Practical Methods to Support Research Prioritisation Decisions: Rapid Assessment of the Need for Evidence

David Patrick Glynn

PhD

University of York

Economics and Related Studies

March 2019

Abstract

This thesis aims to further the use of value of information (VOI) methods in research prioritisation. It comprises seven chapters and a freely available online tool. The Chapter 1 provides an introduction to VOI and an overview of the methods that have thus far been developed to address the practical barriers to its use. Chapter 2 illustrates the rapid VOI methods which are the focus of this thesis using a case study from one of six retrospective proposals provided by the national institute for health research (NIHR) trials and studies coordinating centre (NETSCC). This provides an estimate of the value of research in terms of research cost required to gain an additional unit of a disease specific binary outcome. Chapter 3 provides an approach to addressing the challenge of prioritising across disease areas by linking disease specific binary outcomes to a generic measure of health outcome: quality adjusted life years (QALYs). The implications of using a generic health outcome in research prioritisation are explored. Chapter 4 extends the methods from Chapters 2 and 3 to allow for analysis of a wider range of outcomes (including continuous and survival outcomes) and develops a method for estimating the VOI provided by a feasibility studies. Chapter 5 introduces the online tool which implements each method described in previous chapters (https://shiny.york.ac.uk/rane/.) Chapter 6 addresses the question of how to make early access decisions without a full economic model. Granting early access means that treatments can be quickly provided without delay. The cost of this is that it can remove the possibility of research. This chapter provides a framework for early access decisions which is built upon the rapid methods described in previous chapters. Chapter 7 concludes by providing a brief overview of the entire thesis and identifies avenues for future research and policy.

Contents

TABLES

FIGURES

Acknowledgements

Reflecting on all those who have supported me to get to this point I am struck by the fact that it would be impossible to thank them all. However, even though the task is impossible, it is not always right or possible to opt out.

First my family; Mum, Dad, Colm and Laura. In addition to the gift of life, care and financial support Mum and Dad have always supported my studies and have never left me in any doubt about their beliefs in my abilities. Any achievement of mine can never be separated from my family and so I share this PhD with them.

I am deeply indebted to my supervisors; Claire Rothery and Karl Claxton in addition to my thesis advisory panel member David Torgerson. I would like to thank Claire, as my primary supervisor, for her guidance, patience, honesty, support and encouragement throughout this entire process. I would like to thank Karl for teaching me to think as an economist. For giving me opportunities and encouraging me in my transition from pharmacist to health economist I would also like to thank Mandy Ryan, Ifigeneia Mavranezouli and Bhash Naidoo.

For emotional and spiritual support I thank: my fellow PhD students; my meditation teachers Rob Burbea and Thích Nhất Hạnh; and the York Sangha. Thanks to the whole Centre for Health Economics team at York for making me feel so welcome and at home. I could not ask for a better place to do a PhD. Thanks to all the friends I have made here and the friends I have been privileged enough to know for many years.

Finally my partner Alison, my soulmate, my best friend. You are an ineffable source of support, insight and encouragement to me.

To all mentioned and not mentioned here, I feel the debt I owe is truly unpayable.

Declarations

I confirm that the work presented in this thesis is my own and I am the main author. I have conducted all analyses and written each draft. Claire Rothery, Karl Claxton and David Torgerson advised on methods, structure, writing style and commented on chapter drafts. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References. Francesco Ramponi, Georgios Nikolaidis, Natalia Kunst and James Gaughan provided comments on early drafts.

Funding for my studies was provided by a studentship funded by the Centre for Health Economics, University of York. Funding for this project was also received from the UK National Institute for Health Research (NIHR) Health Technology Assessment (HTA) Programme as project number 16/29/01. The views expressed in this thesis are those of the authors and not necessarily those of the NIHR HTA Programme or the University of York.

Abbreviations

Approval With Research (AWR)

Australian Medical Research Advisory Board (AMRAB)

Canadian Institute of Health Research (CIHR)

Cancer Drugs Fund (CDF)

Confidence Interval / Credible Interval (CI)

Corticosteroid Randomisation After Significant Head injury (CRASH)

Corticosteroids (CS)

Deep Vein Thrombosis (DVT)

Efficacy and Mechanism Evaluation (EME)

Expected Value of Partial Perfect Information (EVPPI)

Expected Value of Perfect Information (EVPI)

Food and Drug Administration (FDA)

Glasgow Outcome Scale Extended (GOSE)

Health Services and Delivery Research (HS&DR)

Health Technology Assessment (HTA)

Incremental Net Health Benefit (INHB)

Intensive Care Unit (ICU)

Intra-Cerebral Haemorrhage (ICH)

Invention for Innovation (i4i)

Mini Mental State Examination (MMSE)

Minimum Clinical Difference (MCD)

National Institute for Health and Care Excellence (NICE)

National Institute for Health Research (NIHR)

Net Health Benefit (NHB)

NIHR Evaluation, Trials and Studies Coordinating Centre (NETSCC)

Non-Governmental Organisation (NGO)

Odds Ratio (OR)

Only In Research (OIR)

PATIENT CENTRED OUTCOME RESEARCH INSTITUTE (PCORI)

PHARMACOLOGICAL THROMBOEMBOLISM PROPHYLAXIS (PTP)

PROGRAMME GRANTS FOR APPLIED RESEARCH (PGFAR)

PROGRAMMED DEATH RECEPTOR 1 (PD1)

PROGRESSION FREE SURVIVAL (PFS)

PUBLIC HEALTH RESEARCH (PHR)

QUALITY-ADJUSTED LIFE YEARS (QALYS)

RANDOMISED CONTROLLED TRIAL (RCT)

RAPID ASSESSMENT OF NEED FOR EVIDENCE (RANE)

RESEARCH FOR PATIENT BENEFIT (RFPB)

RISK DIFFERENCE (RD)

RISK RATIO (RR)

SHEFFIELD ACCELERATED VALUE OF INFORMATION (SAVI)

STANDARD ERROR (SE)

TRAUMATIC BRAIN INJURY (TBI)

TRAUMATIC BRAIN INJURY (TBI)

TREATMENT FOR 12 MONTHS (TF12)

TREATMENT FOR 6 MONTHS (TF6)

TREATMENT TO DISEASE PROGRESSION (TTP)

UNITED STATES OF AMERICA (USA)

UNITED KINGDOM (UK)

VALUE OF INFORMATION (VOI)

VALUE OF INFORMATION FOR CARDIOVASCULAR TRIALS AND OTHER COMPARATIVE RESEARCH (VICTOR)

VENOUS THROMBO-EMBOLISM (VTE)

Chapter 1

1.1 INTRODUCTION

Empirical research provides the scientific foundation for modern medicine. Whether carried out by the public or private sector, budgets to fund research and the real resources to carry it out are limited. Research can take many forms such as randomised controlled trials (RCTs), observational studies, feasibility studies. Research prioritisation is the practice of choosing to fund certain research proposals at the expense of not funding others.

Research prioritisation is an important activity. It was estimated that global investment in applied biomedical research was over US$97 billion (adjusted for purchasing power parity) in 2010 (Røttingen et al., 2013). Approximately 90% of this is spent in high income countries. For these high income countries, the same study estimated that 60% of health research and development investments are allocated by the business sector, 30% are allocated by the public sector and about 10% are allocated by charity and non-profit organisations.

Though the majority of research and development investments are made by the private sector, a substantial proportion is made by the public and charity sectors. There are two related reasons for the involvement of the public and charity sectors in health care research. The first, is because a large amount of health care is not delivered through pure private markets. This is due to a variety of reasons including but not limited to: incomplete information, equity concerns, increasing returns to scale and political history (Arrow, 1963; Culyer, 2012; Mooney, 2012). The second reason is that research carried out by the private sector can be seen as insufficient to meet health needs. Research decisions by the private sector are guided with regard to generating profits through sales of products. Research commissioning by public bodies is important in developed countries in areas in which expected profits are insufficient to incentivise adequate research investment. For example, research on antimicrobials (Rothery et al., 2018), novel indications for off patent drugs or new types of psychological therapy. Private incentives are especially inadequate to address medical need in low and middle income countries. In 2010, approximately

10% of global investment in biomedical research was spent in low and middle income and only 1% of all global research funding was allocated to the diseases most prevalent in low and middle income countries, such as tuberculosis and malaria, despite the fact that these diseases are responsible for more than 12% of the global burden of disease (Kieny et al., 2016; Røttingen et al., 2013).

Research prioritisation by the public sector in developed counties is carried out by a range of national bodies. These include the patient centred outcome research institute (PCORI) in the United States (US), the Australian medical research advisory board (AMRAB) and the Canadian institute of health research (CIHR). In the United Kingdom research is prioritised by a branch of the national institute for health research (NIHR) called the NIHR evaluation, trials and studies coordinating centre (NETSCC). In developing countries research is funded and allocated by a mix of national aid organisations, charities and non-governmental organisations (NGOs) such as the UK department for international development (2018), Bill and Melinda Gates foundation (2018), European and developing countries clinical trials partnership (2018) and the US national institutes of health (2018).

The research proposals which these bodies choose between must first be chosen before being presented to decision makers. This involves making research recommendations, i.e. choosing which research questions deserve attention in different disease areas. In the UK the James Lind alliance is an organisation which is dedicated to this question. It is a priority setting partnership which brings together patients, clinicians and carers to identify and prioritise unanswered questions across a range of disease areas (James Lind alliance, 2019).

In making research prioritisation decisions on the basis of need in both developing and developed countries the aim is to fund research which will improve health outcomes. This is a complex task which requires scrutiny. Quantitative methods which estimate the health impact of research projects can improve the transparency and accountability of this process. To this end, a number of methods to aid research prioritisation have been proposed. A review by Fleurence and Torgerson (2004) identified five approaches to aid research prioritisation: subjective methods, burden of disease methods, clinical variations, payback methods and value of information (VOI) methods. A description and comparison of each of these approaches is beyond

the scope of this thesis, however this review concluded that VOI was the only approach which was consistent with the objective "to provide the most health benefits to the population that it serves within the budget constraint and while respecting equity considerations".

The aim of this PhD is to further the use of VOI in research prioritisation by developing and demonstrating methods to calculate VOI which fit within the practical constraints of the bodies responsible for decision making. The remainder of this chapter provides an introduction to VOI, a summary of the barriers to its wider use, an overview of the methods developed to address these barriers and finally a summary of the structure and contributions of this PhD.

VOI is a method to estimate the value of reducing uncertainty before making a decision. It was originally developed in the 1950's to address decision making in industrial engineering and then further developed in the 1960's (Howard, 1966; Myers et al., 2012; Schlaifer and Raiffa, 1961).

In the context of health, VOI has been developed as an approach to dealing with uncertainty in the economic evaluation of health technologies. Economic evaluation involves constructing an "economic model" or "decision analytic model" which is a mathematical model used to inform decision making by comparing the costs and health effects associated with the relevant treatment options (Briggs et al., 2006). By taking account of the likely costs and benefits of different treatment alternatives the optimal treatment choice can be identified in a way which takes account of budget constraints.

The mathematical structure used to predict the relevant costs and benefits will depend on the disease being modelled and the data available. Decision trees represent possible prognosis (e.g. continued disease vs cure) using a series of pathways (Briggs et al., 2006; Drummond et al., 2015). Costs, health consequences and probabilities for each of the possible pathways are attached using the existing evidence. The expected benefits and costs of a given treatment can then be estimated. Decision trees are widely used but they are not well suited to capturing the natural history of more long term and/or chronic diseases (e.g. breast cancer). This is because with chronic diseases patients face a series of competing risks (e.g. different types of recurrence, death from breast cancer, death from natural causes)

and therefore a large number of possible pathways are required to represent this natural history. This complexity results in decision trees becoming "bushy" with many possible pathways. Markov models are used overcome this limitation. Markov models characterise disease natural history as a series of "states". These represent the possible health states that can be experienced by a patient. In the case of breast cancer these may be; local recurrence, metastatic disease, death from breast cancer and death from natural causes. At each time point in the model there is a probability of moving from one state to another. Each state is associated with different costs and health consequences. Different treatments affect how patients move through between states and so Markov models can be used to estimate the expected cots and health consequences associated with different treatments (Briggs et al., 2006; Drummond et al., 2015)[1].

In addition to differences in model structure, there exist different *types* of economic evaluation in which different units used to capture treatment benefits. If benefits are kept in clinical outcomes (e.g. life years gained or dental cavity avoided) then this is known as a cost effectiveness analysis (CEA). This approach can be used when there is a single common effect which can be compared across treatment alternatives (Drummond et al., 2015). The output of this analysis is typically presented as the additional cost associated with an additional unit of outcome (e.g. £1,000 per dental cavity avoided). The limitation of CEA is that often decision makers need to compare the additional benefits of a more expensive treatment against the opportunity cost (i.e. the benefits foregone) associated with higher costs. This creates a difficulty as the health benefits foregone by increased costs may be in a different clinical area to the treatments assessed. To make an informed decision, the benefits gained must be compared to the benefits foregone and this is difficult to achieve in the absence of a generic measure of health outcome which captures health benefits across disease areas. Cost utility analysis (CUA) is often invoked to overcome this issue. In this approach health benefits are captured using a generic measure of health outcome such as quality adjusted life years (QALYs). These measures capture both length and quality of life and so can be used to capture the relative benefits of treatments across disease areas. Results are typically presented as the additional cost

[1] Other more complex model structures also exist but these are beyond the scope of the present thesis (Brennan et al., 2006).

associated with an additional QALY (e.g. £2,000 per QALY gained). There also exists another form of economic evaluation, cost benefit analysis (CBA) which converts all benefits into monetary terms. For example, the monetary value of an additional QALY may be used to convert a CUA into a CBA. In principle this could be used to compare the benefits in one sector of the economy such as health care to the benefits gained in another sector such as transport (UK Treasury, 2003).

Regardless of the type of economic evaluation or the decision model structure, an economic evaluation will always contain uncertainties. For the purposes of this thesis it is important to distinguish between two types of uncertainty; structural uncertainty and input uncertainty (Strong, 2012)[2]. Structural uncertainty is the inherent uncertainty associated with constructing a mathematical approximation of reality. It reflects the fact that we cannot be certain that we have the "true" model i.e. a model that would correctly predict reality if the true values of our inputs were known (Strong, 2012).

