use; and
(iii) Continuing research for precision medicine in optimising palliative patient net benefit given their individual daily needs, tolerance and preferences.

Hence, international evidence points to cultivation decisions being pivotal to optimising compassionate access patient needs and health-care system effects and costs.

### *12.5.2 International Scientific, Trial and Practice Evidence*

Since the early 1990's when the human bodies endocabbinoid system was discovered it has been established as a key regulator of homeostasis for many of the bodies functions and organs. This is reflected in the USA National Academy of Science (2017) review findings that cannabis and cannabinoids has:

(i) Conclusive or substantative evidence of effectiveness for treatment of chronic pain in adults, anti-emetics in the treatment of chemotherapy induced nausea and vomiting and patient -reported multiple sclerosis spasticity symptoms;
(ii) Moderate evidence of Cannabis being effective for Short term sleep disturbance, Fibromyalgia and chronic pain with MS; and
(iii) Emerging evidence of potential for effectiveness in increasing appetite and decreasing weight loss with HIV/AIDS and addressing Tourette Syndrome, anxiety, PTSD, symptoms of dementia and eye pressure in Glaucoma.

Scientific research highlights that entourage effects between terpenes and cannabinoids (CBD and THC and potential other minor cannabinoids – CBG, CBN, THC-V, etc.) to magnify therapeutic impacts and minimise side effects (Gallily et al. 2015; Russo 2011; Wagener and Ulrich-Merzenich 2009) and are reflected in comparative trial results and practice (Johnson et al. 2010; Tikun Olam 2016). For example, for the most common current use condition of chronic pain, in palliative and cancer populations the most adequately powered RCT, a three-arm study for the most prevalent palliative and medicinal condition of severe or intractable pain (Johnson et al. 2010), shows 43% of the terpene-rich CBD/THC arm had significant pain response (greater than 30% improvement), approximately double that of, and statistically significantly higher ($p = 0.014$) than, placebo (21%) or THC as a single agent (23%).

For more than a decade now, Tikun Olam in Israel (Tikun Olam 2016) and Bedrocan (Bedrocan 2016) in Europe have cultivated and manufactured CBD, THC and terpene-rich varieties and extracts in accordance with good agricultural practice (GAP) and high-quality assurance standards, under secure conditions. In compassionate access populations in Israel and Europe, these CBD-, THC- and terpene-rich plant varieties and extracts have been shown to enable optimisation with precision medicine approaches to individual needs and tolerance.

Individual patient needs are optimised in choosing appropriate CBD-, THC- and terpene-rich varieties and mode of administration (tinctures, oils, tablets as well as smoking modalities) and titrating up dosing to optimise benefits given patient needs, tolerance and potential for side effects. Typically, such optimisation is achieved with patients taking CBD-rich varieties during the day and THC-rich varieties at night in optimising symptom benefits while avoiding potential for THC side effects having a negative effect (drowsiness particularly) on daily activities.

US medicinal cannabis cultivation, production and distribution are less optimised compared to that in Israel (Procon 2016), while more prevalent (de Bruin et al. 2015). Nevertheless in the USA, there is compelling population-level evidence from 13 states where medicinal cannabis programmes were in place between 1999 and 2010, showing that mortality rates from opiate overdoses reduced on average by 25% compared to other US states without these programmes in place (Buchhuber et al. 2014). These impacts also increased over time, with 33% reduction by the sixth year of programmes. More recently between 2010 and 2013, US states with medicinal use of cannabis were demonstrated to have a 12% lower rate of pain relief prescriptions in US Medicaid patients (>65 year) and between 8 and 13% lower rates of anxiety, depression, nausea, psychosis and sleep disorder prescriptions (Bradford and Bradford 2016).

Hence, for palliative and other chronic disease populations, medicinal cannabis compassionate access programmes offer distinct potential for clinical and palliative benefits, as well as health system cost savings relative to alternative therapies. This is particularly the case with appropriately optimised and regulated cultivation, distribution and provision of medical cannabis varieties. These benefits need to be considered relative to unintended consequences of current therapies such as opioids (Buchhuber et al. 2014) and ketamine (Hardy et al. 2012), alongside potential medicinal cannabis risks. As Mather et al. (2013) note in relation to medicinal cannabis, their risks are generally modest and particularly compared against those of not treating the symptoms or alternative treatments in relevant pain or other compassionate access palliative or chronic disease populations.

In Europe, medicinal cannabis has been approved in the Netherlands for over a decade, and medical practitioners have been able to prescribe cannabis preparations supplied by Bedrocan BV. These preparations are made from plants grown according to good agricultural practice (GAP) without using pesticides, characterised by different active constituents (cannabinoids, terpenes, etc.), and extract produced with good manufacturing practice (GMP) to be taken by vaporisation or as herbal tea. The Office for Medicinal Cannabis at the Dutch Ministry of Health, Welfare and Sport also supplies Bedrocan BV preparations as exports to other European countries and is sole import supplier to date of cannabis products to Germany,

Finland, Italy and Norway. Medicinal cannabis has also been legalised in 12 other EU member states, most recently in France, Romania and the Czech Republic, while the use of Sativex – a 1:1 mixture of THC and CBD with broader-spectrum terpenes at a standardised dose has been approved in 17 European countries (Bifulco and Pisanti 2015).

In the USA medicinal cannabis programmes have been operating at a state level since 1996, with 1.02 million medicinal cannabis patients covered under such legislation across 24 states and Washington DC in 2015 (Procon 2017), representing 0.83% of those state respective populations (de Bruin et al. 2015). While US states differ in the scope of indications for medical cannabis use, the overwhelming majority of indications are for severe pain (de Bruin et al. 2015). Severe pain is a qualifying condition for over 90% of medicinal cannabis patients and the primary qualifying condition for 60–70% by state, with muscle spasms and nausea a distant second and third. Given current daily costs in the USA of USD $8 per day for chronic pain therapies (USD $2900 per patient annually), medical cannabis was estimated to be a USD $2.9 billion industry per year in the USA for pain management alone in 2015 (de Bruin et al. 2015).

Based on a current Australian population of 24 million and US prevalence of 0.83%, 90% with pain management needs (de Bruin et al. 2015), Australian medicinal cannabis programmes would be expected to cover approximately 200,000 patients. As in the USA, this population is expected to expand rapidly with ageing of the baby boomer cohort to be in the order of two and a half times this population by 2050 (500,000 in Australia). Hence, a local medicinal cannabis industry for the treatment of chronic pain conditions in Australia, assuming prices as in the USA (USD 8 per day, equivalent to A$11), could be extrapolated following current US prevalence at USD $584 million per year or about AUD$800 million at current exchange rates, while increasing to nearer AUD$2.0 billion by 2050.

However, the expected cost, effect and net benefit impact on the health system of any medical cannabis programme, in any jurisdiction, are dependent on the cultivation method and associated quality and range of varieties cultivated for research and practitioner provision (Caulkins 2010; Bifulco and Pisanti 2015; Pick 2016; Oldham 2015; Mather et al. 2013; Grotenhermen 2012).

### *12.5.3 Opportunity Cost, Cost and Energy Use of Outdoor Versus Indoor Cultivation*

To optimise research and treatment to patient needs, compassionate access providers need access to appropriate cannabinoid- and terpene-rich varieties, the appropriate range and quality of which is optimised with outdoor and greenhouse farming rather than indoor cultivation. Higher quality cannabinoid and terpene-rich cultivars grow better where local growing conditions are appropriate for varieties (narrow vs broad leaf), with natural sunlight, air and space (Caulkins 2010; Grotenhermen 2012; Wagner et al. 2009). Hence, a decision to only allow indoor cultivation and

prevent outdoor cultivation (including greenhouse with supplemental lighting) would have significant opportunity costs for the quality of medicinal cannabis varieties for medicinal cannabis programme and research use. There are also high opportunity costs associated with the much higher direct cultivation costs and environmental impacts from energy consumption with indoor cultivation, the evidence for which we now consider.

Considering the cost of medicinal cannabis cultivation, best current US estimates of factor costs of outdoor cultivation and drying are between 10c and 20c per gram of cannabis, 40% of the estimated cultivation costs of 25c–50c per gram with equivalent THC content in greenhouses (Caulkins 2010). Indoor cultivation with hydroponics is more like $1–2.50 per gram and faces issues of high energy use, plant diseases and mould and more generally issues of plant quality in attempting to replicate growing conditions with natural air and sunlight. The much higher cost of growing cannabis indoors rather than outdoors or in greenhouses relates in substantial part to the huge amounts of equipment and energy used, with indoor lighting rigs as well as ventilation, air conditioning, air filtration and/or heating and watering system. As Oldham (2015) notes the atmosphere is calibrated to mimic outdoor conditions in an unvirtuous cycle, the intense heat from the lights requires air conditioning and fans to keep grow rooms at 75° F, a dehumidifier to prevent mould and a carbon dioxide injection system. Energy use required for indoor production in the USA is also noted to pose significant fire safety risks from overloaded electrical transformers.

This ‘unvirtuous’ cycle also extends to adverse economic and environmental impacts of indoor versus outdoor grown cannabis, with significantly higher cultivation cost as well as energy use with associated detrimental environmental impacts. Each kilogram of cannabis grown indoors is estimated as having energy consumption equating to driving across America seven times (Pick 2016). Total energy use from growing cannabis indoors is currently estimated as being more than 1% of US national energy consumption and producing greenhouse gas emissions (CO2) of three million cars, 15 million tons annually.

While electricity costs can represent up to 50% of an operator’s overhead with indoor cultivation, profits still far outweigh costs, with a pound of medicinal cannabis fetching about US$2500 on the wholesale market, more than fourfold the indoor cultivation costs estimates of US$600 a pound (Caulkins 2010, Oldham 2015). Nevertheless, indoor cultivation is tenfold order of magnitude more expensive than outdoor cultivation (and fourfold more expensive than greenhouses). Hence, where outdoor cultivation conditions are ideal for optimal medicinal varieties for patient needs (i.e. Australia), this substantially and unequivocally dominates indoor cultivation. That is, outdoor cultivation has substantially lower cost of production, in addition to greater expected quality of the terpene, CBD and THC varieties required for optimising the net clinical benefit of medicinal cannabis therapy, given patient preferences, needs and tolerance. Outdoor (and to a lesser extent greenhouse) production consequently enables lower factor cost or health shadow pricing (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014) Hence, it is clear that publicly provided medicinal cannabis programmes optimise net benefit with GAP outdoor cultivation.

In contrast, indoor production of single-agent cannabis extracts and pharmaceutical value-based pricing up to threshold values lead to orders of magnitude higher costs while preventing the ability to optimise net clinical benefit for individual patient's needs and tolerance. Positive entourage effects from CBD-, THC- and terpene-rich varieties are lost, along with the ability to undertake associated research. If compassionate access patients are required to pay for medicinal cannabis, then these arguments for outdoor GAP cultivation and not growing indoors or turning into single-agent pharmaceutical become even stronger on access and equity as well as cost, effect and efficiency grounds.

In Australia, an indoor cultivation and production model such as that of Canada, which is mandated by their lack of outdoor growing conditions, is not justifiable on economic nor clinical grounds, nor sustainable on ecological grounds. Indoor growing in Australia with ideal outdoor growing conditions would face substantial opportunity cost on combined clinical, palliative preference, economic and environmental grounds.

Outdoor (and to a lesser extent greenhouse) cultivation minimises unnecessary energy costs and pollution (Caulkins 2010; Oldham 2015; Pick 2016) and permits organically grown (fertiliser and pesticide free) crops of optimal terpene- and cannabinoid-rich varieties for medicinal use. Attempts to imitate Australia outdoor sunlight and open-air condition indoors elevate cultivation costs by a tenfold order of magnitude, increase dangers of plant disease and drastically increase the carbon footprint. In the USA, cultivators are reaching the same conclusion, with the economic and environmental cost that indoor cannabis cultivation is having. As Alex Cooley from Solstice noted in the Guardian (Sevcenko 2016): 'You can't justify... US$500 a pound to cultivate inside when you can cultivate outside for US$50 a pound'.