The second source of uncertainty is input uncertainty (also known as parameter uncertainty), this reflects the fact that the inputs to our decision model will have some uncertainty associated with them which can be expressed mathematically. For example relative treatment effects may be estimated from a clinical trial, in which case uncertainty will be captured by the standard error (SE). Reducing uncertainty in inputs helps inform decisions about the optimal intervention for subsequent patients and this has health consequences (Claxton, 1999; Eckermann and Willan, 2009; Wilson, 2015).

For example, given existing evidence intervention A may be judged to be the optimal choice; however, due to uncertainty in the relative treatment effect, there is a chance that intervention B is in fact more effective. Therefore, when the existing evidence is uncertain there is always a chance that one of the alternative interventions could improve health outcomes to a greater extent than the intervention which is considered best on average. This means that there are adverse health consequences associated with uncertainty. The importance of this uncertainty is indicated by the scale of these health consequences. The scale is dependent on the likelihood that a

[2] There exist other forms of uncertainty such as "methodological" and "code" uncertainty but an exploration of these is beyond the scope of this thesis (Bojke et al., 2009; Briggs et al., 2006; Strong, 2012).

particular intervention is not the most effective option, how much less effective it is likely to be (in terms of some measure of health outcome), and the size of the patient population facing the uncertain intervention choice.

VOI methods calculate the expected health consequences of the current (input) uncertainty and these expected health consequences can be interpreted as an estimate of the health benefits that could be gained each year if the uncertainty surrounding treatment choice were resolved, i.e., it provides an expected upper bound on the health benefits of further research[3]. These potential expected benefits increase with the size of the patient population whose treatment choice can be informed by additional evidence and the time over which evidence about the effectiveness of the interventions is expected to be useful (see Chapter 2 for further explanation and illustration).

As it estimates the value of reducing uncertainty VOI is well suited to the task of research prioritisation as it can be used to calculate the health gain from reducing uncertainty by funding specific research proposals. However, VOI can also be used in other applications which require trading off the benefits of additional information with the costs of acquiring information such as in the case of early access decisions. Granting immediate access to a new technology may provide expected health benefits but if further research is not possible with approval then the benefits of approval must be compared to the benefits of research foregone. VOI can provide a consistent and transparent approach to illustrating the consequences of this trade-off (see Chapter 5 for further details).

Despite benefits to VOI and evidence that decision makers find the results useful in decision making (Bennette et al., 2016; Bindels et al., 2016; Carlson et al., 2018, 2013; Claxton and Sculpher, 2006), there are a number of well documented barriers to its use in decision making these include: i) human resources, ii) time, iii) computing resources and iv) familiarity with methods.

Human resources represent a barrier as VOI analysis is often based on the results of a decision analytic model which reports results in terms of costs and QALYs

[3] In this thesis VOI is used from a payer perspective to estimate the value of further research in terms of health outcomes. However, VOI can also be used from a manufacturer perspective in which case the benefits of further research will be in terms of expected profit (Willan, 2008).

(Meltzer et al., 2011; Myers et al., 2012). Constructing a full decision analytic model typically requires access to clinical expertise in addition to skills in critical appraisal and synthesis of evidence and decision modelling. In addition to access to skilled labour these models also require a large amount of time to construct and validate. Decision makers often operate within rigid timelines and so cannot easily delay decisions until a full economic model has been constructed. Even in cases in which a decision model exists, conducting VOI analysis using these models can demand access to substantial computing resources (Myers et al., 2012; Strong et al., 2015). This is due to simulations required to calculate VOI (discussed in Chapter 2). In order to reduce the computing resources required, important advances have been made in efficient calculation of VOI metrics (Brennan and Kharroubi, 2007; Heath et al., 2018, 2017; Jalal and Alarid-Escudero, 2018; Menzies, 2016; Strong et al., 2015). However these methods facilitate the efficient calculation of VOI metrics given an economic model and so do not address the time and human resources required to build full economic models. Another barrier is familiarity with VOI methods which have been identified in a number of pilot studies (Bennette et al., 2016; Bindels et al., 2016; Claxton and Sculpher, 2006; Fleurence and Meltzer, 2013) Though economic modelling has become more widely understood and utilised in health care, decision makers remain relatively unfamiliar with VOI methods. This creates an important barrier to their uptake as results may be difficult to interpret and it is irresponsible to base decisions on methods which are poorly understood.

In one respect this lack of familiarity reflects a catch-22 in which VOI methods are not utilised because decision makers are unfamiliar with the methods and decision makers are unfamiliar with the methods because they are not utilised (Heller, 1961). However it is also relevant to note that much academic work on VOI has been focused on developing technical extensions rather than on communicating the basic insights of the approach and making the methods more practical for use within existing decision making structures.

In response to this collection of barriers Meltzer et al., (2011) have pioneered a "minimal modelling approach" in which VOI metrics can be calculated "without constructing a decision model of the disease and treatment process". This minimal modelling approach aims to surmount the barriers described earlier by creating simple economic models which are easier to understand and do not require large

human, time or computing resources to calculate VOI metrics. The authors provide methods to calculate VOI in two scenarios. First is the case in which data from a clinical study is available which directly characterises uncertainty in both costs and outcomes and is sufficient to inform a decision about the benefits of alternative interventions. This requires that the study includes all the relevant treatment alternatives, follows individuals up to the point of death or to full recovery and collects information on a comprehensive measure of benefit such as QALYs. In this case the value of future research can be calculated by bootstrapping the results of the clinical study. The second scenario is one in which the treatments of interest are expected to affect quality of life only with negligible effects on survival. In these cases, if quality of life is directly measured by a clinical trial, then a survival model can be built which will allows for VOI to be calculated. A limitation of these approaches is that prior clinical studies do not always exist, not least studies which are sufficient to inform a decision about the benefits of all relevant treatments. It can also be difficult to access the individual patient data required to carry out the analysis described by Meltzer and colleagues.

In providing support to a US based cancer research prioritisation group, Bennette et al., (2016) and Carlson et al., (2018) found that there did not exist any prior studies on which to base VOI analysis and so the Meltzer approach was not feasible in their context. In response, these authors developed a "hybrid between full decision analytic models and the conceptualization of ''minimal modelling'' by Meltzer and others." This involved developing individual Markov models for each research proposal. The aim of these models is to relate changes in the primary endpoint to estimated lifetime costs and benefits (measured in QALYs). As model structure was kept simple it was possible to develop the required models for each research proposal within the time constraints of the prioritisation process. This approach is very well suited to cancer research prioritisation as the high level disease process is relatively similar across different types of cancer. The limitation of this approach is that research prioritisation bodies typically receive proposals from a range of disease areas and so must prioritise research across disease areas. To apply the approach of Bennette, Carlson and others to this prioritisation task would require a tool box of "disease specific minimal models" to address research proposals arising from a diverse pool of pathologies.

An alternative approach to calculating VOI metrics to aid research prioritisation across disease areas has been developed by Claxton et al., (2015a, 2013) and McKenna et al., (2016). This method is based on a simple decision tree and places the primary outcome at the centre of analysis. This means that this approach is not specific to any particular disease area. The primary outcome reported in existing studies, or a proposed new study, usually captures the most important aspects of health outcome. Uncertainty in the primary outcome is used as a starting point in order to understand the health consequences of uncertainty, e.g., the distribution of values describing uncertainty about the relative effect of an intervention on a specific endpoint. Starting with a primary outcome does not mean that other outcomes are unimportant, it simply places the focus on a specific outcome of interest as a starting point in order to establish the value of reducing uncertainty in that outcome. The health benefits of research, and the value of implementing the findings of existing evidence are expressed as the number of events avoided for a harmful outcome (e.g., death) or gained for a benefit outcome (e.g., cure). In situations where there are a number of other important aspects of outcome that are not captured in the primary outcome (e.g., adverse events, quality of life impacts or resource implications), a minimum clinical difference (MCD) in effectiveness in the primary outcome may be specified in order to capture these additional considerations. For example, a larger MCD in effectiveness in the primary outcome may need to be detected in a new research study before there is confidence that health outcomes will be improved. This is analogous to the concept of an effect size, which has been central to the design of clinical research and determines the sample size used in most clinical trials. Where the primary endpoint of a study is not sufficient to capture all valuable aspects of outcome, external evidence can be used to link the endpoint to a comprehensive measure of outcome. This is analogous to the hybrid minimal modelling approach described by Bennette et al., (2016) and Carlson et al., (2018) where intermediate endpoints may be mapped to a meaningful comprehensive measure of outcome through simple extrapolation or modelling efforts. Because these methods are simpler and quicker to calculate than the others they are referred to as “rapid methods”.

Each of the approaches discussed attempt to simplify the task of creating decision models so that it is feasible to calculate VOI metrics within the time provided by

institutional constraints. In addition to reducing the expertise and time required, such simplification has the benefit of making the VOI methods easier to understand. However, simplification necessarily reduces the sophistication of the analysis. This induces the risk of over-simplification of complex clinical processes. Therefore, the extent to which the simplified approach adequately addresses the need for further evaluative research is an important consideration. This can be addressed by ensuring that the assumptions underpinning the analysis are made as explicit as possible and consideration is given to the likely impact that these assumptions might have on the findings.

1.2 CASE STUDIES

In this thesis we develop the approach of Claxton, McKenna and others to address the diversity of research proposals received by prioritisation bodies. To motivate the appropriate methodological developments this work has been carried out with input and case studies from NETSCC. This reflects the intention of this work to make both an academic contribution and to provide practical methods usable by decision makers within the current institutional constraints.

The NIHR is the main body that commissions and funds applied health and social care research in the UK. There are a number of NIHR funding streams, which include Programme Grants for Applied Research (PGfAR), Programme Development Grants, Research for Patient Benefit (RfPB), Health Technology Assessment (HTA), Efficacy and Mechanism Evaluation (EME), Health Services and Delivery Research (HS&DR), Public Health Research (PHR), Invention for Innovation (i4i), and NIHR Training Awards (NIHR, 2018a). Research prioritisation decisions are made in the vast majority of these programmes. However, for the purposes of this thesis, we focus on those made by the NETSCC HTA programme, which funds research that delivers information about the clinical and cost-effectiveness of developments in health care technologies (including drugs, devices, procedures, diagnosis and screening) and the impact of treatment and tests to NHS patients. There are two main workstreams within the HTA programme: i) the researcher-led workstream, which offers researchers the opportunity to submit proposals on topics or research questions within the programme's remit; and ii) the commissioned workstream, which invites applications in response to calls for

research on specific questions that have been identified and prioritised for their importance to the NHS and patients (NIHR HTA, 2018). Research proposals may include primary research, evidence synthesis, or feasibility and pilot studies.

The assessments required for prioritising research will be demonstrated using a set of six historical proposals that were considered as part of the NIHR HTA programme. NETSCC provided the set of six retrospective proposals with all confidential and personal information removed including costing information (only total costs of the proposal were provided). The six proposals were specifically chosen to demonstrate the diverse variety of types of studies and broad spectrum of issues that NETSCC is typically asked to commission, including: i) feasibility studies; ii) complex multi-arm adaptive trials; iii) disinvestment decisions associated with discontinuation of treatment; iv) inexpensive interventions; v) expensive trials; and vi) non-randomised safety trials. The six proposals are briefly summarised below:

- Proposal 1 (P1): Trial of early versus late treatment of prophylaxis to reduce venous thromboembolism following traumatic brain injury.
- Proposal 2 (P2): Trial discontinuation of a very high price medicine used for treating late stage cancer melanoma.
- Proposal 3 (P3): Non-randomised safety trial discontinuation of a very high price medicine for treating atypical haemolytic uraemic syndrome.
- Proposal 4 (P4): Complex multi-arm adaptive trial investigating treatments to modify the course of Alzheimer's disease.
- Proposal 5 (P5): Feasibility study investigating treatment for first episode psychosis in children and young people.
- Proposal 6 (P6): Trial of a low cost educational booklet, which aims to provide information to family carers of patients with cancer to facilitate death in their preferred location.

1.3 OVERVIEW OF THESIS

Chapter 2 introduces the place of VOI methods in research prioritisation in addition to outlining the rapid methods of research prioritisation which are the focus of this thesis. These methods were introduced by Claxton et al., (2015a) and are based around uncertainty in the primary outcome. The rapid method is illustrated for a

binary primary outcome (functional recovery) using a case study from one of six retrospective proposals provided by NETSCC. As no suitable studies exist on which to base the uncertainty in the primary outcome, a novel method for characterising uncertainty in the absence of available evidence is outlined. This provides an estimate of the value of research in terms of research cost required to gain an additional unit of a primary outcome (cost per additional unit of functional recovery). The results of applying this approach to the full set of six retrospective proposals are illustrated and the challenges of prioritising across research proposals without a common metric of value are discussed.

Chapter 3 provides an approach to addressing the challenge of prioritising across disease areas by linking primary outcomes to a generic measure of health outcome such as QALYs. This approach has been described in Claxton et al., 2013 and McKenna et al., (2016) and is demonstrated here by applying it to the case study from Chapter 2. This provides an estimate of the value of research in terms of research cost required to gain an additional QALY. The result of applying this method to the full set of six NETSCC proposals is a table of research proposals which can be ranked in terms of value for money. As research budgets are limited, the benefits of a particular research proposal must be compared to the benefits of other research proposals which could have been funded with these resources. In order to reflect these trade-offs appropriately, we apply the "bookshelf" approach described by Culyer (2016) and Remme et al., (2017) to the research prioritisation task. This allows us to rank research proposals from highest to lowest health impact and so identify the "best buys" for decision makers. It also provides a basis to understand whether health outcomes could be improved by expanding the research budget relative to the budget for general health expenditure. This approach to research prioritisation also shows the population health implications for charitable and industry contributions to research funding. In addition, this chapter also explores issues of using research to change clinical practice and how to use information on relative prices to determine an appropriate MCD.

Though the methods discussed in Chapters 2 and 3 can assist decision makers in addressing the key tasks of research prioritisation, they are directly applicable to only two of the six research proposals provided by NETSCC. Chapter 4 extends these methods to allow for analysis of binary primary outcomes when costs of

treatment depend on the primary outcome, continuous primary outcomes and survival primary outcomes. The analysis for each type of primary outcome is illustrated using a case study from the original six NETSCC proposals. Chapter 4 also develops and applies a novel method for estimating the value of information provided by a feasibility/pilot studies. These are a research design which involves carrying out a small initial study to determine whether a larger comparative effectiveness research project (a "full trial") is possible. This extension allows the expected health impact of proposals for feasibility/pilot studies to be compared directly to proposals for RCTs or other comparative effectiveness research which is essential if funding for feasibility/pilot studies and full trials come from the same research budget.