In considering potential costs to medicinal cannabis programs added to the factor prices of cultivation are those of potential extraction processes, packaging and distribution and the costs associated with security arrangements at each stage of production and distribution. Distributed costs historically have been in the order of fivefold cultivation costs (Caulkins 2010) and hence in the order of US$0.50c to $1 per gram with outdoor production, US$1–$2.50 with greenhouse production and US$5–$12.50 with indoor production. The upper end of these estimates is more likely for high-quality contaminant-free production required in Australia (Lee 2016). Given the average use of cannabis runs to 1 g/day for pain management (de Bruin et al. 2015), these per gram costs also reflect the expected distributed cost per day of different modes of production. The distributed factor price of outdoor cultivation of terpene-, CBD- and THC-rich plant varieties in Australia can therefore be estimated to be in the order of USD$1 per day for pain patients, about 10% of the USD$12.50 distributed costs of indoor cultivation locally or a value-based price of USD$8 proposed in the USA based on the price of current alternative pain management therapies there.

In relation to cultivation and distribution for medicinal cannabis programme needs, Australia could learn much from Italy's approach to cultivation and their medicinal cannabis programme. Bifulco et al. (2015) highlight that in Italy, access since 2014 has been open to primary physician prescription where the cost of

accessing medicinal cannabis is covered by hospital pharmacy or health insurance. In doing so, the programme has sought to secure patient rights and safety, prescribing when other available medications have proven to be ineffective or inadequate to the therapeutic needs of the patient. To reduce import costs, the Italian Ministry of Health cultivates cannabis plants directly at military secured facilities.

In Australia, supplying medicinal cannabis for pain management across an estimated 200,000 patients, outdoor production could be expected to save in the order of A$730 million annually (AUS$10 per day for 200,000 patients) relative to pain management medications at current US prices, or AUD$1.1billion annually (AUS$15 per day) relative to expected distributed prices of indoor cultivation in Australia. This saving can be expected to rise to by two and half-fold to around AUS$2.7 billion annually by 2050. In addition, for the health system better intractable pain management with such medicinal cannabis in populations of palliative care patients (Johnson et al. 2010; Carter et al. 2011) can be expected to allow a higher proportion to stay at home rather than being in institutional settings such as hospitals, increasing cost savings further. If even half of the absolute additional 22% of palliative patients expected to have clinically significant (>30%) better pain management (Johnson et al. 2010) were enabled to stay at home, then a cost saving of $1430 per-patient treated would be expected, given Australian palliative inpatient care cost are estimated to be A$13,000 more than care at home (A$19,000 vs. A$6000, Swerissen and Duckett 2014). Importantly, patient need and tolerance customised therapy with optimal use of CBD/THC + terpene-rich medicinal cannabis therapies such as that in Israel also reflect palliative care population primary preferences for finalising their affairs with family and friends in a place they want to be – usually their own home or community (McCaffrey et al. 2014, 2015, 2016).

The expected health and economic benefits in palliative intractable pain populations from increasing pain symptom relief and reducing direct and downstream treatment costs could also be expected to translate in chronic pain populations. This is particularly the case given the same issues faced with alternative opioid chronic pain management therapies and US population-level evidence of 33% reduction by 6 years in overdose deaths from opioids (Bachuber et al. 2014) in states with medicinal cannabis programmes, as well as 12% lower use of pain relief prescriptions (Bradford and Bradford 2016). Consequently, appropriate low-cost and net clinical benefit maximising medicinal cannabis programmes in both these high prevalence palliative and chronic pain settings have distinct potential to substantially aid budget-constrained successful ageing or ‘gerentolescence’ (Kalache 2013),

Finally, outdoor GAP cultivation and secure, quality-assured production of appropriate terpene-, CBD- and THC-rich varieties for medicinal cannabis compassionate access patient and programme needs domestically also provide distinct scope for long-term export potential from countries such as Australia, with comparative advantages of ideal outdoor growing conditions, both on grounds of both lower cost and higher quality.

GAP cultivation and secure Medical Grade Cannabis production of appropriate varieties/cultivars and subsequent extracts as in Israel for more than a decade

(Tikun Olam 2016, Better 2017) best supports precision medicine in optimising clinical benefit of treating compassionate access populations in practice. Growing such varieties and cultivars outdoors where climatic conditions are ideal, as over much of Australia, can improve quality further while also minimising costs.

Indoor only cultivation in Australia would only benefit pharmaceutical companies with single-agent therapies and alternative black market and organised crime supply of cannabis without appropriate quality control. These are the only groups who benefit in keeping production costs and prices higher than they should be, while harming the interests of compassionate access and palliative care patients, physicians, the environment and taxpayers who public policy should be representing (Dunmore 2017).

In conclusion, scientific, trial and practice evidence (USA National Academy of Science 2017, Gallily et al. 2015, Johnson et al. 2010, Tikun Olam 2016, Bedrocan 2016) point to GAP cultivated Medicinal Grade Cannabis varieties (particularly terpene-, CBD- and THC-rich cultivars) and subsequent extracts being more effective than current therapies in meeting compassionate access programme palliative pain populations when appropriately used for individual needs and tolerance (Gallily et al. 2015, Tikun Olam 2016). Cultivation outdoors where the climate allows (as in Australia) or where necessary in greenhouses rather than indoors enables optimising net benefit of medicinal cannabis programme via both lower cost and greater effects in the higher quality and appropriate range of terpene-, CBD- and THC-rich varieties required for patient populations in practice and further research alike.

### *12.5.4 Other Promising Palliative Preference Supportive Factor Priced Therapie for Delerium and Cancer Care*

The recent placebo-controlled and well-powered RCT of antipsychotic agents in palliative care patients with delirium (Agar et al. 2017) shows their dangers in this setting on grounds of worsening delirium symptoms (statistically worse for both risperidone $p = 0.02$ and haloperidol, $p = 0.01$ relative to placebo), extrapyramidal side effects (risperidone $p = 0.03$ and haloperidol, $p = 0.01$) and also in terms of survival (HR for placebo survival 1.73 vs. haloperidol, $p = 0.003$, while 1.29 vs. risperidone, $p = 0.14$) in addition to associated higher direct and downstream costs.

The survival findings also support similar RCT findings in dementia patients (Maust et al. 2015). For delirium in both palliative and dementia settings, non-pharmacological therapies addressing individual precipitants and environmental factors and best supportive care with pharmacological use restricted to rescue medication are suggested (Agar et al. 2017, Maust and Kales 2016). The health economics implications of these findings for approaches to maximising net clinical and economic benefit in addressing delirium in dementia and palliative populations as

part of successful ageing are significant in pointing to a better model of care than current practice with pharmacological therapy. Non-pharmacological environmental approaches are preferable on multiple domains of effect as well as cost grounds and reinforce those from Sects. 12.3 to 12.4 in better designing community and aged care environments to be dementia friendly.

However, note that following the health shadow price implications in Chap. 11 (Pekarksy 2012, 2015; Eckermann and Pekarsky 2014), active economic processes of resource shifting and incentive are required to enable such appropriate shifts in practice from current pharmacological practices to net benefit maximising non-patented non-pharmacological therapies. It is in this context that Maust and Kales (2016) despite the clear findings and practice implications of the Agar et al. (2017) study were pessimistic and cautioned about the prospect for reducing use of antipsychotics in palliative care, noting that the non-drug alternatives were time-consuming and not financially incentivised. As Agar (2016) highlights in relation to implications of the Agar et al. (2017) trial results in practice, 'There is no safe or effective medication to manage delirium. We need to invest in our hospitals to focus attention on excellent care. We should value seemingly simple things that can be life changing for a person at risk of delirium'.

In practical terms to optimally treat delirium in settings such as hospitals, Agar (2016) further notes these findings imply 'We need to create hospitals where supporting patients to minimise their risk of delirium is a priority; hospitals where all staff, whether doctor, nurse, or cleaner introduce themselves and remind patients what day of the week and what time it is. We need to take a patient's glasses out of their bedside drawer and help them put them on, remind them to take a sip of water, and help them walk a lap of the ward'. Incentives for such policies pointed to by the health shadow price would be actively supported with use of the net benefit correspondence theorem to compare in practice performance of providers and strategies (Chaps. 8, 9 and 10) or fund with net benefit maximising quality incentives, as Sect. 12.6 highlights.

In considering cancer chemotherapeutics, an all-in-one pH neutral formulation of 5-FU and folinic acid has been developed (Locke et al. 2009). This promises distinct cost, compliance, side effect and tumour response advantages over existing 5-FU formulations (highly alkaline, separate administration of 5-FU and folinic acid) used in many solid cancer treatment regimens.

Expected benefits from animal preclinical studies (Stutchbury et al. 2011) and current phase 1 clinical trial (ongoing) include:

(i) Improved net patient effects with lower side effects and better compliance-related outcomes; and
(ii) Lower direct and administration costs (bolus vs. infusion lines).

Importantly these impacts support use for palliative patient preferences particualrly, as well as wider cancer care settings with a therapy that should be adopted early with best evidence as part of global trial across jurisdictions (as per Chapt. 6–7) and as a factor priced therapy (as per Chapt. 11) in oprimising across research, reimbursement and regulatory processes.

Chapter 11 highlighted the research of Pekarsky (2012, 2015; Eckermann and Pekarsky 2014) showing that the highest value alternative to reimbursing (adopting and financing) new technology is to undertake displacement of least cost-effective services in financing and the most cost-effective expansion of existing services in adoption. Use of medicinal cannabis in palliative care, better supportive care and addressing of environmental and other causes of delirium in palliative and dementia patients and a stable and pH neutral formulation of 5-FU in cancer therapies provide good examples of promising factor price solutions for adoption in expansion that new technology adoption should be compared against.

The key to optimising research to support the potential benefits of such promising strategies, whether better use of existing technologies or new technologies, and avoiding costs of delay to societal decision makers while obtaining best research is use of globally optimal trial designs as identified in Chaps. 6 and 7 (Eckermann and Willan 2009, 2013). More generally research into best expansion and contraction of existing programmes or technologies is required to make best investment and disinvestment decisions and appropriately compare and price new technology. Current market failure to produce evidence in relation to better use of existing programmes point to the highest value for research funding as that related to better use of existing programmes and technology – programme budgeting and marginal analysis (PBMA). That is, PBMA research (Ruta et al. 2005) into the most cost-effective expansion of current programmes, services and technology and contraction of least cost-effective current programmes, services and technology.

If new technology is to aid rather than hinder the affordability of successful ageing in twenty-first century, then new technology needs to be appropriately priced. That is, reflect opportunity costs (best alternative actions) with the health shadow price (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014), the most cost expansion of existing programmes and technology financed by contraction of the least cost-effective programmes with existing technology.

Importantly, the amount we spend at proximity to death could be significantly reduced while also better reflecting palliative population preferences if we culturally dealt better with palliative populations and their care options in our health and social system and supported low cost options that reflect palliative patient preferences. The expected impact on the budget bottom line does not have to be as great as predicted allowing for proximity to death, while better meeting population needs and preferences in palliative reforms as part of successful ageing.

More generally if the use of new and existing technology is to create incentives for budget-constrained net benefit maximising quality of care in practice, then appropriate funding mechanisms as well as the monitoring of provider and health system net benefit efficiency highlighted in Chap. 9 are required. That is, funding mechanisms as well as efficient measures that jointly incorporate and account for the quality as well as cost of care consistent with maximising net benefit and prevent incentives for provider cost and effect shifting and cream skimming. The final policy example for meeting the quality and cost challenge of ageing extends hospital efficiency measures consistent with net benefit maximising quality of care applying the net benefit correspondence theorem in Chap. 9 to a sequential two-stage funding

mechanism (Eckermann 2004, 2009a, b; Eckermann and Coelli 2013). Importantly, this enables budget-controlled and managed transition from current case-mix funding ignoring quality of care to enable active incentives for highest health system relevant quality of care within budget constraints.

## 12.6 Bridging the Silos: Funding for Budget-Constrained Optimal Quality of Care

In Australia legislation was introduced in 2011 to establish the Independent Hospital Pricing Authority (IHPA) and the National Health Performance Authority (NHPA) (Parliament of Australia 2011).

The NHPA was established to monitor and report on performance of health services including public hospitals, private hospitals, local hospital districts and primary health-care organisations (inpatient, non-inpatient, community) (Roxon 2011) with a mission to:

(i) Improve quality, increase transparency and drive value for money in the health-care system and;
(ii) Identify high performers and transfer their successes to other areas, identify poor performers so that action can be taken and provide information for more informed choices.

The IHPA was established to advise on the 'efficient price' of hospital activities in:

(i) Comparing and benchmarking hospitals nationally via analysis of actual activity and costs in public hospitals; and
(ii) Considering the need to ensure reasonable access to public hospital services; safeguard clinical access and quality and ensure the efficiency, effectiveness and financial sustainability of the public hospital system.

Hence, NHPA and IHPA joint missions and objectives were to promote efficiency, quality and accountability of hospital and their impacts across levels and systems of care in monitoring performance and funding. That is, create appropriate incentives for quality of care and prevent incentives for cost shifting and cream skimming and for each hospital provider as well as across sectors (prevention, specialist, hospital, rehab, aged care) and associated funding systems (State, Federal).

More recently with closure of the NHPA on the 30 June 2016 (NHPA 2016), the IHPA notes (IHPA 2016) it's role is to work in partnership with the Australian Commission on Safety and Quality in Health Care (ACSQHC) to ensure that pricing, quality and performance measures for public hospitals are complementary and facilitate a strong national framework for the delivery of hospital services.

These organisations have and continue to operate in the context of activity based funding of hospital inpatient services. As Chap. 9 highlighted activity-based funding of hospital per admission (case-mix adjusted or otherwise) ignores quality of care.

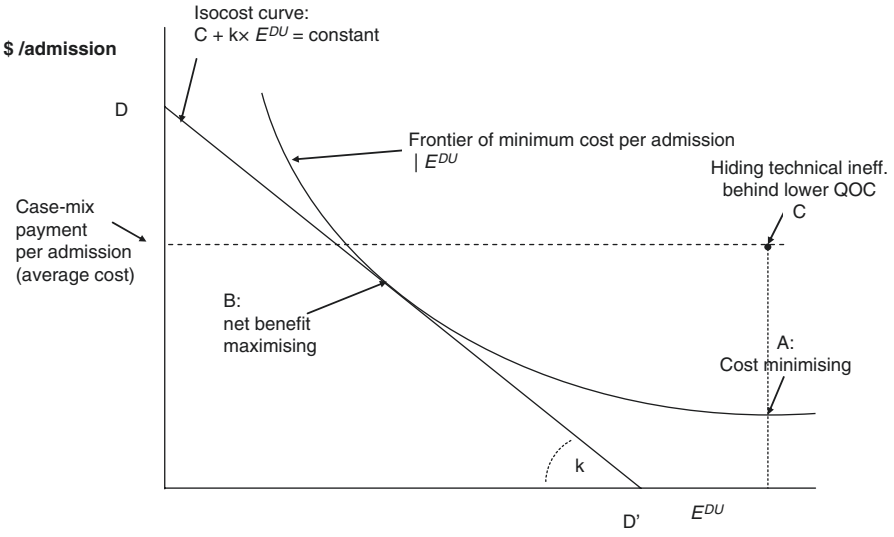


**Fig. 12.1** Case-mix incentives with average cost funding ignoring quality (Source: Adapted from Eckermann 2004)

Hence, while it creates incentives for minimum cost per admission it also creatyes incentives for (Eckermann 2004):

(i) Minimum cost per admission quality of care (while not necessarily minimum cost per admission);
(ii) Cost shifting; and
(iii) Cream skimming.

Hence case-mix funding makes hospitals accountable for the expected average cost of their mix of clinical activities – but not the quality or effects of care. Case-mix proponents try to describe the lack of accountability for effects of care as 'clinically neutrality of case-mix funding' Brook (2002).

However, it is clear that making hospitals accountable for the cost of their care but not the value of the quality of their care creates incentives for cost per admission minimising quality of care and hence hospital and health system allocative inefficiency. In making hospitals accountable to the average expected cost of their admissions, case-mix funding also creates the scope to hide inefficiency behind low quality of care. Figure 12.1 illustrates this on the cost-disutility plane where, as in Chap. 9, quality of care increases with reduction in effect framed from a disutility perspective (e.g. mortality rather than survival). Providers such as those at point C can quite happily operate under case-mix funding at cost-minimising quality while have average expected costs.

Further, such minimum cost per admission quality of care in turn has expected impacts post separation on hospital readmission, treatment in other institutional health-care settings (e.g. general practice, specialist and aged care services) and

informal care in non-institutional settings. Hence, minimum cost per admission does not equate to minimising health system costs, let alone maximising net benefit. To create incentives for appropriate QOC, funding mechanisms need to make providers accountable for the value as well as cost of quality under an appropriate trade off.

To meet their objectives, the IHPA and NHPA would have needed to integrate cost and effect data in efficiency comparison in a way that:

(i) Provided joint accountability for costs and quality of care outcomes in efficiency comparison and pricing;
(ii) Created incentives for net benefit maximising rather than cost min. quality of care; and
(iii) Prevented cost-shifting and cream-skimming incentives. That is, fix the three holes in the case-mix activity-based funding bucket.

### *12.6.1 What Funding Mechanism Provides Appropriate Accountability for Quality?*

In Chap. 9 we showed measuring efficiency consistent with maximising net benefit with the net benefit correspondence theorem (Eckermann 2004; Eckermann and Coelli 2013) creates economic incentives to the extent that identifying peers and relative performance measures influence what is valued. However, economic incentives are directly created by funding mechanisms.

Pay for performance funding measures have previously been used to supplement case-mix funding with block funding at target levels of quality. For example, Lindenauer et al. (2007) describes pay for performance in US public and private hospitals with 1–2% additional yearly payments for hospitals in the top two quality deciles for each of five clinical conditions:

(i) Heart failure;
(ii) Acute myocardial infarction;
(iii) Community-acquired pneumonia;
(iv) Coronary artery bypass grafting; and
(v) Hip and knee replacement.

What incentives for quality of care do such supplementary payments at target levels create? Payments at target or threshold levels create localised incentives for quality of care. The extent of localised incentives depends on the size of block funding at the target level, relative to the expected cost (and probability) of increasing quality to above the target level. However, this points to lack of a theoretical basis for setting the target level and amount of block funding. This in turn naturally leads to questions such as:

(i) Does a theoretical basis exist for funding consistent with maximising NB?
(ii) Can the net benefit correspondence theorem efficiency measurement method be extended to funding?

### *12.6.2 Funding for Net Benefit Maximising Incentives*

Hospitals have control of cost and quality of care (e.g. standardised mortality, morbidity, readmission rates). Hence, funding conditional on differences in quality of care impacts (mortality, morbidity, iatrogenic events, readmission rates, waiting times etc.) can create active incentives for maximising NB, when payments are relative to the NB maximising peer and reward relative performance according to the NB maximising value. However, as funders want to maintain budgetary control, the net benefit maximising value should be determined in a budget-constrained way (as per Chap. 11 and Eckermann and Pekarsky 2014) and hospitals are likely to need a period of time to adjust to economic accountability for quality, as well as costs, of care.

Consequently, a sequential two-stage quality of care funding mechanism is proposed. This starts with a scheduled value for quality at the current industry shadow price for quality across providers (as per Sect 9.3 and Fig. 9.6). This scheduled value rewards current quality. It also ensures remaining within budget, given a second stage buffer payment at a fixed rate for hospital admission across hospital up to the casemix funding level represents industry technical inefficiency at the current industry shadow price for quality. In combination the fist stage quality payment and second stage buffer payment mechansim can be used sequetially where the:

(i) First stage funds according to scheduled price for quality (effects) across providers;
(ii) Second stage apportions the remainder of case-mix funding budget pro rata/admission (buffer payment); and
(iii) Subsequent periods increase the scheduled price for quality.

Steps 1–3 can be repeated until the second stage buffer payment is exhausted, as shown in Fig. 12.2.

Compared to case-mix funding, where quality of care is implicitly valued at 0, low quality of care providers are paid less, while high quality of care providers more. This is appropriate both as their expected costs are better reflected with such payments and a positive value is being ascribed to their quality of care to create increasingly appropriate incentives for quality of care.

Sequentially moving industry towards maximising net benefit within the current budget allows a funding mechanism which provides achievable and more appropriate hospital incentives at each stage; while remaining within overall budget conducive to planning of administrators.

In general this allows a manageable culture change to joint accountability for outcomes and costs of care.

In terms of internal hospital organisation, under the proposed funding mechanism:

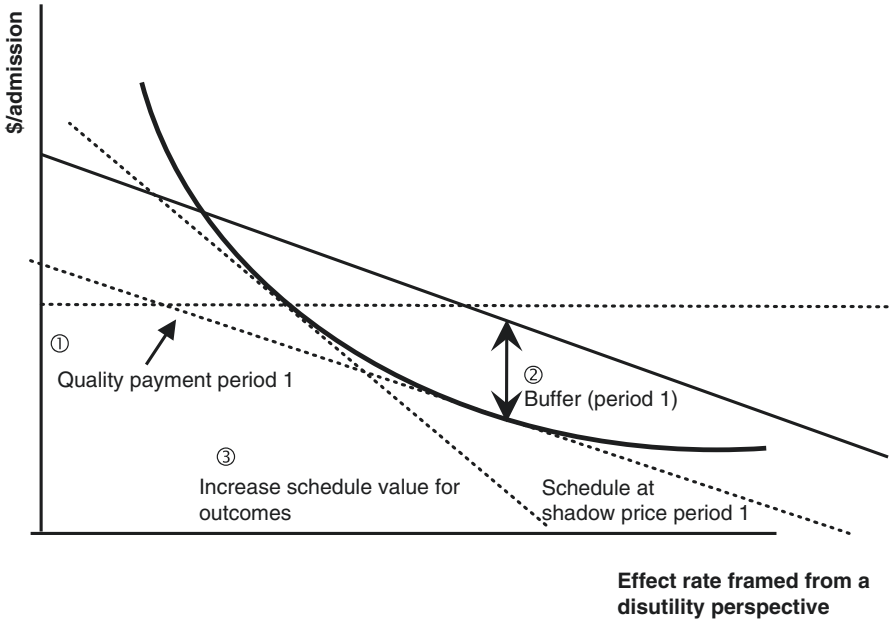


**Fig. 12.2** Sequential two-stage quality funding mechanism moving from current industry shadow price to maximum quality-given case-mix funding (Source: Adapted from Eckermann (2004) – PhD thesis)

(i) Providers have their quality of care valued while becoming accountable for quality of care.
(ii) Administrators no longer act as accountants minimising cost per admission and need to consider trade-offs between the value and cost of quality.

Hence, the funding mechanism encourages meaningful dialogue between administrators and clinicians – trading off cost and value of quality within hospital – rather than administrators cost minimising in battle with clinicians' quality maximising, and related inefficiencies. These inefficiencies include perverse practices such as administrators' cost shifting and clinicians' hoarding resources (their beds, testing, pharmacy, etc.) (Harris 1977).