The methods described in Chapters 2 to 4 provide a means to rapidly estimate the value of research for a range of primary outcomes for both comparative effectiveness research and feasibility/pilot studies. The "rapid approach" to VOI analysis is to provide decision makers with models which are practical, built around primary outcomes and are quick to implement. The aim of Chapter 5 is to introduce a tool which has been designed to reduce the technical barriers to implementation of VOI methods. This tool is called Rapid Assessment of Need for Evidence (RANE) is an important contribution of this PhD. RANE is open source, hosted by the University of York and is freely available for use at https://shiny.york.ac.uk/rane/. The RANE tool embeds the methods described in Chapters 2 to 4 and so allows users to quickly carry out VOI calculations to help inform research prioritisation decisions without having to code new models for each research proposal. This is vitally important as reducing the time and technical barriers to VOI analysis can facilitate its use more widely in the health system thus improving the transparency and accountability of research decision making. Chapter 5 reviews the software currently available for research prioritisation, provide an overview of the RANE tool and its capabilities and provides a step by step illustration of how to use the tool using a NETSCC proposal as an example.

Chapter 6 extends beyond a HTA research funding panel setting (NETSCC) into a more comprehensive HTA decision making context in which approval and research decisions are made simultaneously. This chapter address the question of how to make early access decisions without a full economic model. There is pressure on

decision makers to allow early access to medications and devices with high prices when the evidence base is highly uncertain. The benefit of granting early access to new treatments which appear to be effective based on current evidence is that potentially worthwhile treatments can be quickly provided to patients without undue delay (Claxton et al., 2016; Eckermann and Willan, 2008a). However, the cost of granting early access is that it can reduce or remove the possibility of further research (Griffin et al., 2011). This implies a trade-off between expected health benefits for current patients from early access and health benefits to future patients from further research. The literature on conditional coverage of health technologies provides a coherent and transparent basis to trade-off price, uncertainty and effect size when making early access decisions and so can provide a basis to link evidence to pharmaceutical pricing (Claxton et al., 2008; Rothery et al., 2017). Conditional coverage recognises that decision makers can make not only approve or reject decisions but also have options such as "Only in Research" (OIR) and "Approval with Research" (AWR). The former only allows the use of new treatment in a research setting. The latter approves the treatment for widespread use on the condition that additional evidence is collected (Claxton et al., 2012; McKenna et al., 2015; Walker et al., 2012). Currently the literature on conditional coverage assumes that decision makers have access to a full economic model. In this chapter we provide a framework for early access decisions which is built upon the rapid methods described in Chapters 2, 3 and 4. The approach is illustrated for a binary primary outcome using a case study from the set of six NETSCC proposals introduced in Chapter 2.

Chapter 7 concludes by providing a brief overview of the entire thesis and identifies avenues for future research and policy.

Chapter 2

Rapid assessment of need for evidence: estimating the value of research in terms of cost per primary outcome

2.1 INTRODUCTION

Research is central to the functioning of modern medicine but resources to fund it are limited. Therefore choices must be made about which research projects should be funded and which should not. Research prioritisation is the practice of choosing to fund certain research proposal at the expense of not funding others.

There are a number of bodies worldwide which are charged with deciding which research projects to fund given limited budgets for research. These include national agencies such as NETSCC in the UK and the PCORI in the US, and international bodies such as the Bill and Melinda Gates foundation (2018) and the European and developing countries clinical trials partnership (2018). Value of information (VOI) analysis is a method which can potentially add to the transparency and accountability of the research prioritisation process by providing quantitative estimates of the value of different research proposals. The value of research is understood here as the health gain from reducing uncertainty in health care decision making. There is greater value in resolving the uncertainty in some clinical decisions rather than others as some decisions are more uncertain and have greater health consequences. These methods for calculating the value of research make use of the available evidence and will be described in detail in the next section.

Historically VOI analysis has required building a full health economic model which reported results in terms of costs and quality adjusted life years (QALYs). A health economic model is a mathematical model which draws on a range of data sources to estimate the costs and health effects associated with an intervention over an appropriate time horizon (Briggs et al., 2006). Constructing such models typically requires a large amount of time and a range of expertise including clinical advisors and experts in decision modelling and evidence synthesis. For resource constrained resource prioritisation bodies this is an important barrier to their use. In response

Claxton et al., (2015a) have developed "rapid" methods to calculate VOI which are based around uncertainty in the most important clinical endpoint (the primary outcome). These methods can be applied to binary primary outcomes (i.e., an outcome which either occurs or it does not occur) to calculate the value of reducing uncertainty in the relative effect and the baseline probability of the primary outcome. They do not require a full economic model and so can provide a practical approach for decision makers to inform their decision making process. In this chapter we apply the VOI methods to estimate the benefit of research developed in Claxton et al., (2015a) to a retrospective set of proposals received by NETSCC to understand their implications for applied research prioritisation.

First we outline the current processes used in identifying high priority research. Second, we outline the most important determinants of the health impact of research and how VOI takes account of these. Third we describe the need for rapid approaches to calculate VOI and the evidence required to carry out this analysis. Fourth, we apply the rapid approach to a case study from the set of 6 retrospective NETSCC proposals. As part of this we outline a method to characterise uncertainty in the absence of previous studies. Finally, we explore the implications of this approach when prioritising across a set of research proposals.

2.2 IDENTIFYING HIGH PRIORITY RESEARCH

Research prioritisation involves developing a consensus on a number of priority areas to address key questions which need to be underpinned by future investment in research. The task involves assessing the value of the full range of research topics so that each topic can be compared against each other, or considered against numerous criteria. The resulting priorities are highly dependent upon the value of addressing the research question and the consensus opinion regarding how the research can improve the health of the population (patients, the health system and/or general public). Research funders have limited resources and the prioritised schedule of research must be traded against the available funding in order to identify high-priority areas for further research.

Figure 2.1 illustrates a typical research prioritisation process. The first step is topic generation, where topic ideas may come from the medical community, patients, academics, other stakeholders (such as the James Lind alliance), or research projects

identifying a future need for research in a particular area. Research ideas are then filtered and topics are generated for further consideration. Once topics are selected based on those that are expected to be important to patients, clinical practice, policy or decision makers, and where study outcomes could potentially lead to direct benefits to patients and/or the wider population, the topics are developed further. This may be via a commissioned workstream, where applicants are invited to respond to calls for research on specific questions, which have been identified and prioritised for their importance to the health system and patients, or via a researcher-led workstream, where researchers have the opportunity to submit proposals on the topic or research question.

Once topics are selected, the second step to research prioritisation involves identifying and articulating, as far as possible, why the research question is important to patients and clinical practice in terms of improving health of the population. Health outcomes can be improved by either, conducting research to reduce uncertainties in the existing evidence base in terms of which interventions to use for the treatment, prevention or diagnosis of disease, or by implementing the findings of existing evidence about the best intervention to use. Therefore, the most important starting point is an understanding of how the existing literature or evidence supports the research question. New research should never be undertaken without knowledge of the existing evidence base, on the grounds that it would be considered unethical to enrol patients into a research study without prior knowledge of the effects of the interventions to be evaluated, e.g., enrolment into a RCT means that a certain proportion of patients are likely to be allocated to a suboptimal intervention (Chalmers and Nylenna, 2014; Ioannidis et al., 2014). The principle of equipoise implies that there must be genuine uncertainty over whether an intervention will be beneficial. Therefore, participants of a research study should never knowingly be offered less than the best intervention for their condition. Despite this basic requirement, however, Cooper et al., (2005) found that only 46% of a sample of 24 responding authors of trial reports included in a Cochrane review were aware of the relevant existing reviews at the time when they were designing their new studies. The NIHR's Health Technology Assessment Programme routinely requires research proposals to include a summary of why the research is important in terms of

improving the health of the population and how the existing literature supports the proposal.

Once research questions requiring prioritisation are identified, the third step involves prioritising the proposals over other topics that could be commissioned with the same resources. The approach used for this purpose varies by different decision making bodies. One common approach is to attach an overall score to research proposals based on meeting certain criteria under different categories. For example, the NIHR Invention for innovation (i4i) panel in the UK attaches an overall score to proposals based on an assessment of: the expertise and track record of the applicants; the importance of the research area; the expected impact on NHS practice; and the quality of the project plan, including value for money. This score is based on a scale from 0 to 10, with 0 representing research that has poor or little merit compared with other research; 5 representing research that is comparable with other research; and 10 representing research that has exceptional benefit compared with other research (NIHR, 2015). Similar overall scores are also used in other NIHR programs such as Research for patient benefit (RfPB) and Efficacy and mechanism evaluation (EME) (NIHR, 2019a, 2019b). A second categorisation approach involves the use of a "traffic light" system of colours to indicate the status of each research proposal relative to the others. For example, red may indicate fair existing knowledge on a topic and a ranking of lower priority compared to other research; amber may indicate existing knowledge could improve and research is comparable with other topics; while green may indicate a knowledge gap exists and further research is definitely warranted compared with other research.

A third approach is to explicitly assess the 'value' of each research proposal using a specific metric of value, which can be used to directly compare the research proposals to each other. For example, the metric of value may be net QALYs gained representing the net health impact of an intervention on population health.[4] This third

[4] Quality-adjusted life years (QALYs) gained is a generic measure of disease burden, which incorporates the impact of an intervention relative to a comparator in terms of both the quality and quantity of life lived. The cost per QALY gained provides a useful summary measure of how much additional resource is required to achieve the measured improvement in health (quality and quantity of life lived) of the intervention relative to the comparator. The *net* impact of the intervention on health outcomes overall is judged relative to the likely health opportunity costs, which is the health that is forgone elsewhere in other health care programmes by diverting resources to the intervention, rather than other uses. The value of a research proposal can be expressed as the scale of the potential net

approach explicitly quantifies the value of the additional information generated by each of the proposed research projects competing for the same limited resources. By the VOI associated with each proposal, it is possible to identify those which offer the greatest return in terms of net health impact after taking account of the costs of research. When the value of all research proposals are assessed using the same metric of value, a straightforward exercise can be used to rank the proposals in order of value to determine the areas of highest priority.

A combination of 'scoring' approaches may also be pursued in order to prioritise between the research proposals. Importantly, the process of research prioritisation should aim to be as transparent as possible so that there is clear justification and rationale for the funding of research. Once research proposals are ranked in order of priority, the final step involves selecting the proposals for funding based on the available research budget. The total budget should be focused on the high-priority areas as identified by the ranking process, with any remaining budget left to pursue other policy focused evidence-based research.

The process by which research is identified as high-priority usually involves extensive and structured stakeholder engagement in order to secure both the best possible evidence base for the task, as well as identifying the "best buys" from the available research budget. These deliberations have a direct bearing on the topics that are selected as priority areas and, consequently, the manner in which the research is conducted. The selected priority areas should be reviewed on a regular basis in order to ensure that the proposed research delivers on the promised question and is of continued relevance. This will help ensure that resources are not wasted, which could be used elsewhere to identify and address new research opportunities.

Establishing the potential benefits of new research to help inform its priority level requires a number of considerations; some of which represent value judgements, while others represent scientific beliefs about the evidence to date. Value judgements are made when more weight is placed on a particular outcome compared with another, while scientific beliefs are reasonably held views about a particular state of the world and the degree of uncertainty or knowledge about it. No quantitative

---

health impact of the intervention on health outcomes overall, i.e., the difference between QALYs gained by the intervention and QALYs that could have been gained elsewhere with the additional resources which are required to fully implement the intervention.

analysis or scoring approach that establishes the value of a research proposal, no matter how well it is conducted, will be sufficient to capture all aspects of scientific and social value relevant to making decisions about research priorities; largely because these aspects can be readily disputed. The more relevant question is whether they offer a practical and useful starting point for deliberation so that a consensus opinion can be reached, and whether they add to the transparency and accountability of the decision making process (Claxton et al., 2013).

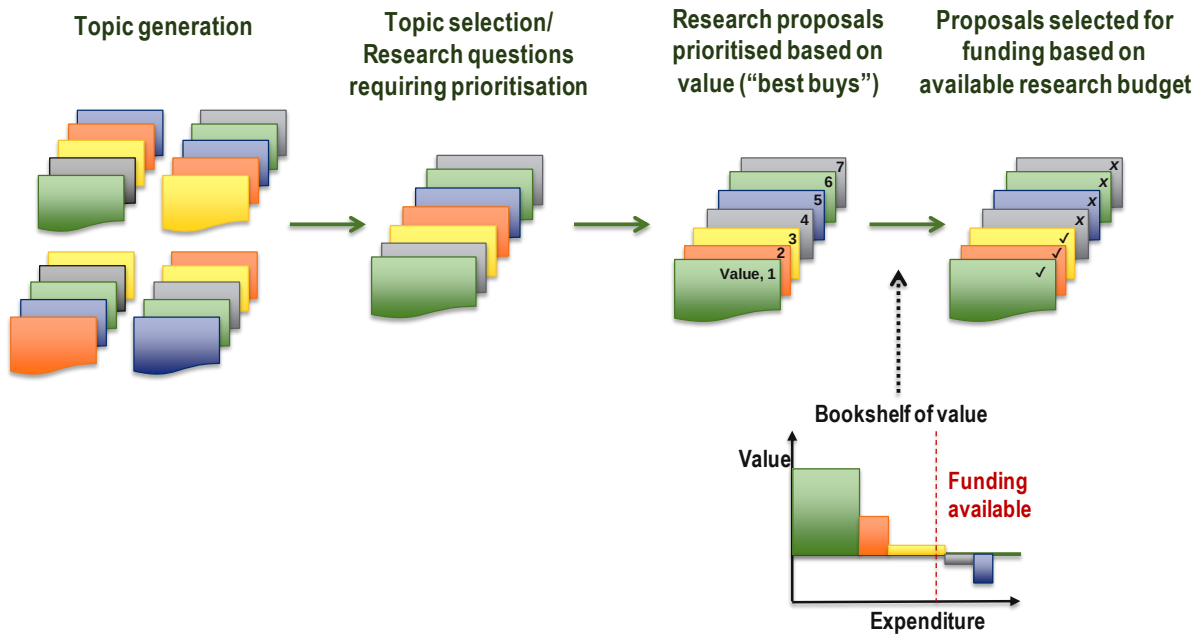


Figure 2.1: Illustration of a research prioritisation process in which research projects are chosen based on their value to the health system.