Analysis thus far suggests quality of care can be improved within budget while there is a second stage buffer payment. However, increased hospital quality of care per admission has been conservatively assumed to have neutral impact on costs post separation. Increased quality of care associated with improving technical and allocative efficiency can be expected to reduce expected need for downstream services post separation, except in the notable case where quality is mortality related and increasing mortality with reduced quality of care is cheap. Consequently, unless quality improvement is related to avoiding low cost mortality, QOC can increase further within the health system budget.

The Quality in Australian Health Care Study (Wilson et al 1995, 1999) evidence suggests this assumption is highly conservative, with deaths less than 5% (4.9%) of hospital adverse events (AEs), while death related AE's were on average more expensive.

Of the 16.6% of hospital admissions associated with an adverse event (Wilson et al 1995), 46.6% of these were associated with temporary disability, 13.7% with permanent disability and 4.9% with deaths.

While AE's on average added 7.1 days to hospital stays, deaths related AEs had 8.2 additional days, while minimal disability (<1 month) added 3.3 days; moderate disability (1-12 months) 8.9 days and <50% and >50% permanent disability added 11.7 and 23.1 days respectively.

Overall, while most quality improvement is non mortality related, even if improving quality is related to reducing mortality, this is pointed to as avoiding high cost, not low cost adverse event related deaths. Consequently the impacts of improving hospital quality can be expected to usually lower health system costs, and enable quality of care to be improved further.

The bottom line advantages of the proposed 2-stage sequential funding mechanism over case-mix funding within any DRG pointed to are:

(i) A first stage payment conducive to planning of administrators;
(ii) Second stage payments that allow a buffer for actual performance and budget control;
(iii) Sequential increases in the unit value of quality related payments until the 2nd stage buffer payment is exhausted provides achievable while more appropriate incentives for improved quality of care at each stage – inefficient hospitals can change behaviour over time;

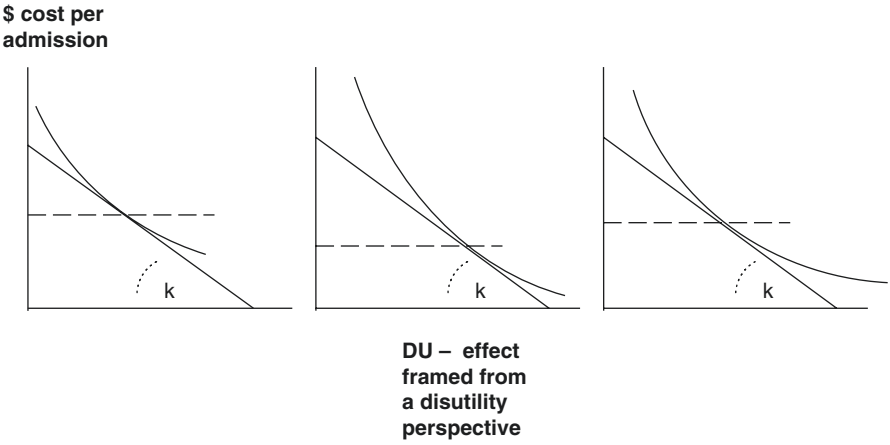


**Fig. 12.3** Equalising values for quality payments across DRGs

(iv) Internally the two stage sequential funding mechanism allowing clinicians and administrators to adapt to economic accountability for QOC; and
(v) Higher quality for current hospital funding reduce expected demands on the wider health system with associated savings providing potential for further quality of care improvement.

For each DRG the proposed performance measurement and funding mechanism allows quality of care to increase as much as possible within that DRGs current budget. Nevertheless, to maximise NB across DRGs given a global budget, resources need to be shifted to equalise final quality value for effects (and incentives for industry shadow price of quality in practice) across DRGs. That is, a process akin to programme budgeting and marginal analysis (PBMA) in marginally shifting resources until the quality values equal across DRGs is required at which point the health shadow price would reflect that of allocative efficiency in practice. For example at a value of k per unit effect across 3 activities in Fig. 12.3, If this value reflects the health shadow price.

In conclusion, case-mix funding of hospitals currently creates incentives for minimum cost per admission QOC and hence allocative inefficiency in not appropriately valuing quality of care and its downstream impacts, while the average cost basis allows technical inefficiency to be hidden behind lower quality of care. The proposed sequential two-stage funding mechanism with first stage quality adjusted payments relative to the cost of the net benefit maximising peer and second stage buffer payment creates appropriate budget controlled joint accountability for cost and quality of care. Hence, the sequential two-stage mechanism allows managed transition in improving quality as much as possible within a case-mix or health system budget.

Measuring efficiency (as in Chap. 9) and funding to make providers jointly accountable for their cost and quality of care addresses both allocative inefficiency (not valuing quality) and technical inefficiency (hidden behind lower quality care with industry average payments for each DRG) with case-mix funding. In doing so, the 2-stage sequential funding mechanism supports budget constrained net benefit choice and use of available strategies, methods and technology (allocative and technical efficiency). Importantly, the coverage and comparability conditions of the net benefit correspondence theorem (Eckermann 2004, 2009a, b; Eckermann and Coelli 2013) underlying these methods as discussed in Chap. 9 provides a robust framework that prevents cost and effect shifting and cream-skimming incentives.

More generally the net benefit correspondence theorem naturally links across research, reimbursement and regulatory decisions in aiding optimisation with:

(i) Robust comparison of multiple strategies and multiple effect domains and summary measures such as expected net loss frontier and planes – directly linking research and reimbursement in HTA (Eckermann et al. 2008; Eckermann and Willan 2011; McCaffrey 2013; McCaffrey et al. 2015);
(ii) Support for joint nature of optimal research and reimbursement decisions using VOI methods (Eckermann and Willan 2007, 2008a, b, 2009; Willan and Eckermann 2012); and

(iii) Performance (efficiency) measurement of providers and their funding in regulation in practice consistent with net benefit maximisation practice (Eckermann 2004, 2009a, b; Eckermann and Coelli 2013). The potential of the NBCT for improving hospital and health system technical and allocative efficiency while appropriately valuing budget constrained quality of care and also mutually supporting research, reimbursement and regulatory decisions reflects the type of integrated reform required to aid budget constrained successful ageing.

## 12.7 Ageing Policy Conclusions

Successful baby boomer ageing (gerentolescence) pioneers for all, but effective and affordable reform needs to address dementia in the community and aged care facilities and end-of-life and new technology challenges. Key reforms pointed to are:

(i) Dementia-friendly communities to provide safe and active ageing in community – dementia- and community-friendly cities, walks, community gardens, transport, shops, etc. (Kalache 2013; Phillipson et al. 2016).
(ii) Dementia-friendly, safe and functional architecture for age care facilities – circular communal area radiating out to corridors with clear line of site to a central communal kitchen area and unobtrusive safety features and line of sight access to a circular garden (Fleming et al. 2010; Zeisel et al. 2003) – encouraging and enabling active and meaningful individual and community interactions.
(iii) Palliative care options and strategies that reflect palliative patient preferences and domains for finalising affairs with family and friends in their community of choice – usually at home and minimising carer and family distress (McCaffrey 2013; McCaffrey et al. 2014, 2015, 2016) such as use of medicinal cannabis in intractable pain palliative populations.
(iv) Research identifying better use of existing programmes and technology and pricing new technology appropriately (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014) and;
(v) Improving health system quality of care within current budgets with robust net benefit maximising efficiency measures and funding mechanisms in monitoring and creating incentives in practice (Eckermann 2004; Eckermann and Coelli 2013).

A major challenge as ever is getting the community's voice heard and having health systems serve community preferences not vested interests, as hilighted by Mooney (2012), where as he suggested in overcoming such vested interests there is a valuable role for citizen juries.

### *12.7.1 Health Economic Tools Aiding Health Reform Gets There*

To meet the reform challenge, robust methods that satisfy coverage and comparability principles for joint research, reimbursement and regulatory decisions are needed including:

(i) Multiplier/network impacts on community programmes highlighted in Chap. 4 (Shiell, Hawe; Eckermann et al. 2014);
(ii) Multiple strategy and domain comparisons in Chaps. 8 and 10 (Eckermann et al. 2008; Eckermann and Willan 2011; Eckermann and Coelli 2013; McCaffrey et al. 2015);
(iii) Downstream quality of care impacts for practice comparisons and policy analysis consistent with maximising net benefit as highlighted in Chap. 9 and Sect. 12.6 (Eckermann 2004, 2009, Eckermann and Coelli 2013) and policy examples (Eckermann 2014a, b; Eckermann et al. 2016; Eckermann and Sheridan 2016);
(iv) Unbiased evidence translation, synthesis and extrapolation as highlighted in Chap. 3 (Eckermann et al. 2009, 2011);
(v) Efficiently designed global research (Chaps. 6 and 7, Eckermann and Willan 2009, 2013) on better use of existing programmes and technology for best expansion and contraction options to emerge and appropriate pricing of new technology relative to best alternative actions (Chap. 11); and
(vi) The health shadow price allowing for decision context – allocative and displacement inefficiency highlighted in Chap. 11 (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014).

## References

ABS 2016. Life Tables, States, Territories and Australia, 2013–2015. ABS cat. no. 3302.0.55.001. Canberra: ABS.

Access Economics. Making choices, Future dementia care: projections, problems and preferences. Report by Access Economics for Alzheimer's Australia. 2009. https://www.fightdementia.org.au/sites/default/files/20090400_Nat__AE_FullMakingChoices.pdf.

Agar M. Antipsychotic medications for delirium don't work and could harm, a new study shows. So what options are left? The Conversation. 2016. http://theconversation.com/drugs-for-delirium-dont-work-and-may-in-fact-harm-69143.

Agar M, Lawlor P, Quinn S, Draper B, Caplan BA, Rowett D, Sanderson C, Hardy J, Le B, Eckermann S, McCaffrey N, Devillee L, Fazekas B, Hill M, Currow DC. 2017. Efficacy of oral Risperidone, Haloperidol or Placebo for symptoms of delirium among patients in palliative care: a randomised trial. JAMA Intern Med. 2017 Jan 1;177(1):34–42. doi:10.1001/jamainternmed.2016.7491. First Published online December 5, 2016.

AIHW. 2015. Aged Care. viewed 18 Feb 2017, http://www.aihw.gov.au/aged-care/.