2.2.1 THE HEALTH BENEFITS OF RESEARCH

Research is important for improving health outcomes since it can resolve existing uncertainty about the effectiveness of the interventions available for the treatment, prevention or diagnosis of disease. Resolving uncertainty helps inform decisions about the optimal intervention for subsequent patients. For example, on the balance of existing evidence, a particular intervention may be judged to be the most effective option; however, due to uncertainty in the evidence base, there will be a chance that the other alternative interventions used for the same condition are in fact more effective. Therefore, when the existing evidence is uncertain there is always a chance

that one of the alternative interventions could improve health outcomes to a greater extent than the intervention judged to be the most effective on average. This means that there are adverse health consequences associated with uncertainty. The importance of this uncertainty is indicated by the scale of these health consequences. The scale is dependent on the likelihood that a particular intervention is not the most effective option, how much less effective it is likely to be (in terms of some measure of health outcome), and the size of the patient population facing the uncertain intervention choice.

A judgement about the level of uncertainty in the existing evidence base can come from a systematic review of what is already known or being researched about a particular topic: a meta-analysis or statistical modelling approach that combines the results from multiple studies in an effort to increase power over individual studies and improve estimates of the size of effect (Chalmers and Nylenna, 2014; Higgins and Green, 2011); expert elicitation where relevant experts are asked to provide their judgement regarding the magnitude of effect size (O'Hagan et al., 2006; Soares et al., 2011); meta-epidemiological studies that adopt a systematic review or meta-analysis approach to examine the impact of study design characteristics on effect size (Bae, 2014; Rhodes et al., 2015); or a combination of these sources.

The level of uncertainty in the decision arises from the range of plausible values that the outcome of interest can take. This is usually represented by the confidence interval (CI)[5], or standard error (SE), around the mean or median estimate of effect. A wide CI implies a large amount of uncertainty; however, it is only when the CI crosses the line of no difference between the alternative interventions that this uncertainty creates the potential for adverse health consequences, i.e., only the consequences of uncertainty that will change the decision are important. For example, uncertainty about the estimate of treatment effect only matters in so far as it influences the decision; if the decision is the same for all plausible values of the treatment effect then the uncertainty is unimportant.

As an example, consider the evidence on the use of corticosteroids following traumatic brain injury (TBI) before the large definitive trial of CRASH (CRASH trial

[5] Technically, a Bayesian interpretation required to estimate the value of research as described in this report, therefore this may considered a credible interval (CI).

collaborators, 2005). Before CRASH, a meta-analysis of 19 RCTs indicated that the effects of corticosteroids (CS), compared with not using them, on death and disability were unclear. The odds ratio for death was 0.93 in favour of the use of CS, but with a 95% CI crossing the line of no difference (odds ratio equal to one) from 0.71 to 1.18 (see bottom panel of Figure 2.2) (Claxton et al., 2013). This uncertainty means that every decision about the use of CS following TBI is associated with a chance that it may not have be the most effective treatment choice. Based on a Bayesian interpretation of the CI, there was a 75% chance that CS were effective and improved survival; however, there was a 25% chance that CS resulted in excess deaths per annum (see top panel of Figure 2.2). This uncertainty can be translated into the consequences for patient outcomes in number of expected deaths per annum, by combining the uncertain estimate of relative effect with an estimate of the baseline risk (derived from either the control arms of the trials, or from an external source on baseline risk relevant to the population of interest) and multiplying by the incidence of TBI per year (8,800 individuals). In this case, the expected (average) number of deaths per annum due to uncertainty in the use of CS following TBI was 39 additional deaths per year and is represented by the grey section in Figure 2.2. This estimate of the consequences of uncertainty is derived from the fact that there is a low probability of a large increase in deaths with CS (say, greater than 500), compared to a larger probability of smaller increases in deaths (say, below 100). The average over these consequences gives the number of deaths per annum due to uncertainty in the use of CS.

The expected health consequences can be interpreted as an estimate of the health benefits that could be gained each year if the uncertainty surrounding treatment choice were resolved, i.e., it provides an expected upper bound on the health benefits of further research, which would confirm whether CS following TBI increases or reduces the number of deaths per annum. These potential expected benefits increase with the size of the patient population whose treatment choice can be informed by additional evidence and the time over which evidence about the effectiveness of the interventions is expected to be useful.

The health benefits that can be gained through research are called the “information value”. The information value will vary between research proposals as some intervention decisions will be associated with large uncertainty and large health

consequences, while others may have large uncertainty but with relatively modest consequences (e.g., wide CI but with less decision uncertainty). Some decisions will be associated with modest uncertainty but with very important health consequences, while others will have small uncertainty with modest consequences. The value of information for a particular decision will depend on both the level of uncertainty and the consequences of this uncertainty. By quantifying the value of conducting further research in this way, the potential information value of a particular research proposal can be compared to the value of other research proposals competing for the same resources.

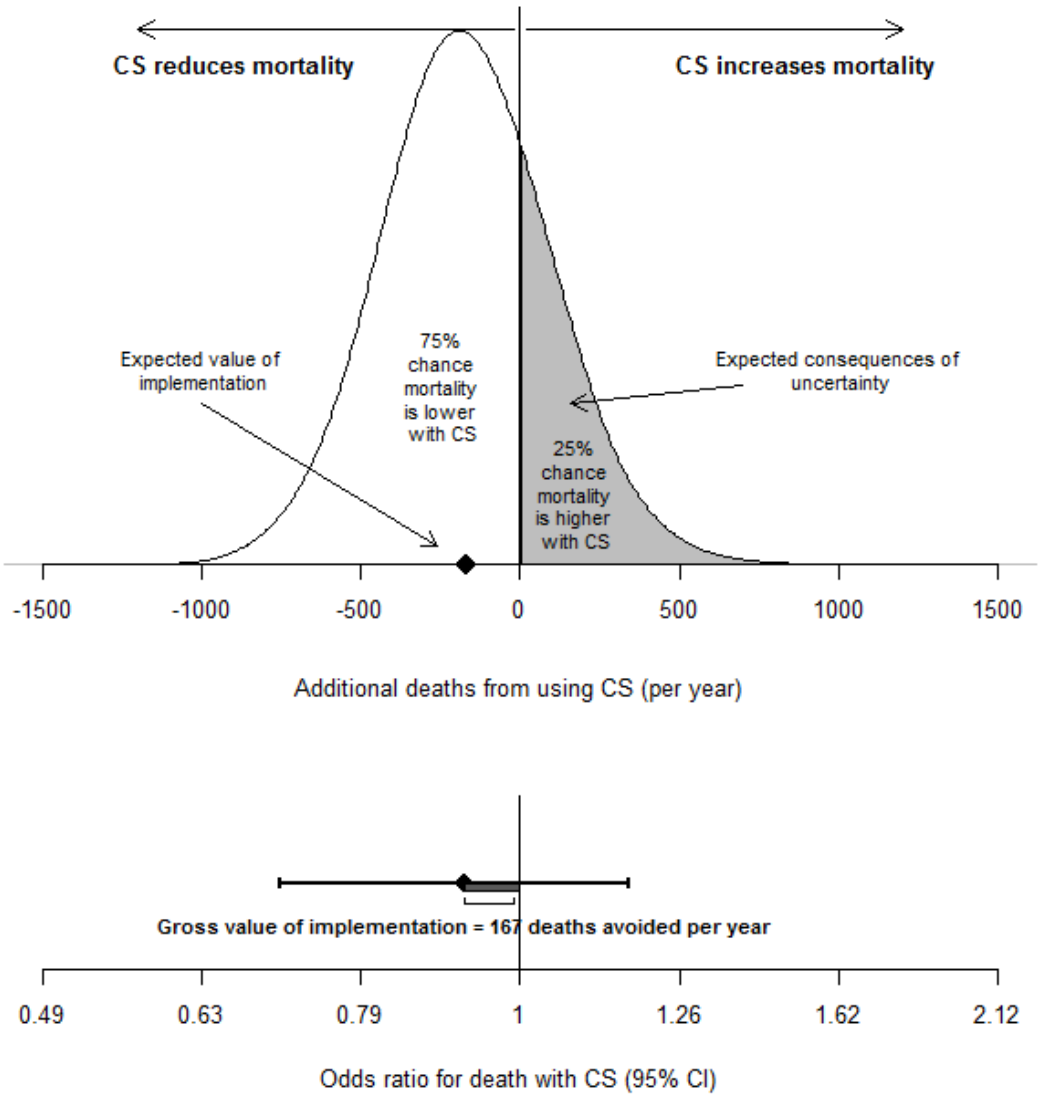


Figure 2.2: Uncertainty in the outcome of mortality from the use of CS following traumatic head injury. The upper panel shows the range of plausible outcomes associated with the use of CS in terms of deaths. The lower panel illustrates the 95% CI which generates the upper panel. The impact of implementing CS following traumatic head injury based on existing evidence is illustrated on each panel. CS, corticosteroids. CI, confidence / credible interval.

2.2.2 THE HEALTH BENEFITS OF IMPLEMENTING THE FINDINGS OF EXISTING RESEARCH

Funding research is not the only way to improve health outcomes. It is also possible to improve health outcomes by ensuring that the intervention option that is expected to be best based on the findings of existing evidence is implemented into clinical practice. In fact, the improvements in health outcomes from implementing the

findings of existing evidence, which is known as the "implementation value", may be greater than the potential improvements in health outcomes through conducting further research (Claxton et al., 2013).

Drawing a distinction between information value and implementation value is important because conducting further evaluative research is not the only way to change clinical practice or improve health outcomes. The results of a new research study may influence clinical practice and may contribute to the implementation of research findings, but this is not the only, or the most effective way to do so. There are other mechanisms (e.g., more effective dissemination of existing evidence) and policies which are within the remit of other bodies (e.g., incentives and sanctions) to affect implementation. Therefore, conducting research to influence implementation, rather than because there is real value in acquiring additional evidence itself, may be inappropriate as there is limited research capacity and this could be used elsewhere to conduct research in areas where there are genuine uncertainties. Conducting research can also have negative health effects for those patients enrolled in research and allocated to interventions that are expected to be less effective (e.g., patients enrolled to the suboptimal arm of a clinical trial).

Figure 2.2 illustrates the implementation value in the outcome of mortality from the use of CS following TBI. On the balance of existing evidence, the odds ratio of 0.93 favours the use of CS. Treating all patients with CS following TBI compared to usual care is expected to be worth 168 deaths avoided per annum. However, prior to CRASH, approximately 12% of patients receiving CS following TBI in the UK, while 88% did not receive CS (McKenna et al., 2016). If clinical practice were to switch fully to CS to reflect the balance of existing evidence, we would expect to observe (167 x 88% =) 147 fewer deaths per annum. This represents the expected value of implementation efforts to change clinical practice, which is the difference between the expected value of a decision based on existing evidence that is fully implemented and the expected value of a decision with implementation at its current level (i.e., 12% in the case of CS following TBI). Implementation efforts can be difficult and costly to reverse if the results of subsequent research find that the intervention is not as effective as the previous evidence suggested (Claxton et al., 2016; Eckermann and Willan, 2008a). Therefore, in some circumstances (particularly, if there is a complete absence of evidence) it may be better to delay the

use of the intervention until additional research is undertaken. This may involve collecting a small amount of information to inform beliefs where there is no existing evidence before committing to changing practice or funding a large clinical trial.

2.2.3 THE TIME HORIZON FOR RESEARCH DECISIONS

The information generated by research will not be valuable indefinitely because future changes are expected to occur over time that impact on the value of information. For example, new and more effective interventions may become available, making the current intervention and comparators obsolete and possibly rendering information about their effectiveness irrelevant to future clinical practice. Research may also fundamentally change our understanding of disease processes, e.g., the mechanisms that cause resistance to antibiotics, thereby impacting on the future value of the information generated by research that is commissioned today. Furthermore, other evaluative research may already have been commissioned by other bodies or health systems, which may resolve much of the uncertainty. Therefore, the actual time horizon for evidence generated by research will depend on the anticipated shelf-life of the alternative interventions and expected future changes over time.

The actual time horizon for research decisions is unknown since it is a proxy for a complex and uncertain process of future changes (Philips et al., 2008). However, some judgement about the time horizon is required in order to make decisions about research priorities. This assessment is possible based on historical evidence and judgements (e.g., beliefs elicited from experts) about whether a particular area of research is likely to experience these changes. Information about clinical trials or studies that are already planned or underway may be obtained from various trial registries, applications for marketing authorisation, and funding bodies. Where there are limited or no data available, sensitivity or scenario analyses should be undertaken to highlight the extent to which the value of information is influenced by the time horizon. The health benefits of research should also be discounted over this time horizon so that more weight is given to decisions that are informed by the research in the near term and less weight given to decisions in the more distant future.

2.2.4 SIZE OF THE POPULATION THAT CAN BENEFIT FROM RESEARCH

The health benefits of research depend on the size of the population that can benefit from the new information. The size of the beneficiary population is typically derived using epidemiology data based on an understanding of the decision making context and the scale of the health care decisions that are likely to be affected by the new research. Those who could potentially benefit from new information include the prevalent cohort with the disease in question and/or the future incident cohorts over the appropriate time horizon for the research decision.

The health benefits of research will not be realised until the study is completed and the results become available. Therefore, the eligible population, based on prevalence and/or incidence, is usually adjusted to reflect the time it will take for the study to complete. If treatment decisions cannot be reversed, then it is only those patients incident after the research reports that will realise any of the potential benefits. However, some study participants who are enrolled in the optimal intervention arm will receive the benefits of the best intervention while the study is conducted (McKenna and Claxton, 2011).

The size of the beneficiary population that can benefit from research also depends on whether decisions are made at a local, national or international level, and the extent to which information is valuable across jurisdictions and populations (Eckermann and Willan, 2009; Woods et al., 2018). When decisions are made at a local level, it is usually with reference to the health benefits for a specific local population subject to resources available in the local setting. However, information generated by publicly funded research is a public good. Therefore, in some instances, information is more generalisable and the health benefits from a local research activity may be realised over a much broader population, e.g., at national or international levels, or wider risk group of individuals. In this case, the global value of research may be assessed (Eckermann and Willan, 2009). In other circumstances, research may offer the potential to inform multiple decisions, e.g., improved surveillance data may inform investment decisions across a range of prevention and treatment decisions, which could affect multiple populations.