Alzheimer's Disease International. Dementia friendly communities: key principles. London: Alzheimer's Disease International; 2016.
Al-Zaidy A. Models for paying providers: capitation. British Medical Association. 2015. https://www.bma.org.uk/connecting-doctors/policy_debate/new_integrated_care_models/m/bma_briefings. Accessed 6 July 2016.
A Place for Mom. Best practices for designing Alzheimer's care facilities [internet]. Seattle: A Place for Mom; 2013 [updated 2013 April 4. Available from: http://www.aplaceformom.com/senior-care-resources/articles/alzheimers-care-facilities.
Arrow KJ. Uncertainty and the welfare economics of medical care. Am Econ Rev. 1963; LIII(5):941–73.
Australian Bureau of Statistics. Older people in cared accommodation. Canberra; 2007. https://www.abs.gov.au/ausstats.
Australian Bureau of Statistics. 2013a. Australian social trends 4102.0. http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4102.0Main+Features20April+2013#p2.
Australian Bureau of Statistics. Disability, ageing and carers Australia, summary of findings 2012. 4430.0. 2013b. http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4430.02012?OpenDocument.
Australian Bureau of Statistics, Population pyramid Australia to 2050. http://www.abs.gov.au/websitedbs/d3310114.nsf/home/Population%20Pyramid%20-%20Australia.
Australian Government Department of Health Office of Drug Control 2016. Guideline: security of medicinal cannabis V1.0 October 2016. https://www.odc.gov.au/sites/default/files/guideline-security-medicinal-cannabis.pdf.
Australian Institute of Health and Welfare. 25 years of health expenditure in Australia 1989–90 to 2013–14, Health & welfare exp series 56. Cat. no. HWE 66. Canberra: AIHW; 2016.
Bachuber M et al. Medical Cannabis Laws and opioid analgesic overdose mortality in the United States, 1999-2010. JAMA Intern Med. 2014;174(10):1668–73.
Bedrocan. 2016. http://www.bedrocan.nl/english/products.html. Accessed 21 Oct 2016.
Better. Better medical grade cannabis. 2017. http://www.better-global.com/en/doctors.
Bifulco M, Pisanti S. Medicinal use of cannabis in Europe: the fact that more countries legalize the medicinal use of cannabis should not become an argument for unfettered and uncontrolled use. EBMO Rep. 2015;16(2):130–2. doi:10.15252/embr.201439742.
Bradford AC, Bradford WD. Medical marijuana laws reduce prescription medication use in Medicare Part D. Health Aff (Millwood). 2016;35:1230–1236. doi:10.1377/hlthaff.2015.1661. pmid:27385238.
Britt H, Miller GC, Henderson J, Charles J, Valenti L, Harrison C, Bayram C, Chambers T, et al. A decade of Australian general practice activity 2003–4 to 2012–13, General Practice Series number 34. Sydney: Sydney University Press; 2013.
Brook C. Casemix funding for acute hospital care in Victoria, Australia. Victorian state government; 2002. Available at casemix.health.vic.gov.au/about casemix updated 1 October 2002. Last accessed 16 Sept 2003.
Buffel T, Phillipson C, Scharf T. Ageing in urban environments: developing 'age-friendly' cities. Crit Soc Policy. 2012;32(4):597–617.
Burgio L, Scilley K, Hardin MJ, Hsu C, Yancey J. Environmental "white noise": an intervention for verbally agitated nursing home residents. J Gerontol Psychol Sci. 1996;51B:P364–73.
Butler JGR. Policy change and private health insurance: did the cheapest policy do the trick? Aust Health Rev. 2002;25(6):33–41.
Capezuti E, Strumpf NE, Evans LK, Grisso JA, Maislin G. The relationship between physical restraint removal and falls and injuries among nursing home residents. J Gerontol Med Sci. 1998;53A:M47–52.
Carter GT, Flanagan AM, Earleywine M, et al. Cannabis in palliative medicine: improving care and reducing opioid-related morbidity. Am J Hosp Palliat Care. 2011;28:297–303.
Caulkins JP. Estimated cost of production for legalized cannabis. Rand Corporation Working Paper 764, July 2010. 2010. Available online 21 July 2016: http://www.rand.org/content/dam/rand/pubs/working_papers/2010/RAND_WR764.pdf.

Charter D. For the Alzheimer's victims lost in time, a new village of care. The Times 31 March 2012. 2012.
Cohen-Mansfield J, Werner P. The effects of an enhanced environment on nursing home residents who pace. Gerontologist. 1998;38:199–208.
Colombo F, Tapay N. Private health insurance in OECD countries: the benefits and costs for individuals and health systems. In: Towards high-performing health systems. Paris (FRA): Organisation for Economic Co- operation and Development; 2004. Chap 6.
Cooper. Growing medicinal cannabis is now legal in Australia. Huffington Post 30th October 2016. http://www.huffingtonpost.com.au/2016/10/30/growing-medicinal-cannabis-crops-is-now-legal-in-australia/.
Davis K, Stremikis K, Squires D, Schoen C. Mirror, mirror on the wall: how the performance of the U.S. Health Care System compares internationally: 2014 update. The Commonwealth Fund pub. no. 1755; 2014.
de Bruin D, Ahmad T, Avendano JE, Sajanlal M, Edelstein A, Ryskin M. Medical cannabis has high potential: a joint biotech & tools primer. New York: Bank of America Merril Lynch; 2015.
Deeble JS. Medicare: where have we been? Where are we going? Aust N Z J Public Health. 1999;23:563–70.
Dementia Village advisers. 2016. Accessed 17 Oct 2016 at http://www.dementiavillage.com.
Department of Health. Prime Minister's challenge on dementia: delivering major improvements in dementia care and research by 2015. London: D.o.H. Publications. 2012; 2015.
Department of Social Services. Home care packages program, Operational manual. Canberra; 2015.
Dickinson JI, McLain-Kark J. Wandering behaviour and attempted exits among residents diagnosed with dementia-related illnesses: A qualitative approach. J Women Aging. 1998;10:23–35.
Eckermann S. Hospital Performance including quality: creating economic incentives consistent with evidence-based medicine. Ph.D. Dissertation, University of New South Wales; 2004. http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/.
Eckermann S. Measuring health system efficiency and funding for net benefit maximisation: the health economics of quality of care. Clinical Change Working Paper number 4; 2009a. http://clinicalchange.flinders.edu.au/publications.html.
Eckermann S. Funding to maximise quality of care within a budget: bringing net benefit to casemix funding. Clinical Change Working Paper number 5; 2009b. http://clinicalchange.flinders.edu.au/publications.html.
Eckermann S. Over and underservicing: further reasons to scrap the GP co-payment. The Conversation,August72014.2014a.http://theconversation.com/over-and-under-servicing-further-reasons-to-scrap-the-gp-co-payment-30199.
Eckermann S. Avoiding a health system hernia and the associated outcomes and costs. ANZJ Public Health. 2014b;38(4):303–5.
Eckermann S, Coelli T. Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice. Soc Sci Med. 2013;76:159–68.
Eckermann S, Pekarsky B. Can the real opportunity cost stand up: displaced services the straw man outside the room. PharmacoEconomics. 2014;32(4):319–25. http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s40273-014-0140-3.
Eckermann S, Sheridan L. Supporting medicare health, equity and efficiency in Australia: policies undermining bulk billing need to be scrapped. Applied Health Economics and Health Policy; 2016. Published online 20 July 2016 available at http://link.springer.com/article/10.1007%2Fs40258-016-0258-x.
Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16:195–209.
Eckermann S, Willan AR. The option value of delay in health technology assessment. Med Decis Making. 2008a;28:300–5.
Eckermann S, Willan AR. Time and EVSI wait for no patient. Value Health. 2008b;11:522–6.

Eckermann S, Willan AR. Globally optimal trial design for local decision making. Health Econ. 2009;18:203–16.
Eckermann S, Willan A. Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies. Pharmacoeconomics. 2011;29(7):563–77.
Eckermann S, Willan A. Optimal global VOI trials: better aligning manufacturer and decision maker interest and enabling feasible risk sharing. Pharmacoeconomics. 2013;31:393–401.
Eckermann S, Briggs A, Willan A. Health technology assessment in the cost- disutility plane. Med Decis Making. 2008;28:172–81.
Eckermann S, Coory, M, Willan AR. Indirect comparison: relative risk fallacies and odds solution. J Clin Epidemiol. 2009. Published Online: January 2009. doi:10.1002/hec.1353.
Eckermann S, Karnon J, Willan AR. The value of value of information: best informing research design and prioritization using current methods. Pharmacoeconomics. 2010;28(9):699–709.
Eckermann S, Coory M, Willan A. Consistently estimating absolute risk difference when translating evidence to jurisdictions of interest. Pharmacoeconomics. 2011;29(2):87–96.
Eckermann S, Dawber J, Yeatman H, Quinsey K, Morris D. Evaluating return on investment in a school based health promotion and prevention program: the investment multiplier for the Stephanie Alexander Kitchen Garden National Program. Soc Sci Med. 2014;114:103–12. doi:10.1016/j.socscimed.2014.05.056.
Eckermann S, Sheridan L, Ivers R. Which direction should Australian health system reform be heading? ANZJ Public Health. 2016;40(1):7–9.
Ellis R, Savage E. Run for cover now or later? The impact of premiums, threats and deadlines on private health insurance in Australia. Int J Health Care Finance Econ. 2008;8(4):257–77.
Fernandes E. Dementiaville: how an experimental new town is taking the elderly back to their happier and healthier pasts with astonishing results. Daily Mail, 5 March 2012. 2012. Available online 14 October 2016 at: http://www.dailymail.co.uk/news/article-2109801/Dementiaville-How-experimental-new-town-taking-elderly-happier-healthier-pasts-astonishing-results.html#ixzz4N1BDZz2K.
Fleming R. Improving the design of facilities for people with dementia. Austr J Dementia Care. 2013;1(4):35–7.
Fleming R, Bennett K. The dementia friendly community environmental assessment tool. University of Wollongong; 2015. http://www.enablingenvironments.com.au/audit-tools/http://www.enablingenvironments.com.au/uploads/5/0/4/5/50459523/alzheimers_enviro_assess_tool_a4-signoff.pdf.
Fleming R, Purandere N. Long-term care for people with dementia: environmental design guidelines. Int Psychogeriatr. 2010;22(7):1084–96.
Fleming R, Cookes PA, Sum S. A review of the empirical literature on the design of physical environments for people with dementia. 2008. Retrieved from: http://ro.uow.edu.au/cgi/viewcontent.cgi?article=3923&context=hbspapers.
Fuchs VR. Though much is taken': reflections on aging, health, and medical care. Milbank Memorial Fund Q/Health Soc. 1984;62(2):143–66.
Gallily R, Yekhtin Z, Ondřej HL. Overcoming the bell-shaped dose-response of cannabidiol by using cannabis extract enriched in cannabidiol. Pharmacol Pharm. 2015;6:75–85.
Glendinning C. Home care in England: markets in the context of under-funding. Health Soc Care Community. 2012:292–9.
Godwin B. I did it! Did you see. J Dementia Care. 2004;12(6):16–7.
Godwin B. Hogewey: a 'home from home' in the Netherlands. J Dementia Care. 2015;23(3):28–31.
Godwin B, Walters H. In solitary confinement: planning end of life wellbeing with people with advanced dementia, their family and professional carers. Mortality. 2009;14(3):265–85.
Goss J, Eckermann S and Pinyopusarerk M. Health care for the elderly: costs and some institutional issues. In: EPAC, economic and social consequences of Australia's ageing population, preparing for 21st century. Canberra: AGPS; 1992.
Gotestam KG, Melin L. Improving well-being for patients with senile dementia by minor changes in the ward environment. In: Levi L, editor. Society, stress, and disease. Oxford: Oxford University Press; 1987. p. 295–7.

Grotenhermen F, Müller-Vahl K. The therapeutic potential of cannabis and cannabinoids. Dtsch Arztebl Int. 2012;109(29–30):495–501. Epub 2012 Jul 23

Hardy J, Quinn S, Fazekas B, Plummer J, Eckermann S, Agar M, Spruyt O, Rowett D, Currow D. A randomised, double-blind placebo controlled study to assess the efficacy and toxicity of subcutaneous ketamine in the management of cancer pain. J Clin Oncol. 2012;30(29):3611–7.

Harris J. The internal organization of hospitals: some economic implications. Bell J Econ. 1977;8:467–82.

Harris A. Things you should know about the private health insurance rebate. The Conversation. 2013.

Harrison C, Byram C, Miller G, Britt H. The cost of freezing general practice. Med J Aust. 2015;202(6):313–6.

Hawe P, Shiell A. Social capital and health promotion: a review. Soc Sci Med. 2000;51(6):871e885.

Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. Am J Community Psychol. 2009;43(3e4):267–76.

Helms T, Newhouse JP, Phelps CE. Copayments and demand for medical care. The California Medicaid experience. Bell J Econ. 1978;9(1):192–208.

IHPA 2016. Available at: https://www.ihpa.gov.au/who-we-are.

International Longevity Centre Brazil. Active ageing: a policy framework in response to the longevity revolution. International Longevity Centre Brazil. Ist Edition. Rio de Janeiro; 2015. Available 14 Oct 2016 at: http://ilcbrazil.org/news/active-ageing-a-policy-framework-in-response-to-the-longevity-revolution/.

Johnson JR, Burnell-Nugent M, Lossignol D, Ganae-Motan ED, Potts R, Fallon MT. Multicenter, double-blind, randomized, placebo-controlled, parallel-group study of the efficacy, safety, and tolerability of THC:CBD extract and THC extract in patients with intractable cancer-related pain. J Pain Symptom Manage 2010;39(2):167–179. doi: 10.1016/j.jpainsymman.2009.06.008. Epub 2009 Nov 5.

Kalache A. The longevity revolution: creating a society for all ages. Adelaide thinker in residence 2012–2013. Adelaide: Government of South Australia; 2013. Available at: http://www.flinders.edu.au/sabs/fcas-files/Publications/The%20Longevity%20Revolution.pdf.

Kitwood T. Towards a theory of dementia care: Ethics and interaction. J Clin Ethics. 1998;9(1):723–34.

Kremer P. Alzheimer: Une vie presque ordinaire. Le Monde, 13 November 2013. 2013

Lee J. Senate passes medicinal cannabis legislation. Sydney Morning Herald, 24 February, 2016. 2016. http://www.smh.com.au/federal-politics/political-news/senate-passes-medicinal-cannabis-legislation-20160224-gn2gjk.html.

Lindenauer P, Remus D, Roman S, Rothberg M, Benjam E, Ma A, Bratzler D. Public reporting and pay for performance in hospital quality improvement. New Eng J Med. 2007;356(5):486–96.

Locke JM, Stutchbury TK, Vine KL, Gamble AB, Clingan PR, Bremner JB, Ranson M. Development and assessment of novel all-in-one parenteral formulations with integrated anticoagulant properties for the concomitant delivery of 5-fluorouracil and calcium folinate. Anticancer Drugs. 2009;20:822–31.

Lu L, Pearce H, Roome C, Shearer J, Lang IA, Stein K. Cost effectiveness of oromucosal cannabis-based medicine (Sativex®) for spasticity in multiple sclerosis. Pharmacoeconomics. 2012;30(12):1157–71.

Lubitz J, Riley G. Trends in medicare payments in the last year of life. N Engl J Med. 1993;328:1092–6.

Lyapustina T, Alexandre GC. The prescription opioid addiction and abuse epidemic: how it happened and what we can do about it. Pharma J. 2015. http://www.pharmaceutical-journal.com/opinion/comment/the-prescription-opioid-addiction-and-abuse-epidemic-how-it-happened-and-what-we-can-do-about-it/20068579.article.

Mather LE, Rauwendaal ER, Moxham-Hall VL, Wodak AD. (Re)introducing medicinal cannabis. Med J Aust. 2013;199(11):759–61. doi:10.5694/mja13.10728.

Maust DT, Kales HC. Medicating distress. JAMA Intern Med. 2017;177(1):42–3. doi:10.1001/jamainternmed.2016.7528. First published online December 5, 2016.

Maust DT, Kim HM, Seyfried LS. Antipsychotics, other psychotropics and the risk of death in patients with dementia: number needed to harm. JAMA Psychiat. 2015;72(5):438–45.
May J. The new age of old age. Sydney Morning Herald. May 1, 2012. 2012. Accessed 14 Oct 2016 at: http://www.smh.com.au/national/the-new-age-of-old-age-20120430-1xuzb.html#ixzz45MDtfdyS.
McCaffrey N. Modelling joint cost and outcomes uncertainty on the cost-disutility plane – case studies in palliative care. PhD, Flinders University; 2013.
McCaffrey N, Skuza P, Breaden K, Eckermann S, Hardy J, Oaten S, Briffa M, Currow D. Preliminary development and validation of a new end-of-life patient-reported outcome measure assessing the ability of patients to finalise their affairs at the end of life. PLoS One. 2014;9(4):e94316. doi:10.1371/journal.pone.0094316.
McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Better informing decision making with multiple outcomes cost-effectiveness analysis under uncertainty in cost-disutility space. PLoS One. 2015;10(3):e0115544. doi:10.1371/journal.pone.0115544.
McCaffrey N, Bradley S, Ratcliffe J, Currow DC. What aspects of quality of life are important from palliative care patients' perspectives? A systematic review of qualitative research. J Pain Symptom Manage. 2016. pii S0885–3924(16)30078–1. doi:10.1016/j.jpainsymman.2016.02.012. [Epub ahead of print].
McKay M, Thompson C, Ward D, Green D, Campbell D, McDonell G, et al. GP copayments would increase emergency department waiting times. The Conversation. 2014. http://theconversation.com/gp-co-payment-would-increase-emergency-department-wait-times-28658.
Minde R, Haynes E, Rodenburg M. The ward milieu and its effect on the behaviour of psychogeriatric patients. Can J Psychiatry. 1990;35:133–8.
Mooney G. The health of nations: towards a new political economy. London: Zed books; 2012.
Morgan DG, Stewart MJ. Multiple occupancy versus private rooms on dementia care units. Environ Behav. 1998;30:487–504.
National Hospital performance Authority. Closure of the National Health Performance Authority. 2016. Accessed 21 Oct 2016 at http://www.nhpa.gov.au/internet/nhpa/publishing.nsf.
Newton C. How can we best design housing for Australia's ageing population? The Conversation. 2015. http://theconversation.com/how-can-we-best-design-housing-for-australias-ageing-population-50304.
Oldham J, December 21 2015. http://www.bloomberg.com/news/articles/2015-12-21/as-pot-growing-expands-power-demands-tax-u-s-electricity-grids.
Organisation for Economic Co-operation and Development (OECD). Health at a glance 2013: OECD indicators. Paris: OECD Publishing; 2013.
Parliament of Australia. Report: National Health Reform Amendment (Independent Hospital Pricing Authority) Bill. 2011. www.aph.gov.au/Parliamentary_Business/...13/nhpa/.../nhpa/.../c01.ashx.
Parliament of Australia (2014). Budget papers 2014–15. Retrieved July 25, 2015 from http://www.budget.gov.au/2014-15/content/bp1/html/index.htm.
Peacock SJ, Richardson JR. Supplier-induced demand: re-examining identification and misspecification in cross-sectional analysis. Eur J Health Econ. 2007;8(3):267–77.
Pekarsky BAK. Trusts, constraints and the counterfactual: reframing the political economy of new drug price. Dissertation, University of Adelaide; 2012.
Pekarsky BAK. The new drug reimbursement game: a regulator's guide to playing and wining. London: Springer; 2015.
Phillipson L. Making flexible respite a practical reality. Key note address at Alzheimer's Australia National Respite Summit. Ryde, NSW, April 1 2016; 2016.
Phillipson L, Hall D, Cridland L, Fleming R, Brennan-Horley C. Dementia-friendly Kiama. Final evaluation report. October, 2016. University of Wollongong; 2016.
Pick R. 16th February 2016. Motherboard. http://motherboard.vice.com/read/growing-marijuana-uses-1-percent-of-americas-total-electricity-industry-says.
Private health Insurance Administration Council. Quarterly statistics march 2014. Canberra: Commonwealth of Australia; 2014.

Procon.org. 23 Legal Medical Marijuana States and DC: Laws, Fees, and Possession Limits. Accessed 7thFebruary2016at:http://medicalmarijuana.procon.org/view.resource.php?resourceID=000881.
Procon.org. 28 Legal Medical Marijuana States and DC: Laws, Fees, and Possession Limits. Accessed 5th March 2017 at: http://medicalmarijuana.procon.org/view.resource.php?resourceID=000881.
Productivity Commission. Implications of an ageing Australia: cost of death and health expenditure. Productivity Commission Technical paper 13. Canberra; 2006.
Productivity Commission. Caring for older Australians. 2011. http://www.pc.gov.au/inquiries/completed/aged-care.
Richardson JR, Peacock SJ. Supplier-induced demand: reconsidering the theories and new Australian evidence. Appl Health Econ Health Policy. 2006;5(2):87–98.
Roxon N. 2011. Second reading speech: National Health Reform Amendment (National HealthPerformance Authority) Bill 2011, 3 March 2011 p. 219, viewed 18 Feb 2017. http://parlinfo.aph.gov.au/parlInfo/search/display/display.w3p;query=Id%3A%22chamber%2Fhansardr%2F2011-03-03%2F0012%22.
Russo EB. Taming THC: potential cannabis synergy and phytocannabinoid-terpenoid entourage effects. Br J Pharmacol. 2011;163(7):1344–64.
Ruta D, Mitton C, Bate A, Donaldson C. Programme budgeting and marginal analysis: bridging the divide between doctors and managers. BMJ. 2005;330(7506):1501–3. doi:10.1136/bmj.330.7506.1501.
Seshamani M, Gray A. Time to death and health expenditure: an improved model for the impact of demographic change on health care costs. Age Ageing. 2004;33(6):556–61.
Sevcenko M. The Guardian. 2016. http://www.theguardian.com/us-news/2016/feb/27/marijuana-industry-huge-energy-footprint.
Shiell A, Hawe P. Health promotion community development and the tyranny of individualism. Health Econ. 1995;5(3):241e247.
Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. Br Med J. 2008;336(7656):1281e1283.
Simon H. A behavioral model of rational choice. In:Models of man, social and rational: mathematical essays on rational human behavior in a social setting. New York: Wiley; 1957.
Sloane PD, Mathew LJ, Scarborough M, Desai JR, Koch GG, Tangen C. Physical and pharmacologic restraint of nursing home patients with dementia. Impact of specialized units. JAMA. 1991;265(10):1278–82.
Stewart JT. Management of behaviour problems in the demented patient. Am Fam Physician. 1995;52:2311–20.
Stutchbury TK, Vine KL, Locke JM, Chrisp JS, Bremner JB, Clingan PR, Ranson M. Preclinical evaluation of novel all-in-one formulations of 5-fluorouracil and folinic acid with reduced toxicity profiles. Anticancer Drugs. 2011;22:24–34.
Swerissen H, Duckett S. Dying well. Grattan Institute; 2014. ISBN: 978–1–925015-61-4
Tagliabue J. Taking on Dementia with the Experiences of Normal Life. New York Times April 25 2012; 2012. Accessed 14 Oct 2016 at: http://www.nytimes.com/2012/04/25/world/europe/netherlands-hogewey-offers-normal-life-to-dementia-patients.html?_r=0.
Tikun O. Tikun Olam Medicinal Cannabis; 2016. http://www.tikun-olam.info/. Last viewed 21 Oct 2016.
Tinker B. 'Dementia village' inspires new care. CNN December 27, 2013; 2013. http://edition.cnn.com/2013/07/11/world/europe/wus-holland-dementia-village/.
Trivedi A, Moloo H, Mor V. Increased ambulatory care copayments and hospitalizations among the elderly. N Engl J Med. 2010. 28;362(4):320–8. doi: 10.1056/NEJMsa0904533.
USA National Academy of Science. 2017. The Health Effects of Cannabis and Cannabinoids: The Current State of Evidence and Recommendations for Research. Washington DC. The National Academies Press. https://doi.org/10.17226/24625.
van Dijk CE, van den Berg B, Verheij RA, Spreeuwenberg P, Groenewegen PP, de Bakker DH. Moral hazard and supplier-induced demand: empirical evidence in general practice. Health Econ. 2013;22(3):340–52. doi:10.1002/hec.2801. Epub 2012 Feb 20.
Wagner H, Ulrich-Merzenich G. Synergy research: approaching a new generation of phytopharmaceuticals. Phytomedicine. 2009;16:97–110.