2.2.5 INFORMING RESEARCH PRIORITIES

The health benefits of research provide an estimate of the extent to which research can potentially reduce the occurrence, and expected consequences, of uncertainty in the decision between alternative interventions. Acquiring information through research, however, can be costly. Therefore, the expected benefits of research must be compared to the costs of conducting the research. The costs of research include the fixed and variable costs of conducting the research activity and the costs that fall on the health care system, e.g., NHS support costs associated with implementing the research activity. These costs can be in terms of resources (use of more expensive treatments in research) or in terms of health (enrolling patients to a treatment which is expected to be suboptimal). The expected benefits of research must be greater than the costs of the research for the activity to be considered as potentially worthwhile; it is only when the costs of the research exceed the expected benefits that research can be completely ruled out. This criterion is used as a first hurdle to identify research that is potentially worthwhile.

The costs of research should include a consideration of the health opportunity costs associated with the research expenditure. This means that the health gain from the research should be compared to the health that could have been gained elsewhere by making the resources available for other health care activities (i.e., the health opportunity costs is the health that is forgone elsewhere by resources not being available for other activities because they are accommodating the costs of the research activity). Therefore, the expected *net* impact of the research activity on health outcomes is judged relative to the likely health opportunity costs. Recent empirical work in the UK has estimated the relationship between changes in NHS expenditure and health outcomes (Claxton et al., 2015b; Lomas et al., 2018). This work suggests that the NHS spends approximately £15,000 to gain one QALY and £100,000 to avoid one death. Using these estimates, if the proposed research costs £2 million, this means that the cost of the research could have been used to gain approximately 133 QALYs or avoid 20 deaths elsewhere in the NHS. If these opportunity costs of research are substantially less than the expected benefits of the research then it would suggest that the proposed research is potentially worthwhile to the NHS.

Once it is established that the expected health benefits of research are sufficient to regard a particular research proposal as potentially worthwhile, the second important question is whether the research should be prioritised over other topics (or research proposals) that could be commissioned with the same resources. Insofar as there is funding dedicated to research activities, e.g., a dedicated research funding pot held by NETSCC specifically for new research to support NHS decision making, the opportunity costs incurred by the research funder is the funding (and associated health benefits) that are diverted away from other types of research in order to fund the specific activity. Most research funding bodies have limited resources and, therefore, it is likely that not all potentially worthwhile research projects can be commissioned. In this case, the benefits of some research projects must be foregone in order to commission others. Therefore, estimates of the costs and potential health benefits of research projects competing for limited resources can be helpful to inform the priority level of each proposal. If a research funder is concerned with the total health of the population served by the research budget, then they can identify the research that generates the most health per expenditure and order proposals in order of value for money. The research funder works down the ordered list, funding all proposals until the research resources run out. It the research funding available to decision makers runs out before all worthwhile research projects have been funded then this provides a case for increasing research funding.

It is important to highlight that research prioritisation decisions require an assessment of the expected health benefits of research before the actual results of the research that will be reported in the future are known. Therefore, it might seem intuitive to look back at the historical proposals and ask whether a particular research prioritisation decision was correct based on the results of the research. However, this use of hindsight is inappropriate because the findings of the research represent only one realisation of the uncertainty that could have been found when the decision to prioritise and commission research was taken. For example, the expected health benefits of conducting a trial on the use of CS following TBI based on the evidence prior to CRASH was estimated to be 1,375 deaths averted over the 15 year time horizon. With an expected research cost of £2.2 million the value of CRASH was estimated to be (£2.2 million/1,375 =) £1,600 per death averted (McKenna et al., 2016). Given that the NHS spends around £100,000 to avert one death (Claxton et

al., 2015b), this suggests that the CRASH trial was worthwhile. As it turns out, the trial was worthwhile to avoid unnecessary deaths with a definitive finding that CS increase the risk of death following TBI (CRASH Collaborators, 2005). However, it would be inappropriate to say whether the Medical Research Council that funded the CRASH trial made the 'right' decision to commission CRASH in the year 2000 only because it showed a surprising and consequential result. This is because i) the value of the other research proposals which were on the table for consideration on the day that CRASH was funded are unknown; ii) the actual findings of the CRASH trial represent only one realisation of the uncertainty that could have been found when the decision to prioritise and commission research was taken. Therefore, it is important to evaluate the quality of decisions based on the information that was available when the decision was made i.e. without the benefit of hindsight[6].

[6] It is also important to bear in mind the distinction between implementation value and information value discussed in Section 2.2.2. The effect of the results of CRASH on motivating changes in practice should distinguished from the uncertainty that it resolved, especially when research funding is limited and there are other mechanisms for changing practice.

2.3 RAPID ASSESSMENT OF THE NEED FOR EVIDENCE

2.3.1 THE NEED FOR A RAPID APPROACH

The health benefits of research have traditionally been established within a net benefit framework used to assess uncertainty surrounding a decision to adopt or reimburse a health technology into the health care system. This typically requires the construction of a decision-analytic model, which brings together relevant evidence on short and long-term costs and health outcomes for the intervention and comparators under consideration, and facilitates the synthesis of data from a variety of sources in order to assess the cost-effectiveness of the intervention and the need for further evaluative research. However, institutions with a responsibility for making research prioritisation and commissioning decisions are often restricted by the time and resources required to generate decision models, making traditional modelling efforts unsuitable for integration into the research prioritisation process. This partly arises as a consequence of the fact that those institutions with the remit for making reimbursement decisions are often separated from those responsible for prioritising and commissioning research.

The need for a practical and feasible method within the time and resource constraints of a deliberative process of research prioritisation has called for the development of 'rapid' or minimal modelling approaches. These approaches allow for rapid estimation of the health benefits of research without the need for constructing a full disease and/or decision-analytic model. This involves simplifying or omitting components of the full modelling approach in order to produce information on the value of research in a timely manner. A rapid approach can offer a quick and practical means for estimating the health benefits of research in a matter of days, rather than weeks or months. The approach may also be viewed as offering a transparent and efficient method for setting research priorities and may provide a workable interface, whereby analysts and stakeholders could potentially validate key inputs and assumptions about the existing evidence base in real time as part of the deliberative process (see Chapter 5 for user friendly rapid VOI tool). The approach may also be relevant to different types of health care systems and decision making contexts, including those that do not explicitly include economic considerations in their decision making process. For example, for those institutions, the expected

health benefits of research may be assessed using a metric of value that is based on health outcomes alone rather than economic considerations.

2.3.2 RAPID APPROACHES FOR ESTIMATING THE HEALTH BENEFITS OF RESEARCH

Minimal modelling has been proposed in the literature as a method for rapid estimation of the value of research (Meltzer et al., 2011). This can be performed if a prior clinical study is available that directly characterises uncertainty in comprehensive measures of health outcome (e.g., both costs and QALYs) and is sufficient to inform a decision about the benefits of alternative interventions (Meltzer et al., 2011). This may be achieved in studies that follow patients up to the point of death or to full recovery, while recording all relevant outcomes between the interventions such that a comprehensive measure of benefit can be assessed. The drawback to this approach is that in many cases the required clinical studies do not exist and so this analysis cannot be carried out. Extending the minimal modelling approach Bennette et al., (2016) propose a method to map progression free survival to a meaningful comprehensive measure of outcome by constructing a simple economic model with relatively few parameters for each research proposal. This approach has been applied successfully in cancer research prioritisation in the USA (Bennette et al., 2016; Carlson et al., 2018). A limitation of this approach is its specificity to oncology and the time and expertise required to construct a customised decision model for each research proposal.

The rapid approach for estimating the health benefits of research that is proposed in this thesis places the focus on a primary outcome of interest. The primary outcome reported in existing studies, or a proposed new study, usually captures the most important aspects of health outcome. Uncertainty in the primary outcome is used as a starting point in order to understand the health consequences of uncertainty, e.g., the distribution of values describing uncertainty about the relative effect of an intervention on a specific endpoint. Starting with a primary outcome does not mean that other outcomes are unimportant, it simply places the focus on a specific outcome of interest as a starting point in order to establish the value of reducing uncertainty in that outcome. For example, mortality was taken as the primary outcome to understand the health consequences of uncertainty in the evidence available before

the CRASH trial. In this case, the expected health benefits of research were expressed in terms of number of deaths averted for the outcome of mortality. The health benefits of research, and the value of implementing the findings of existing evidence, can be expressed as the number of events avoided for a harmful outcome (e.g., death) or gained for a benefit outcome (e.g., cure).

In situations where there are a number of other important aspects of outcome that are not captured in the primary outcome (e.g., adverse events, quality of life impacts or resource implications), a minimum clinical difference (MCD) in effectiveness in the primary outcome may be specified in order to capture these additional considerations. For example, a larger MCD in effectiveness in the primary outcome may need to be detected in a new research study before there is confidence that health outcomes will be improved. This is analogous to the concept of an effect size, which has been central to the design of clinical research and determines the sample size used in most clinical trials. The required effect size does not represent what is expected to be found by the research, but instead it represents the improvement in the primary outcome that would need to be detected for the new treatment to be considered worthwhile and to have an impact on clinical practice. For the example of CS for use in TBI, if CS is more expensive than the current standard of care, then a MCD of 2% may be required. This implies that the probability of death must decrease by at least 2% for the new treatment to be worthwhile relative to the current standard of care.

Specifying a MCD is one way to implicitly account for the other aspects of outcome that are not captured in the primary outcome. This may be used as part of the deliberative process to assess whether the proposed research is a priority at a MCD that is regarded as sufficient to account for these other aspects of outcome. Where the primary endpoint of a study is not sufficient to capture all valuable aspects of outcome, external evidence can be used to link the endpoint to a comprehensive measure of outcome[7]. This is analogous to the minimal modelling approach described above where the primary endpoint is mapped to a meaningful comprehensive measure of outcome through simple extrapolation or modelling efforts. The translation to a comprehensive and comparable measure of health

[7] Section 3.3.4 describes a method to inform the appropriate size of the MCD with reference to a comprehensive measure of health outcome, such as QALYs.

outcome such as QALYs enables the health benefits of research to be compared directly across diverse clinical areas. This helps to address the difficult, but unavoidable question, in research prioritisation about how to estimate and compare the health benefits of research across diverse disease areas (see Chapter 3).

The notable limitation associated with the approach described above, or any of the minimal modelling approaches is the possible over-simplification of complex clinical processes. Therefore, the extent to which the simplified approach adequately addresses the need for further evaluative research is an important consideration. This can be overcome by ensuring that the assumptions underpinning the analysis are made as explicit as possible and consideration is given to the likely impact that these assumptions might have on the findings.

2.3.3 THE MINIMUM EVIDENCE REQUIRED TO ESTABLISH THE HEALTH BENEFITS OF RESEARCH

The minimum evidence required to conduct a rapid assessment of the need for research are described below for a binary outcome measure[8].

Primary outcome measure

The primary outcome measure or endpoint captures the most important aspect of health outcome. The health benefits of research are expressed in terms of 'benefits gained' or 'harms avoided' depending on whether the outcome is a benefit or harm. Alternative endpoints can also be used to consider the impact of additional evidence on different aspects of outcome. Where the primary outcome is not sufficient to capture all valuable aspects of outcome, a MCD in the primary outcome may be specified in order to implicitly account for these other unquantified aspects of outcome and/or costs.

Relative effectiveness

An estimate of the relative effectiveness of the intervention is required for the primary outcome, along with an estimate of its uncertainty. This is usually expressed in terms of an odds ratio or relative risk, with a 95% CI (or SE) representing the

[8] The methods for continuous or survival primary outcomes have slightly different evidence requirements. These outcomes are discussed in detail in the Chapter 4.

range of plausible values that the quantity can take. Importantly, some judgement about the uncertainty in this estimate based on what is already known must be made in order to determine whether additional evidence is required. This judgement can come from a systematic review and standard meta-analysis of the available existing evidence or from alternative sources such as expert elicitation or meta-epidemiological studies (see Section 2.4.4). If an estimate is unavailable or considered inadequate, alternative values can be used to represent different judgements about the uncertain estimate of relative effect.

Baseline event rate

An estimate of the baseline event rate in the absence of the intervention is required. This is used to obtain an estimate of the absolute effect of the intervention on the primary outcome by applying the relative measure of effect to the baseline risk. The baseline probability of an event is also likely to be uncertain. This may be informed by the event rate in the control arms of the trials in the meta-analysis informing the relative intervention effect or, alternatively, from external evidence or judgements relevant to the target population.

Incidence per annum

An estimate of the number of patients facing the uncertain choice between alternative interventions is required in order to establish the size of the benefits to the target population.

Minimum clinical difference (MCD) in primary outcome

Specifying a MCD in the primary outcome that is required for the results of research to have an impact on clinical practice is one way to incorporate concerns that the primary outcome does not capture all important aspects of outcome and/or costs. The MCD represents the improvement in the primary outcome that would need to be detected for the new treatment to be considered worthwhile and to have an impact on clinical practice. Specifying a MCD is one way to implicitly account for the other aspects of outcome that are not captured in the primary outcome.

Costs of the proposed new study

Some assessment of the likely costs of the proposed new study is required in order to establish whether the expected benefits of the study are sufficient to justify the expected costs. It can also be used to establish whether the proposed study represents

a priority compared to other research that could be commissioned using the same resources.

Duration of the proposed new study

An assessment of the duration of time it will take for the proposed research to be conducted and for the results to report is required since the health benefits of research decline the longer it takes research to report. This might be informed by an assessment of study sample size, expected recruitment rates, or historical experience from conducting similar types of studies.

Length of time for which new evidence is expected to be valuable

The information generated by new research will not be valuable indefinitely because other changes occur over time. For example, over time new and more effective interventions become available, which will eventually make those currently available obsolete. This means that new information about effectiveness is only relevant for a specific amount of time. A judgement about the length of time that the evidence from the proposed new study might be valuable is required in order to estimate the expected benefits over an appropriate time horizon. This judgement could be informed by historical evidence or experience about whether a particular research area is likely to see future innovations and/or other evaluative research reporting.

Discount rate

When a time horizon greater than one year is considered, discounting should be used to reflect the fact that resources committed today could be invested at a real rate of return to provide more resources in the future. Guidance from the UK Treasury suggests the use of a discount rate of 3.5% per annum (*HMT Green Book*, 2013).

2.3.4 RAPID ESTIMATION OF THE HEALTH BENEFITS OF RESEARCH

The health benefits of research to resolve uncertainty in the primary outcome are estimated by sampling from the uncertain distributions of relative effect and baseline event rate (i.e., from the range of plausible values specified by the CI or standard error on these quantities) and multiplying by the number of patients per annum whose treatment choice is to be informed by the decision. Each sampled value from the distributions is interpreted as one possible realisation of how patient outcomes

might turn out in practice, as supported by the existing evidence, i.e., each sampled value represents one possible 'true' value of how patient outcomes could turn out. Repeating this process many (e.g., 50,000) times creates a distribution of the health consequences of uncertainty, which is expressed in terms of the primary outcome measure[9]. For example, if the primary outcome is mortality, the consequences of uncertainty are expressed in terms of number of deaths per annum. The distribution of consequences tells us the chance of making an 'incorrect' decision due to uncertainty in the existing evidence base, while the number of patients affected by the decision provides the scale of the health consequences per annum. The average over this distribution provides an expected upper bound on the health benefits that could be gained by conducting further research to resolve this uncertainty.

The process for the estimation of the health benefits of research is illustrated using the CRASH example introduced in Section 2.2.1. Table 2.1 shows 5 random (equally likely) samples taken from independent distributions of relative effect (odds ratio for the intervention relative to a control) and baseline risk (control) for a primary outcome of mortality.[10] These are combined with the incident population (8,800 patients per annum) to understand the absolute health consequences of the decision. The last row of Table 2.1 represents the average across the 5 sampled values. For the 5 sampled values, the balance of evidence based on what is already known about the intervention and control indicates that the intervention is expected to reduce the number of deaths per annum by 80 compared with the control. This is because the expected odds ratio for death is 0.96 which is in favour of the intervention (column A). As a consequence, the mean absolute number of deaths for the intervention is 2,669 per annum (column E) compared with the baseline risk for the control of 2,749 deaths per annum (column D). This indicates that implementing the intervention would be worth (2,749 – 2,669 =) 80 deaths averted per year.

Though current evidence favour the intervention, there is a possibility that the intervention would increase rather than reduce mortality – this is seen in Table 2.1 for the sampled realisations 4 and 5, where the odds ratio for the intervention is

[9] Analytic solutions to calculation VOI without simulations are possible, however this becomes challenging in situations in which more than two treatments are being compared (Claxton, 1999; Schlaifer and Raiffa, 1961; Willan and Pinto, 2005).

[10] Correlation in outcomes between baseline risk and intervention effectiveness should be preserved where possible. For example, a multivariate or bivariate meta-analysis may be more appropriate to account for the dependence between multiple and possibly correlated outcomes (Riley, 2009).

greater than one. The expected health consequences of this uncertainty depends on the likelihood that the intervention is less effective than the control, how much less effective it is, and the size of the eligible population. The chance that the intervention is less effective is simply the chance of observing an odds ratio of greater than one, which is 40%, i.e. in 2 out of the 5 samples. The resulting consequences of this uncertainty is the number of additional deaths incurred if the intervention is used instead of the control in these instances (realisations 4 and 5 of column G). The expected health benefits of additional evidence to resolve this uncertainty is a weighting of the consequences of the uncertainty by the likelihood of them occurring, which is the average of the consequences of uncertainty across the sampled realisations (i.e., the average of column G). Therefore, the expected upper bound on the health benefits of research to resolve uncertainty is 48 deaths averted per annum for this sample.

Table 2.1: Rapid estimation of the health benefits of additional evidence for a primary outcome of mortality

Sampled realisation of uncertainty	Odds ratio for death (intervention vs. control)	Baseline odds of death for control	Odds of death for intervention (=A*B)†	Deaths per annum for an incidence of 8,800 eligible patients		Health benefits of additional evidence	
				Control (=8,800*B/(1+B))	Intervention (=8,800*C/(1+C))	Absolute effect for intervention in number of deaths per annum (=E-D)	Consequences of uncertainty for intervention (=F if A>1)
	[A]	[B]	[C]	[D]	[E]	[F]	[G]
Sample 1	0.83	0.6	0.5	3,300	2,933	-367	0
Sample 2	0.91	0.54	0.49	3,086	2,894	-192	0
Sample 3	0.95	0.49	0.47	2,894	2,814	-80	0
Sample 4	1.05	0.35	0.37	2,281	2,377	95	95
Sample 5	1.08	0.33	0.36	2,183	2,329	146	146
Average	0.96	0.46	0.44	2,749	2,669	-80	48

† Distributions for baseline odds and odds ratio are assumed independent.

It should be noted that the estimate of the value of additional research (48 deaths per year) and the value of implementation (80 deaths per year) reported in Table 2.1 are estimates which result form only 5 samples. This process should be repeated a large number of times in order to propoerly characterise the consequences of uncertianty in the odds ratio and baseline odds of deaths. After 50,000 samples it is estimated that the consequences of uncertainty are 39 deaths per year and the value of implemtation is 170 deaths per year. These estimates are different to those in Table 2.1. This difference is a result of chance and shows why a large number of simulations are necessary. This method is equivalent to doing a probabilistic sensitivity analysis on a decision tree in which the payoffs are quantified in terms of the primary outcome (mortality).

2.3.4.1 *Using MCD to capture other aspects of outcome*

If there are other important aspects of outcome that are not captured in the endpoint of mortality, then a MCD may be specified. Due to increased costs of the intervention a MCD of 0.5% deaths averted per annum may be required meaning that clinical practice should only change if the results of the new study indicate that at least (8,800 x 0.5% =) 44 additional deaths are averted per year with the new treatment. Therefore, the *gross* expected gain from the intervention of 80 deaths averted per year can be adjusted using the MCD to estimate the *net* health benefits of implementing the new treatment which are (80 -44 =) 36 deaths averted per year. The value of research will also be affected by taking account of the MCD to estimate the *net* health benefits of the intervention. Table 2.2 shows uncertainty in *gross* deaths per annum for the control and the intervention (column D and E). Column H shows the uncertainty in *net* deaths per annum after taking account of the MCD.

Table 2.2: Rapid estimation of the health benefits of additional evidence for a primary outcome of mortality taking account of MCD = 44 deaths per annum

Deaths per annum for an incidence of 8,800 eligible patients			Net health benefits of additional evidence	
Control (=8,800*B/(1+B))	Intervention (=8,800*C/(1+C))	Net deaths per annum for intervention (=E+44)	Absolute net effect for intervention in number of deaths per annum (=H-D)	Consequences of uncertainty for intervention (=I if I>0)
[D]	[E]	[H]	[I]	[J]
3,300	2,933	2,977	-323	0
3,086	2,894	2,938	-148	0
2,894	2,814	2,858	-36	0
2,281	2,377	2,421	139	139
2,183	2,329	2,373	190	190
2,749	2,669	2,713	-36	66

Column J of Table 2.2 shows the *net* health consequences of uncertainty. The *net* health consequences of uncertainty are larger than the *gross* health consequences of uncertainty in column G of Table 2.1. This is because the MCD penalises the new intervention and as the new intervention is expected to be superior to the current treatment, this makes the decision more uncertain. The adjustment for MCD has increased the expected upper bound on the health benefits of additional research from 48 to 66 deaths averted per year. Figure 2.3 below extends Figure 2.2 to graphically illustrate the effect on the value of research of increasing the MCD from 0% to 1.8%.

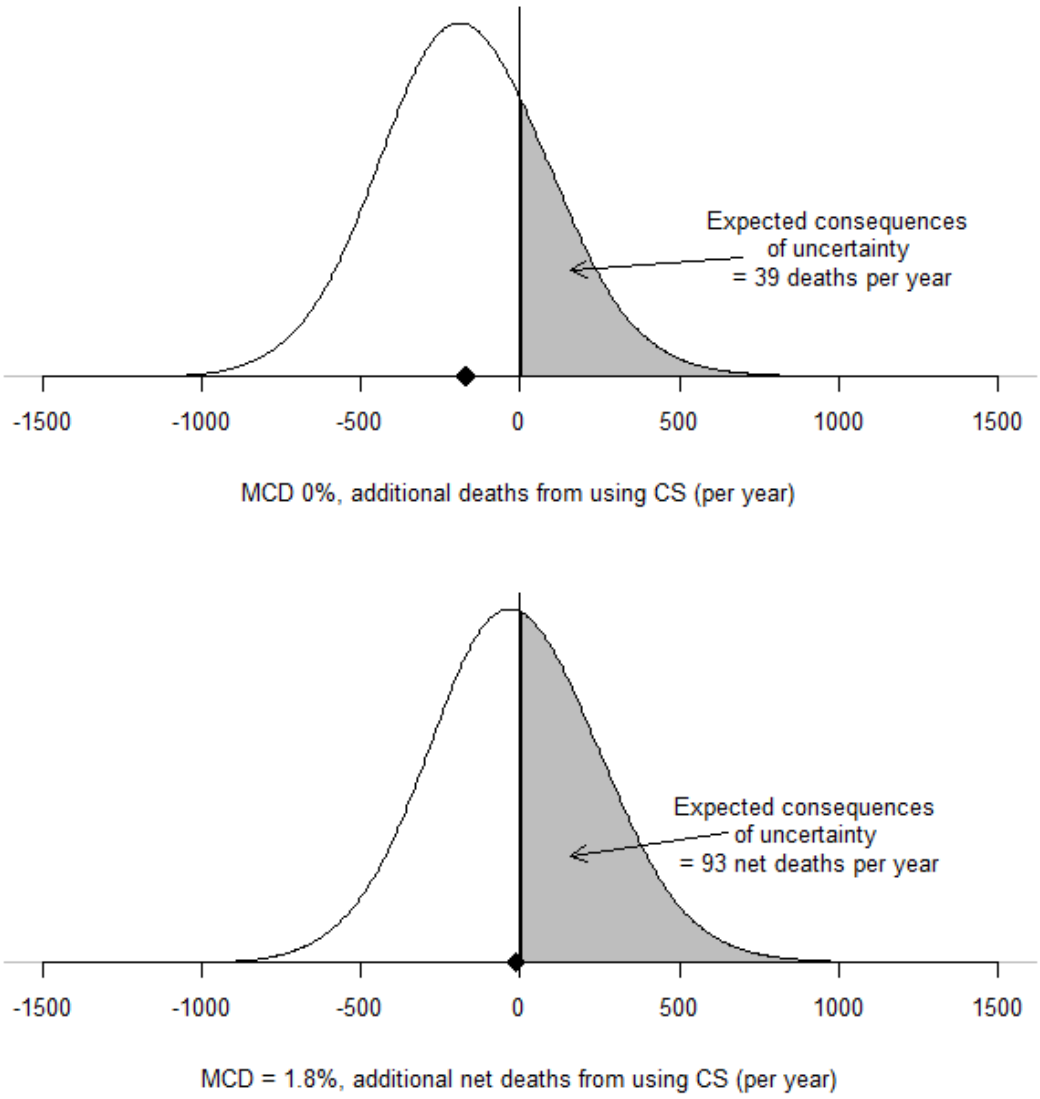


Figure 2.3: The effect of increasing the MCD for CS on the value of additional research in traumatic head injury. The upper panel shows the range of plausible outcomes associated with the use of CS with an MCD of zero. The upper panel shows the range of plausible outcomes associated with the use of CS with an MCD of 1.8%. A larger MCD in this case increases the uncertainty in the decision and so is associated with a larger value of additional research. CS, corticosteroids. MCD, minimum clinical difference.

As shown above, increasing the MCD from 0% to 1.8% adds 158 deaths[11] to the use of CS each year to take account of differences between the treatments. This shifts the distribution of net outcomes with CS to the right. This means that there is a greater chance of increasing the number of net deaths with CS and so makes the decision

[11] (0.018 – 0) x 8,800 = 158

more uncertain. This increase in uncertianty is represented by a larger gray area in Figure 2.3.

Deciding on an appropriate MCD requries trading off the health benefits of the priamry outcome with any other outcomes which the primary outcome does not capture (secondary outcomes). An explicit method to inform the MCD is provided in Section 3.3.4.

2.4 APPLICATION OF THE RAPID EVIDENCE GENERATION APPROACH TO A RETROSPECTIVE SET OF RESEARCH PROPOSALS

2.4.1 ASSISTING NIHR PANEL AND BOARD RESEARCH PRIORITISATION DECISIONS

In the sections that follow, the methods described in Section 2.3 are applied to the P1 case study from the six NETSCC proposals introduced in Section 1.2. The same assessments are completed for all six proposals (Chapter 4 for analysis of P2-P5 and Chapter 5 for analysis of P6). Each proposal used to illustrate different issues that are encountered when establishing the need for research, and to illustrate the types of scientific value judgements that are required in each context.

2.4.2 PROPOSED TRIAL OF EARLY VERSUS LATE TREATMENT OF PROPHYLAXIS TO REDUCE VENOUS THROMBOEMBOLISM FOLLOWING TRAUMATIC BRAIN INJURY

2.4.2.1 *Why is the research needed?*

P1 is for a proposed RCT to examine the efficacy and safety of early (before end of day 3) pharmacological thromboembolism prophylaxis (PTP) compared with late (from day 8 onwards) PTP following traumatic brain injury (TBI). PTP, which is a blood thinner that is used to protect against serious blood clots following TBI, is established standard practice in the UK. However, there is uncertainty about the optimal timing of initiation of the treatment. Early PTP has an associated risk of worsening intracerebral haemorrhage (ICH), while late PTP has an increased risk of venous thromboembolism (VTE) (including deep vein thrombosis, DVT) impacting on recovery and patient outcomes. A summary of the research question and proposed study is presented in Box 2.1.

Summary of proposal 1

Research question: Does early pharmacological thromboembolism prophylaxis (PTP) after traumatic brain injury (TBI) improve functional recovery at 6 months?

Intervention: Early PTP (before end of day 3 after TBI) with Dalteparin 5000u.

Control: Late PTP from day 8 onwards with Dalteparin 5000u, representing standard care until intensive care unit (ICU) discharge.

Outcomes:
Primary outcome - Functional recovery and mortality at six months after TBI, assessed by the Glasgow Outcome Scale Extended (GOSE).

Secondary outcomes - DVT; pulmonary embolism; ICH expansion; pneumonia; heparin induced thrombocytopaenia; length of stay in ICU.

Proposed study: RCT to examine the efficacy, cost-effectiveness and safety of early PTP in 1300 patients after TBI measured by functional recovery at 6 months.