Willan AR, Eckermann S. Expected value of information and pricing new health care interventions. PharmacoEconomics. 2012;30(6):447–59.
Wilson RM, Runciman WB, Gibberd RW, et al. The Quality in Australian Health Care Study. Med J Aust 1995; 163: 458–471.
Wilson RM, Harrison BT, Gibberd RW and Hamilton JD. An analysis of the causes of adverse events from the Quality in Australian Health Care Study. MJA 1999; 170: 411–415.
World Health Organisation. Active aging: a policy framework. Geneva: WHO; 2002a
World Health Organisation. Active ageing: a policy framework. Geneva: World Health Organization; 2002b.
World Health Organisation. Global age-friendly cities, a guide. Geneva: World Health Organization; 2007.
World Health Organization and Alzheimer's Disease International. Dementia: a public health priority. Geneva: WHO; 2012. Available at: http://www.who.int/mental_health/publications/dementia_report_2012/en/.
Yeatman H, Quinsey K, Dawber J, Nielsen W, Condon-Paoloni D, Eckermann S, Morris D, Grootemaat P, Fildes D. Stephanie Alexander Kitchen Garden National Program Evaluation: final report. Wollongong: Centre for Health Service Development, Australian Health Services Research Institute, University of Wollongong; 2013.
Yeatman H, Quinsey K, Dawber J, Nielsen W, Condon-Paoloni D, Eckermann S, Morris D, Grootemaat P, Fildes D. Combining realism with rigour: evaluation of a National Kitchen Garden program in Australia primary schools evaluation. J Australasia. 2014;14(2):17–24.
Zeisel J. Inquiry by design-environment/behaviour/neuroscience in architecture, interiors, landscape and planning -foreword by John b. Eberhard. Revised edition. New York: W.W. Norton & Company, Inc; 2006.
Zeisel J, Silverston N, Hyde J, Levkoff S, Lawton MP, Holmes W. Environmental correlates to behavioural health outcomes in Alzeimers special care units. Gerontologist. 2003;43(5):697–711. doi:10.1093/geront/43.5.697.

# Chapter 13
# Conclusion

This book has shown how a principled approach to health economic evaluation and research can optimise community objectives under resource and budget constraints, but only where key bigger picture structural issues are jointly addressed across research, reimbursement and regulation of practice.

Underlying principles of coverage and comparability and related methods for undertaking robust health economic analysis in optimising across joint research reimbursement and regulatory decisions with budget-constrained community objectives have been introduced and illustrated in relation to addressing key research and policy areas.

Joint principles of coverage and comparability introduced in Chaps. 1 and 2 have been shown to be central to robust methods of analysis whether in:

(i) Within-study analysis (Chap. 2);
(ii) Decision-analytic modelling (Chap. 3);
(iii) Health promotion coverage of multiplier effects across populations over time and comparability with individual-focussed interventions (Chap. 4);
(iv) Palliative care coverage of primary domains of interest and multiple domain comparisons (Chaps. 4 and 10);
(v) Value of information analysis locally (Chap. 5) and Globally (Chaps. 6 and 7) in relation to coverage of key decision contexts, evidence translation and comparability of evidence in relation to location of the INB distribution under uncertainty;
(vi) Multiple strategy and multiple domain of effect comparisons (Chaps. 8 and 10) with comparability in each replicate and at threshold value/s for effect/s relative to the strategy minimising net loss (or equivalently maximising net benefit) and coverage of the scope of strategies and domains of effect compared;
(vii) Efficiency comparisons in practice with explicit comparability (risk factor std.) and coverage (data linkage/modelling) conditions of the net benefit correspondence theorem (Chap. 9) which also underlie robust multiple strategy and multiple outcome comparisons (Chaps. 8 and 10);



S. Eckermann, *Health Economics from Theory to Practice*,
DOI 10.1007/978-3-319-50613-5_13

(viii) The health shadow price and threshold value in relation to coverage of best expansion and contraction of existing technology and integration and appropriate pricing of new technology (Chap. 11); and
(ix) Policy analysis of budget-constrained successful ageing of the baby boomer cohort and beyond (Chap. 12) with coverage of options for better use of existing technology and integration and pricing of new technology across community health promotion and preventative settings, aged care environments (architecture, gardens, etc.) and palliative care settings.

The optimal decision cycle diagram (Fig. 13.1) introduced in Chap. 1 maps the related decisions and the optimal decision pathway for societal decision making to address these related decisions that the book's four parts and associated chapters follow in building across societal decision-maker reimbursement, research, regulation in practice and price and policy decisions.

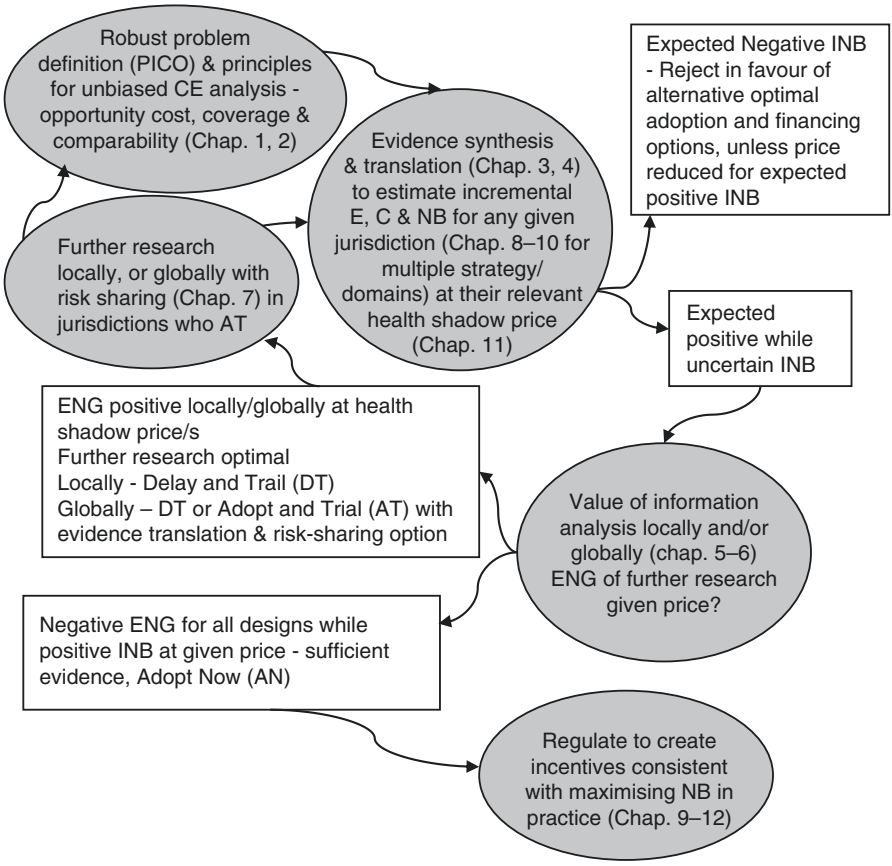


**Fig. 13.1** Optimal decision-making cycles for joint research, reimbursement and regulatory processes locally and globally

Part I (Chaps. 2, 3 and 4) established coverage and comparability principles and related methods for robust analysis in evidence synthesis, translation and extrapolation of joint costs and effects in informing incremental net benefit estimation for two-strategy comparison for individual-based (Chaps. 2 and 3) and community-based interventions (Chap. 4). Importantly, these principles and their consideration to inform unbiased decision making also extend to addressing research, reimbursement and regulatory decisions in HTA and practice in Parts II, III and IV. Key findings in Part I included showing:

(i) Distinct advantages of incremental net benefit over incremental cost-effectiveness ratios as a metric technically, and in making the threshold value for cost-effectiveness-related decision making explicit;
(ii) The need for joint, rather than partial, consideration of costs and effects in informing cost-effectiveness analysis and more generally evidence synthesis and extrapolation to allow adequate coverage of the scope and duration of incremental impacts (costs and effects) following research of O'Brien and colleagues (O'Brien 1996; Briggs and O'Brien 2001; Briggs et al. 2002);
(iii) How to overcome inherent biases arising from use of relative risk in evidence synthesis and translation for binary outcomes commonly required in analysis (survival, progression, etc.) with use of odds ratios (Eckermann et al. 2009, 2011); and
(iv) The importance of multiplier (and more generally network) methods in assessing the long-term success and cost-effectiveness of community health promotion interventions (Hawe and Shiell 2000; Hawe et al. 2009; Shiell et al. 2008; Eckermann et al. 2014) and multiple domain assessment, particularly in areas such as palliative care where key domains cannot be integrated with survival (McCaffrey et al. 2013, 2015).

These findings also started to point to the need to systematically address critical weaknesses of the current political economy in research, reimbursement and regulation biasing towards individual-focussed new technology and away from better use of existing programmes and technology. Findings further reinforced, clarified and established with the health shadow price introduced in Sect. 2.10 and illustrated in detail in Chap. 11 following Pekarsky (2012, 2015), and the failure of community preferences to be reflected in resource allocation and policymaking in key areas such as palliative and end-of-life care (Chap. 10). In each of these areas, community preferences need to be the basis for decision and policymaking if community objectives are to be efficiently and equitably satisfied.

Part II (Chaps. 5, 6 and 7) extended Part I consideration of adoption decisions based on INB under uncertainty to optimising joint research and reimbursement decisions and research design using value of information (VOI) methods. In Chap. 5 Occam's razor was applied to VOI methods to assess their ability to explain relative to their simplicity in address questions such as:

(i) Is further research for a specific HTA potentially worthwhile?
(ii) Is the expected cost of a given research design less than its expected value?

(iii) What is the optimal research design for a specific HTA?
(iv) How can research funding be best prioritised across alternative HTAs?

Value of information methods applying the central limit theorem (CLT) were shown to enable optimising the expected value relative to costs across trial designs in meaningfully addressing these questions while being simple enough to allow for key decision contexts (Eckermann et al. 2010). In particular, for cases of interest, where new options (technologies, strategies or programmes) have expected positive while uncertain INB addressing questions of whether it is optimal to delay and trial, adopt now or adopt and trial where feasible (Eckermann and Willan 2007, 2008a, b, 2009, 2013, 2016; Willan and Eckermann 2010, 2012).

Optimisation in Chap. 5 was illustrated locally where DT and AN are feasible, while AT is usually infeasible. In Chaps. 6 and 7, optimal global trials (Eckermann and Willan 2009, 2013) with coverage of evidence translation and the ability to adopt and trial were shown to provide a circuit breaker that enables first best solutions, respectively, across:

(i) Joint research and reimbursement decisions where a trade-off between opportunity costs of delay and adequate evidence is otherwise faced by societal decision makers and manufacturers alike; and
(ii) Research design, reimbursement, pricing and implementation between manufacturers and societal decision makers in better aligning societal decision-maker and manufacturer interests and incentives for translatable evidence and optimal trial design across jurisdictions.

Without translatable evidence, jurisdictions will not adopt early as part of a global trial. Hence manufacturers need to satisfy their concerns for robust and globally translatable evidence in trial design if they want to both avoid manufacturer opportunity costs of delay and optimise implementation for best global evidence (Eckermann and Willan 2008b; Willan and Eckermann 2010).

The ability to adopt and trial was also shown to enable robust and efficient pricing and risk-sharing arrangements based on robust evidence of incremental net benefit over time, with globally optimal trial evidence alongside practice evidence addressing incomplete contracts for contingencies related to both sets of evidence that otherwise arise (Eckermann and Willan 2013). Nevertheless, the key pricing issue for the usual case of interest with new technology expected to have net incremental cost is the appropriate economically meaningful threshold price for incremental net benefit motivating the health shadow price in Chap. 11 (Pekarsky 2012, 2015). Importantly the health shadow price in optimising decision making reflects opportunity costs under a fixed budget constraint while derived under characteristic health system conditions of allocative and displacement inefficiency (Pekarsky 2012, 2015; Arrow 1963).