Duration of proposed study: 5 years

Costs of proposed study to NETSCC: £2,854,000

NHS support and treatment costs: £490,000

Box 2.1: Summary of proposal 1

2.4.3 EVIDENCE REQUIRED TO ESTABLISH THE HEALTH BENEFITS OF RESEARCH FOR PROPOSAL 1

Primary outcome measure

The primary outcome for the proposed trial is functional recovery and mortality at six months after TBI, as measured by the Glasgow Outcome Scale Extended (GOSE). The GOSE is a scale for functional outcome that rates patient status into one of eight categories: dead (GOSE of 1), vegetative state (GOSE of 2), lower severe disability (GOSE of 3), upper severe disability (GOSE of 4), lower moderate disability (GOSE of 5), upper moderate disability (GOSE of 6), lower good recovery (GOSE of 7), or upper good recovery (GOSE of 8).

The health benefits associated with the proposed trial are estimated by comparing the percentage of patients who are expected to be functionally recovered at 6 months in the early PTP and late PTP treatment groups. For simplicity, and following the approach used by other authors (Nichol et al., 2015), the GOSE is dichotomised such that a patient with a GOSE score of ≥5 is defined as functionally recovered, while those with a GOSE score of 1-4 are classified as not having functionally recovered.

Relative effectiveness

An estimate of the relative effectiveness of early PTP compared with late PTP, based on what is already known about the interventions, is required for the primary outcome of functional recovery, along with an estimate of its uncertainty. The proposal summarises the existing evidence to date based on three systematic reviews that have been conducted examining the efficacy and safety of PTP following TBI. However, none of these reviews have considered the primary outcome of functional recovery. The evidence from the reviews included a number of small studies that assessed VTE rates and ICH progression in patients treated with early (< 72 hours) versus late (> 72 hours) thromboprophylaxis (Chelladurai et al., 2013; Jamjoom and Jamjoom, 2013). These studies concluded that there is insufficient evidence to comment on the effectiveness of early compared with late VTE prophylaxis and its effect on ICH progression. The Brain Trauma Foundation Guidelines also report insufficient evidence to recommend a preferred drug, dose or timing of initiation of PTP following TBI. UK neurological critical care units were surveyed as part of a Delphi panel and this identified the timing of PTP as an important research question

for clinicians in the management of TBI. In summary, this means that there is no existing evidence reported on the primary outcome of GOSE of the proposed trial and there is no quantitative estimate of uncertainty in relative effectiveness of the interventions. In order to quantify the benefits of the proposed RCT, an explicit estimate of uncertainty for the relative difference between early and late PTP in the primary outcome of functional recovery is required. In the absence of this information, some judgement about the uncertainty or range of plausible values that the quantity can take is required, approaches to address this are outlined in Section 2.4.4.

Baseline event rate

An estimate of the baseline event rate for late PTP is required. This is used to obtain an estimate of the absolute effect of early PTP on the primary outcome of functional recovery by applying the estimate of relative effectiveness of early PTP compared with late PTP to the baseline risk. Although event rates for VTE and proximal DVT were reported in P1, the baseline event rate for the primary outcome of functional recovery for late PTP was not reported in the proposal. The baseline event rate was derived from the placebo arm of a recent multicentre RCT (EPO-TBI), which examined the effects of erythropoietin compared with placebo on neurological recovery, mortality, and VTE in patients with TBI (Nichol et al., 2015). The primary outcome assessed at 6 months was the proportion of patients with GOSE of ≥5 (i.e., proportion of patients functionally recovered). Under the assumption that the placebo arm of this trial is a close approximation to standard of care of late PTP, the baseline event rate for functional recovery is 55.1% (=162 patients out of 294 at risk who achieved GOSE ≥5 with placebo). In order to reflect uncertainty in this estimate, a beta distribution based on 162 events out of a sample of 294 patients at risk was used (Briggs et al., 2006). This results in a 95% CI of 49% to 61%.

Incidence per annum

The expected health benefits of additional evidence depend on the size of the patient population whose treatment choice is to be informed by the evidence. Sauerland and Maegele (2004) estimate an approximate annual incidence of TBI in the UK of 8,800. Therefore, the impact of uncertainty on the absolute number of functional recoveries per year is estimated based on an annual incidence of 8,800.

Minimum clinical difference (MCD) in primary outcome

The proposed trial is designed based on the primary outcome of functional recovery. Therefore, the benefits of the proposed research and implementation efforts to change clinical practice are expressed in number of functional recoveries gained. However, functional recovery may not necessarily be the only relevant outcome for assessing the value of the proposed study. For example, early PTP incurs additional treatment costs compared with late PTP, i.e., 5 additional days of treatment costs. The drug cost of Dalteparin 5000u is relatively cheap at £2.82 per dose (BNF, 2018), over an additional 5 days this incurs a cost of £14.10 per person, which is equivalent to an additional cost of £124,000 (= £14.10 x 8,800) per year to the NHS. Specifying a MCD required to change clinical practice is one way to incorporate concerns about increased costs and/or potential adverse events that are not captured in the primary outcome. As additional per patient costs are small, then the required MCD will also be small. For illustrative purposes, since early PTP is more costly than late PTP it might be required to demonstrate a 1% increase in the number of functional recoveries, in addition to those expected when everything else is considered equal.

Costs of the proposed new study

Some assessment of the likely costs of the proposed RCT is required in order to establish whether the expected benefits of the study are sufficient to justify the expected costs. For NETSCC costs are divided into two categories. “Research costs” which fall on the research budget (i.e. they are borne by NETSCC) and cover the costs organising the research project such as research team payments, administrative support, travel costs and dissemination costs. “NHS support and treatment costs” fall on the general health system (NHS) budget and include the costs associated with using experimental treatments as part of the research. For P1, the research costs are estimated to be £2,854,000, while the NHS support and treatment costs are £490,000.

Duration of the proposed new study

Some assessment of the duration of time it takes for the proposed research to be conducted and for the results to report is required since the health benefits of research decline the longer it takes research to report. For P1, the RCT is expected to take 5 years to complete and report.

Length of time for which new evidence is expected to be valuable

The information generated by new research will not be valuable indefinitely because other changes occur over time. A judgement about the length of time that the evidence from the proposed new study might be valuable is required in order to estimate the expected benefits over an appropriate time horizon. It is anticipated that the new information for P1 might be valuable for a long time span of 15 years since standard practice in TBI appears to move relatively slowly. Alternative scenarios may be used to assess the impact of longer or shorter durations on the health benefits of the proposed research.

Discount rate

A discount rate of 3.5% per annum is used based on Guidance from the UK Treasury (*HMT Green Book*, 2013).

2.4.4 INFORMING A JUDGEMENT ABOUT RELATIVE EFFECTIVENESS IN THE ABSENCE OF EXISTING EVIDENCE

When deciding about the need for research, some judgement about the level of existing uncertainty (e.g., estimate of standard error) for relative effect of the alternative interventions is required. An explicit quantification of uncertainty in the relative effect for the primary outcome is required for use in VOI analysis[12]. Ideally, in addition to being based on any existing studies, this estimate should take account of expert knowledge about the disease process, clinical intuition and any relevant weaknesses in the evidence base e.g. from unrepresentative sampling or failure to blind participants. Making such explicit and quantified statements about uncertainty is challenging; however, this judgement must be made when prioritising research, either implicitly or explicitly. None of the six proposals received by NETSCC (listed in Section 2.4.1) reported a suitably explicit quantitative summary of the current level of uncertainty in the primary outcome. Unfortunately, this omission is not unique to this particular set of research proposals. Bennette et al., (2016) also found that none of the 9 proposals investigated as part of a US oncology research prioritisation exercise reported uncertainty in the primary outcome. This is likely because the currently most common approach to medical statistics (Frequentism)

[12] In more technical terminology, this is a Bayesian concept called a "prior" and reflects the state of the decision maker's knowledge given the current state of the evidence.

does not place emphasis on explicitly quantifying current levels of uncertainty, even though this is central to making decisions about the need for additional research. The first step to establishing whether additional research is required should always include a systematic review of what is already known about the interventions (Chalmers et al., 2014). It would be inappropriate and potentially unethical to fund research when sufficient evidence already exists to inform a decision about alternative interventions, especially if an experimental research design such as an RCT is required. However, it is recognised that for many decisions a review of the existing literature will show that there have not been any suitably similar studies carried out. In this case, there are a number of options available to inform a quantitative judgement of the uncertainty in the existing evidence base:

Expert opinion or expert elicitation

This is the process whereby relevant experts are asked to provide their judgement regarding the magnitude of a given quantity and its uncertainty. In its simplest form, this could involve asking experts for their 'best guess' estimates, but more formal methods of expert elicitation are readily available that involve asking for probabilistic belief statements about unknown quantities and using formal processes to combine judgements from multiple experts (Mason et al., 2017; O'Hagan et al., 2006; Wilson et al., 2018). Where some limited data may be available for a given quantity, expert elicitation may be used to supplement this information, and a large literature exists on expert elicitation in Bayesian statistics (O'Hagan et al., 2006). The elicitation task need not be complex to aid decision making. In utilising rapid methods to aid research prioritisation in the US, (Carlson et al., 2018) developed a questionnaire consisting of two questions to inform uncertainty in relative effect: Question 1. "What is the probability that the new treatment is equivalent or better than the control arm for the primary outcome?" Question 2." What is the probability the new treatment offers a substantial improvement over the control arm in the primary outcome?"

Statistical modelling or extrapolation

In the absence of information on a particular quantity in a population of interest, it may be possible to extrapolate data from other sources in order to inform the quantity in the population of interest. For example, in a recent National Institute for Health and Care Excellence (NICE) technology appraisal of biological therapies for

chronic plaque psoriasis in children and young people, data from the adult population on the relative effectiveness of the biological therapies was used to complement the limited evidence base for these treatments in children and young people (Duarte et al., 2017). It is also possible to reflect the uncertainty involved in the extrapolation process.

Meta-epidemiological studies

This is a rapidly developing area of research in which large databases of RCTs and meta-analyses are reviewed to describe the distribution of research evidence to inform a specific question and to understand heterogeneity between studies and control for potential bias (Bae, 2014). It may be used to assess the impact of study characteristics on treatment effect estimates and to identify possible effect modifiers. Meta-epidemiological studies provide a useful quantitative analysis of historical study results that may be used to empirically inform judgments about plausible effects for interventions where study results are judged to be exchangeable. For P1, a meta-epidemiological approach was used to inform a quantitative judgement about uncertainty in relative effectiveness in the absence of alternative information. This study utilised the meta-epidemiological analysis by Djulbegovic et al., (2012), which examined the likelihood that new treatments being compared to established treatments in randomised trials would be superior[13]. The analysis examined 743 publically funded RCTs across a diverse range of conditions and involved 297,744 patients. The results of the study found that, on average, new treatments were slightly more likely to have favourable results than established treatments for the primary targeted outcome, i.e., slightly more than half of publically funded RCTs demonstrate an improvement in the primary outcome for new treatments compared with established practice, while slightly less than half demonstrate less favourable outcomes. The magnitude of effect on primary outcomes indicates that new treatments favour established practice with an average odds ratio of 0.91.[14] A kernel density analysis showed that the pooled trial results were approximately symmetrical

[13] In prioritising research in a US oncology setting Bennette et al., (2016) also draw on the meta-epidemiological analysis by Djulbegovic et al., (2012) to address a lack of data on relative treatment effects. The approach taken by Bennette and colleagues makes use of the judgements implied in sample size calculations to mimic an expert elicitation exercise. Though related, the meta-epidemiological method used in this chapter is distinct from this approach.

[14] In order to combine results across RCTs, primary outcomes in Djulbegovic et al., (2012) were coded as "harm" rather than as "benefit", such that a reduction in the outcome favours the new treatment.

of new versus established treatments centred near 'no effect', but that the results of individual RCTs are unpredictable and tend to fall within an approximate 95% interval from 0.19 to 4.39[15]. The study also reports that trial results have not changed significantly over time, and are not significantly affected by choice of comparator (e.g., active treatment versus placebo).

In the absence of other information, the meta-epidemiological approach based on historical trials represents a reasonable starting point to inform the magnitude of effect and uncertainty for the relative effectiveness of early PTP compared with late PTP following TBI. This suggests an odds ratio of 1.09 in favour of early PTP improving functional recovery compared with late PTP (i.e., reciprocal of 0.91 for harmful outcome) and a 95% CI of 0.23 to 5.24. This effect estimate and CI describe the distribution of possible values for the relative effectiveness of early PTP compared with late PTP.

2.4.5 RAPID ESTIMATION OF THE HEALTH BENEFITS OF THE PROPOSED TRIAL

On the balance of existing evidence, early PTP is judged to be the most effective option. However, due to uncertainty in the evidence base, there is a 46% chance that early PTP is less effective than the baseline treatment of late PTP. The health consequences of this uncertainty, at a population level, are estimated by combining the range of plausible values for the relative effect of early PTP with the baseline probability of functional recovery following TBI, which is also estimated with uncertainty, and the annual incidence of TBI.

The distribution of the health consequences of uncertainty for early PTP following TBI is illustrated in Figure 2.4. The balance of evidence favours early PTP with a 54% chance that it results in more functional recoveries than late PTP. However, this means that there is a 46% chance that late PTP could result in additional functional recoveries gained per year. The health consequences of uncertainty are not uniform; there is a greater chance of more limited consequences (e.g., 9% chance of between zero and 250 functional recoveries lost per year) and a smaller chance of greater consequences (e.g., 5% chance of more than 2,250 functional recoveries lost per

[15] This is based on an estimated standard error of 0.8 on the log scale.

year). The average over this range of potential outcomes provides an estimate of the expected health benefits that could potentially be gained each year if the uncertainty in the decision were resolved. This corresponds to an expected upper bound on the value of further research of 577 functional recoveries per year[16].

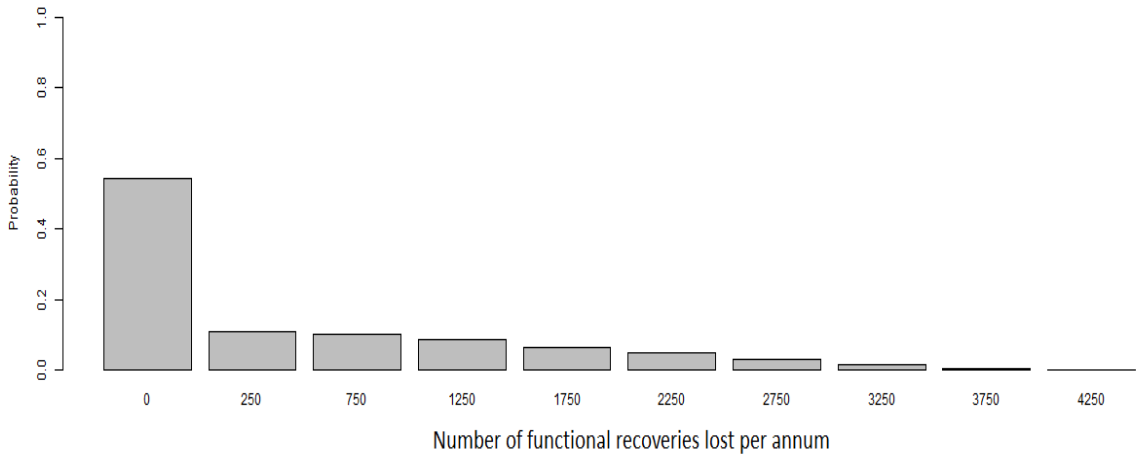


Figure 2.4: Expected consequences of uncertainty in terms of number of functional recoveries lost per annum due to uncertainty.