This points to the imperative of research on best expansion and contraction of existing programmes and technologies alongside displaced services in assessing new technologies and their pricing (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014). The absence of such research currently for non-patented or non-patentable

programmes and existing technologies is highlighted as the key market failure and source of bias in preventing allocative and displacement efficiency of health systems in practice and appropriately pricing new technology.

Optimal societal decision-maker global trials identified in Chaps. 6 and 7 and methods for robust comparison in practice with the next benefit correspondence theorem in Chap. 9 provide robust methods for non-patented and non-patentable strategies and programmes.

The net benefit correspondence theorem (Chaps. 8, 9 and 10) more generally was highlighted as enabling robust and efficient methods for net benefit (cost-effectiveness) comparison of:

(i) Multiple strategies, of increasing importance with multiple diagnostic and treatment pathways whether genetic testing and individualised care, combination therapies, or multiple modalities. Comparing their relative cost-effectiveness with use of flexible axes on the cost-disutility plane and expected net loss curve and frontier summary measures was shown in Chap. 8 to best inform multiple strategy societal decision making under the Arrow-Lind theorem (Arrow and Lind 1970) following Eckermann et al. (2008) and Eckermann and Willan (2011).

(ii) Multiple domains of interest, shown to be particularly important to areas such as palliative care in Chap. 4 where key domains of interest are not able to be integrated with patient survival. Key methods illustrated in Chap. 10 including use of expected net loss planes and surfaces and cost-effective surfaces in best summarising evidence as well as comparison in cost-disutility space. Importantly this was shown to best inform societal decision making for multiple domain comparisons under the Arrow-Lind theorem following McCaffrey et al. (2013, 2014, 2015).

(iii) Provider efficiency in practice, creating appropriate incentives for net benefit maximising rather than cost minimising quality of care while avoiding perverse cost-shifting and cream-skimming incentives as illustrated in Chap. 9, following Eckermann (2004) and Eckermann and Coelli (2013). These efficiency methods were extended to funding mechanisms to address quality of care issues with case-mix funding of hospitals (Eckermann et al. 2009) in policy illustration Sect. 12.6.

Consequently, the net benefit correspondence theorem (NBCT) method uniquely provides a highly flexible, efficient and robust framework consistent with the appropriate underlying net benefit objective for dealing with joint decisions across what can be very complex comparisons in accommodating as many strategies and providers, for as many domains as required to support coverage and comparability principles. Coverage and comparability conditions explicit in the NBCT (Eckermann 2004; Eckermann and Coelli 2013) are required to be met for unqualified analysis, but more generally provide an explicit framework to improve evidence coverage, comparability and synthesis of cost and effect evidence for health economic analysis consistent with net benefit. Explicit coverage and comparability conditions were shown to be particu-

larly key in prevention of cost- and effect-shifting and cream-skimming incentives in practice (Chap. 9), supporting data linkage, modelling and risk factor adjustment.

More generally still, the NBCT while providing a robust framework across technology assessment and evaluation in practice also leads to summary measures that address missing links between research, reimbursement and regulation. In this respect Chaps. 8 and 10 highlight that comparison on the cost-disutility plane underlying NBCT methods naturally lead to expected net loss frontier (multiple strategies) and surface (multiple effect) summary measures, which in each case provide in one diagram both the optimal strategy if no further research is undertaken and the potential value of further research per patient across relevant potential threshold value/s for effect/s (Eckermann et al. 2008, Eckermann and Willan 2011; McCaffrey et al. 2013, 2015). Hence, these summary measures address key missing links in optimising joint research and reimbursement decisions even in the most complex of cost-effectiveness (net benefit) analyses, with multiple strategies and multiple domains of effect.

In relation to appropriate threshold values for effects and new technology pricing, Chap. 11 shows the health shadow price (Pekarsky 2012, 2015) provides the economically meaningful threshold value and a pathway to allocative and displacement efficiency (Eckermann and Pekarsky 2014). It does this by allowing for better use and pricing of existing and new technology starting from characteristic health system conditions of allocative and displacement inefficiency with market failure and imperfect information. The health shadow price makes clear the critical need for societal decision-maker research into best expansion and contraction of existing programmes, services and technologies and particularly in addressing market failure for those services and technologies that are not patented or patentable and providing a pathway to appropriate pricing and allocative efficiency.

Policy issues addressing health and aged care system challenges faced with the ageing of the baby boomer cohort in Chap. 12 make clear the need for such appropriate consideration of better research into, and use of, existing technology, and pricing of new technology. Alongside supporting publicly provided universal health care on health, equity and efficiency grounds (Mooney 2012, Eckermann 2014; Eckermann et al. 2016), research into better use of existing technologies is urgently needed to address current research, adoption, displacement and pricing biases in considering better use of existing versus new technology. Promising low factor cost while effective and community preference informed options considered include:

(i) Community public health promotion programmes for age and dementia friendly services (Kalache 2013; Phillipson et al. 2016) in Sect. 12.3;
(ii) Aged care facility design and environments for active ageing (Fleming and Purandere 2010; Zeisel et al. 2003) in Sect. 12.4;
(iii) Palliative care factor cost options such as medicinal cannabis therapies that better reflect palliative domains (McCaffrey et al. 2015) and have distinct potential to dominate existing therapies for common palliative symptoms such as intractable pain, particularly when optimised on clinical, environmental and economic grounds with highest clinical value, while lowest cost and energy use outdoor cultivated terpene, CBD and THC rich strains grown in climates appropriate to outdoor growing such as Australia in Sect. 12.5; and

(iv) A sequential two-stage funding mechanism to shift from case-mix funding incentives for cost-shifting and cost per admission minimising quality of care to create budget-constrained incentives for health system net benefit maximising quality of care in hospital practice in Sect. 12.6.

In conclusion, the framework and methods presented have been shown to enable optimising of joint research, reimbursement (adoption and financing) and regulatory (pricing and practice monitoring) processes and decision making. Jointly addressing these related decisions has been shown to be key in meeting current and future challenges of baby boomer ageing and more generally in identifying areas for policy reform to enable a pathway to budget-constrained optimisation of community net benefit. The bottom line for such reforms is that better use of existing programmes and technologies and associated research that reflect community preferences is required and particularly now in facing the challenge of budget-constrained successful ageing of the baby boomer cohort.

## References

Arrow KJ. Uncertainty and the welfare economics of medical care. Am Econ Rev. 1963;LIII(5):941–73.

Arrow KG, Lind RC. Uncertainty and the evaluation of public investment decision. Am Econ Rev. 1970;60:364–78.

Briggs A, O'Brien B. The death of cost minimization analysis. Health Econ. 2001;10:179–84.

Briggs A, O'Brien B, Blackhouse G. Thinking outside the box: recent advances in the analysis and presentation of uncertainty in cost effectiveness studies. Annu Rev Public Health. 2002;23:377–401.

Eckermann S. Hospital performance including quality: creating incentives consistent with evidence-based medicine. PhD Dissertation, UNSW, Sydney; 2004. http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/.

Eckermann S. Avoiding a health system hernia and the associated outcomes and costs. ANZJ Public Health. 2014;38(4):303–5.

Eckermann S, Coelli T. Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice. Soc Sci Med. 2013;76:159–68. http://dx.doi.org/10.1016/j.socscimed.2012.10.020

Eckermann S, Pekarsky B. Can the real opportunity cost stand up: displaced services, the straw man outside the room. Pharmacoeconomics. 2014;32(4):319–25.

Eckermann S, Sheridan L, Ivers R. Which direction should Australian health system reform be heading? ANZJ Public Health. 2016;40(1):7–9.

Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16:195–209.

Eckermann S, Willan AR. The option value of delay in health technology assessment. Med Decis Mak. 2008a;28:300–5.

Eckermann S, Willan AR. Time and EVSI wait for no Patient. Value Health. 2008b;11:522–6.

Eckermann S, Willan AR. Globally optimal trial design for local decision making. Health Econ. 2009;18:203–16.

Eckermann S, Willan A. Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies. Pharmacoeconomics. 2011;29(7):563–77.

Eckermann S, Willan A. Optimal global VOI trials: better aligning manufacturer and decision maker interest and enabling feasible risk sharing. Pharmacoeconomics. 2013;31:393–401.

Eckermann S, Willan AR. Appropriately estimating EVSI with imperfect implementation: improving practice and reducing uncertainty with appropriate counterfactual consideration. Med Decis Mak. 2016;36(3):282–3.
Eckermann S, Briggs A, Willan A. Health technology assessment in the cost-disutility plane. Med Decis Mak. 2008;28:172–81.
Eckermann S, Coory, M, Willan AR. Indirect comparison: relative risk fallacies and odds solution. J Clin Epidemiol. 2009. doi: 10.1002/hec.1353. Published Online.
Eckermann S, Karnon J, Willan A. The value of value of information: best informing research design and prioritization using current methods. Pharmacoeconomics. 2010;28(9):699–709.
Eckermann S, Coory M, Willan A. Consistently estimating absolute risk difference when translating evidence to jurisdictions of interest. Pharmacoeconomics. 2011;29(2):87–96.
Eckermann S, Dawber J, Yeatman H, Quinsey K, Morris D. Evaluating return on investment in a school based health promotion and prevention program: the investment multiplier for the Stephanie Alexander Kitchen Garden National Program. Soc Sci Med. 2014;114:103–12. doi:10.1016/j.socscimed.2014.05.056.
Fleming R, Purandere N. Long-term care for people with dementia: environmental design guidelines. Int Psychogeriatr. 2010;22(7):1084–96.
Hawe P, Shiell A. Social capital and health promotion: a review. Soc Sci Med. 2000;51(6):871–85.
Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. Am J Community Psychol. 2009;43(3–4):267–76.
Kalache A. The longevity revolution: creating a society for all ages. Adelaide Thinker in Residence 2012–2013. Government of South Australia: Adelaide; 2013. Available at: http://www.flinders.edu.au/sabs/fcas-files/Publications/The%20Longevity%20Revolution.pdf
McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Is home-based palliative care cost effective? An economic evaluation of the Palliative Care Extended Packages at Home (PEACH) pilot. BMJ Support Palliat Care. 2013;3(4):431–5.
McCaffrey N, Skuza P, Breaden K, Eckermann S, Hardy J, Oaten S, Briffa M, Currow D. Preliminary development and validation of a new end-of-life patient-reported outcome measure assessing the ability of patients to finalise their affairs at the end of life. PLoS One. 2014;9(4):e94316.
McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Better informing decision making with multiple outcomes cost-effectiveness analysis under uncertainty in cost-disutility space. PLoS One. 2015;10(3):e0115544.
Mooney G. The health of nations: towards a new political economy. London: Zed books; 2012.
O'Brien B. Economic evaluation of pharmaceuticals. Frankenstein's monster or vampire of trials? Med Care. 1996;34(12 Suppl):DS99–108.
Pekarsky, BAK. Trusts, constraints and the counterfactual: reframing the political economy of new drugs. PhD Dissertation, University of Adelaide; 2012. http://digital.library.adelaide.edu.au/dspace/handle/2440/79171.
Pekarsky BAK. The new drug reimbursement game: a regulator's guide to playing and winning. New York: Springer; 2015.
Phillipson L, Hall D, Cridland L, Fleming R, Brennan-Horley C. Dementia-Friendly Kiama. Final evaluation report. Oct 2016. University of Wollongong; 2016.
Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. Br Med J. 2008;336(7656):1281–3.
Willan AR, Eckermann S. Optimal clinical trial design using value of information with imperfect implementation. Health Econ. 2010;19:549–61.
Willan AR, Eckermann S. Value of information and pricing new health care interventions. Pharmacoeconomics. 2012;30(6):447–59.
Zeisel J, Silverston N, Hyde J, Levkoff S, Lawton MP, Holmes W. Environmental correlates to behavioural health outcomes in Alzeimers special care units. The Gerontologist. 2003;43(5):697–711. doi:10.1093/geront/43.5.697.