These expected benefits increase with the size of the patient population whose treatment choice can be informed by additional evidence and the time over which evidence about the effectiveness of the interventions is expected to be useful. It is expected that the information from research will be valuable for approximately 15 years. This means that the consequences of uncertainty surrounding the decision increases greatly by the fact that, in the absence of better evidence, the health system is likely to utilise the suboptimal treatment option every year for the next 15 years. Extending the yearly consequences of uncertainty over the 15 year time horizon, means that the expected maximum health benefits of research is estimated to be 6,732 functional recoveries gained over the full time horizon (after discounting appropriately). This is illustrated in Figure 2.5.

[16] This is calculated by multiplying the consequences by the probability of those consequences: 0 x 54.7% + 250 x 11% + 750 x 9.9% +1250 x 8.3% + 1750 x 6.5% + 2250 x 4.6% + 2750 x 2.9% + 3250 x 1.5% + 3750 x 0.5% + 4250 x 0.1% = 574 functional recoveries per year. This is marginally different due to rounding errors.

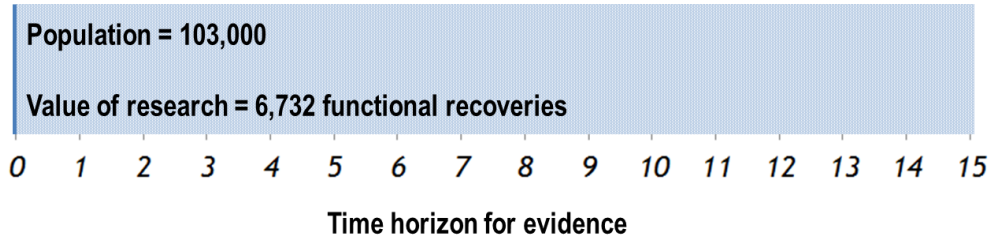


Figure 2.5: Expected maximum health benefits of research to inform the optimal timing of PTP following TBI over a time horizon of 15 years.

For P1, the proposed RCT is expected to take 5 years to complete and report, and so the expected upper bound on the health benefits of research will fall from 6,732 to 4,086 functional recoveries gained. This is illustrated in Figure 2.6.

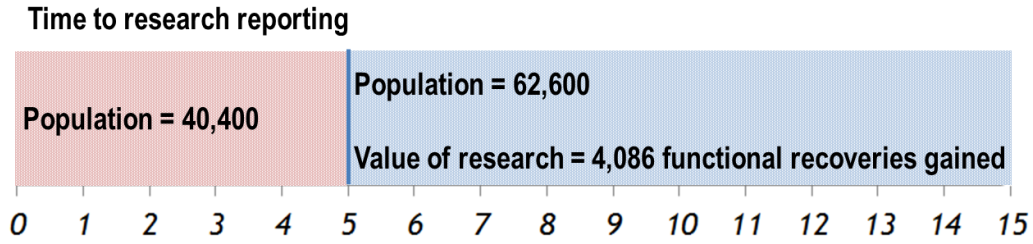


Figure 2.6: Expected maximum health benefits of research to inform the optimal timing of PTP following TBI over a time horizon of 15 years, with research reporting in 5 years.

As the research funding budget is limited, the same analysis should be undertaken for all competing proposals in order to compare the value of research across proposals. However, before doing this, the costs of the research should to be taken account of. The proposed RCT is expected to cost the research funder £2,854,000, which means that the maximum expected value of the proposed research is estimated to be (£2,854,000/4,086 =) £699 per functional recovery gained. Whether this represents good value to NETSCC depends on the health benefits of the other competing research proposals, which may or may not be funded due to the resources required to fund P1.

Taking account of other aspects of outcome

The proposed trial is designed with functional recovery as its primary outcome so the benefits of implementation and the proposed research are expressed as functional

recoveries gained. Although functional recovery may be the most appropriate primary outcome, it is not necessarily the only relevant outcome for assessing the value of the proposed treatments. For example, relative to current practice, early PTP incurs additional treatment costs of £124,000 per year. Specifying a MCD required to change clinical practice is one way to incorporate concerns about increased costs and/or potential adverse events that are not captured in the primary outcome. For example, since early PTP is more costly than late PTP it might be required to demonstrate a 1% increase in the number of functional recoveries, in addition to those expected when everything else is considered equal. Requiring a 1% increase implies that the additional costs are equivalent to 88 additional functional recoveries per year[17].

Taking the 1% increase into account using the method described in Section 2.3.4, the consequences of uncertainty are now estimated to be 618 additional functional recoveries per year. This is higher than the previous estimate of 577 additional recoveries per year because taking account of the MCD penalises early PTP, which is expected to be slightly superior given current evidence, this making the decision between treatments more uncertain and so increases the value of additional evidence. The expected value of the additional evidence is 4,377 functional recoveries gained over the 15 year period (taking account of the 5 years required for research to report). This translates to an expected maximum value of the trial of (£2,854,000/4,377 =) £652 per functional recovery gained.

[17] The trade-off implied by requiring a percentage change in a binary outcome is calculated by multiplying the required increase (e.g. 1%) by the incident population (e.g. 8,800) = 8,800 x 0.01 = 88.

2.5 PRIORITISING ACROSS RESEARCH PROPOSALS

The previous section shows how to estimate the expected benefit of proposed research in terms of the primary outcome (e.g., cost per functional recovery after TBI). Research funders with a fixed research budget (like NETSCC) need to make research prioritisation decisions across a diverse range of research proposals with different primary outcomes. To explore the implications of this, Table 2.3 reports the expected value of the proposed research for each 6 NETSCC proposals in terms of cost per primary outcome (i.e., expected benefits of proposed research net the research costs). It also reports the scale of the additional research costs imposed on the NHS outside of the NETSCC budget (i.e., NHS Support and Treatment costs). The method to estimate the benefit of P1 has been illustrated in the preceding sections of this chapter. Rapid VOI approaches have also been used to estimate the value of funding P2-P6. The methods used are an extension of those illustrated in this chapter. See Chapter 4 for analysis of P2-P5. See Chapter 5 for analysis of P6.

Table 2.3: Expected value of proposed research study expressed in terms of the primary outcome for each of the six NETSCC proposals

Proposal	Upper bound for health impact of research	Cost to NETSCC	NETSCC cost per outcome	Additional NHS Support and Treatment costs
P1: Brain injury	4,377 additional functional recoveries	£2,854,000	£652 per additional functional recovery	£490,000
P2: Melanoma	544 additional months of progression free survival	£2,522,710	£4,641 per additional months of progression free survival	£62,410,967 cost savings
P3: Rare disease	1 composite endpoint** prevented	£855,403	£570,510 per composite endpoint prevented	£10,608,500 cost savings
P4: Alzheimer's	2,390 additional MMSE* points	£3,310,883	£1,385 per additional MMSE point	£1,297,789
P5: Psychosis	245 relapses avoided	£601,481	£2,458 per relapse avoided	£150,000
P6: Information booklet	112 additional deaths in preferred place	£882,177	£7,869 per additional death in preferred place	£4,104
Total costs		£11,026,654		

*MMSE: mini mental state examination. Questionnaire used to assess severity of Alzheimer's disease. Maximum score is 30. Scores in the MMSE are often classified into different categories: 26–30 (normal ageing), 21–25 (mild dementia), 15–20 (moderate dementia), 10–14 (moderately severe dementia) and 0–9 (severe dementia)
** Composite endpoint: death, meningococcal infection or irreversible organ injury

The expected value of the proposed research across the 6 proposals is expressed in terms of a broad range of outcomes. There are substantial differences in the number of primary outcomes expected gained/avoided across the different proposals (column 2). The main factors which account for these differences are:

Incident population

Proposals which affect a large number of patients each year such as P4 (100,000/year) and P1 (8,800/year) result in a large number of primary outcomes gain/avoided, 2,390 additional MMSE points and 4,377 additional functional recoveries respectively. This is compared to research proposals which affect fewer people per year; P5 (1,563/year), P2 (1,137/year) and P3 (26.3/year) which are expected to

result in 245 relapses avoided, 544 additional months of progression free survival and 1 composite endpoint prevented respectively.

Degree of uncertainty in relative effects

P6 is an important exception to the rule that a larger incident population results in a larger health impact (in terms of primary outcomes gained/avoided). P6 has a large incident population (259,150/year) but only results in 112 additional deaths in preferred place. This is because P6 investigates the effect of an information booklet on the ability of terminal patients to die in their preferred place (e.g. at home). As it is highly unlikely that providing the booklet will reduce the number of terminal patients dying in their preferred place there is not much uncertainty in the direction of the relative treatment effect. As the booklet is cheap and is likely to improve the primary outcome, there is not much uncertainty for P6 to address and so it is likely to have a relatively small health impact, despite its large incident population.

In addition to health outcomes per NETSCC spend, the proposals differ in the additional costs of research imposed on the NHS. The health foregone due to these additional research costs should also be considered when choosing among the competing research proposals. For example, P2 and P3 claim substantial NHS cost savings associated with funding these research projects. In both cases, current practice is associated with large treatment costs and, therefore, the cost savings are assumed to follow by allocating patients to less costly treatment options as part of the planned research projects. However, whether these cost savings can be expected to materialise or not is discussed in Section 3.4.

2.6 DISCUSSION

VOI methods can potentially improve the transparency and accountability of research prioritisation decisions, however a number of barriers to their use in policy remain. This chapter investigated the feasibility of the rapid VOI methods developed by Claxton et al., (2015a) for applied research prioritisation using a set of retrospective research funding applications received by NETSCC. This method can potentially improve the transparency and accountability of research prioritisation by highlighting the trade-offs in primary outcomes required when choosing between research proposals while taking account of the size of the population affected by the decision and the costs of research. The approach does not require a large number of inputs and so should be feasible to incorporate into the research prioritisation process.

In order to rank the priority of a set of research proposals decision makers should take account of which projects represent the "best buys" to the research funder (Culyer, 2016). Table 2.3 shows the value of each project expressed in terms of its cost per primary outcome (e.g. £652 per additional function recovery). This information helps decision makers to compare value across proposals by making the trade-offs between the different outcomes clearer. By providing a starting point for deliberation, this analysis represents a clear improvement relative to implicit forms of decision making. However, significant implicit scientific judgements are still required to make decisions. First, some primary outcomes may be more severe than others (e.g., death versus short periods of immobility). Second, some outcomes represent concrete clinical events (e.g., overall survival) while others represent surrogates (e.g. changes in blood pressure), which may or may not be good predictors of relevant outcomes (Kemp and Prasad, 2017). Finally, research will result in additional costs that fall directly on the general health system budget (e.g. NHS) and the differential health opportunity costs associated with these resources must also be accounted for. This means that although the analysis presented in Table 2.3 can greatly improve the transparency and accountability of research prioritisation, the task can be supported further by expressing all primary outcomes in terms of a generic outcome which can be compared across disease areas (such as QALYs). This extension and the advantages provided are demonstrated in the next chapter.

As in most applied VOI analysis (e.g. Bennette et al., (2016) and Carlson et al., (2018)), the methods illustrated here makes two simplifying assumptions about decision making. First, it is assumed that there are no irrecoverable costs associated with the treatments considered. Irrecoverable costs are those costs that, once committed, cannot be recovered if guidance is changed at a later date (McKenna et al., 2015). For example, if large capital investments are required for a treatment to be delivered in a health system then this is associated with irrecoverable costs if the health system cannot easily sell the capital when guidance changes. The second assumption is that decisions to recommend an intervention for widespread use and decisions to carry out additional research on that same intervention are independent. However, there are cases in which research cannot be carried out at the same time as the treatment is available for widespread use. Chapter 6 explores the consequences of relaxing these assumptions.

In applying VOI methods to a retrospective set of real-world proposals received by NETSCC, this chapter has highlighted the importance of methods to explicitly quantify uncertainty in situations in which suitable pre-existing studies do not exist. This is fundamental issue in the use of VOI to aid research prioritisation and has also been highlighted by Bennette et al., (2016) and Carlson et al., (2018). Despite its importance this issue has received very little attention in the VOI literature. To address this, we outlined three available options for decision makers; expert elicitation, statistical modelling and the use of meta-epidemiological studies. In this chapter we demonstrated the use of a meta-epidemiological approach related to that used by Bennette and Carlson. Though providing a useful starting point for deliberation, this method could be greatly improved with further research. The primary limitation of the approach used here is that applying a generic meta-epidemiological estimate to inform uncertainty across all research proposals does not take account of contextual differences between proposals. There may be good scientific reasons to treat proposals differently. This contextual information could be incorporated by (i) integrating expert elicitation and meta-epidemiological methods, (ii) utilising more sophisticated statistical methods to reflect the fact that different disease areas and types of outcome are associated with different distributions of effect sizes (iii) combining approaches i and ii. Developing these methods with the

aim of supporting research prioritisation is potentially an important and fruitful area of further research.

Relatedly, rapid methods of expert elicitation may be useful to develop and integrate into the research prioritisation process. As judgements about uncertainty must be made either implicitly or explicitly, expert elicitation may be especially useful in cases in which previous randomised trials have not been carried out and the use of meta-epidemiological evidence is deemed inappropriate. The importance of expert elicitation in decision making has been increasingly recognised in recent years and number of user friendly tools have been developed to facilitate its integration into decision making (Mason et al., 2017; O'Hagan and Oakley, 2018).

Chapter 3

Prioritising across research proposals by linking primary outcomes to a comprehensive measure of health outcome

3.1 INTRODUCTION

The previous chapter demonstrated that it is possible to rapidly estimate the value of research in terms of cost per primary outcome e.g. £652 per additional function recovery. This provides a useful starting point for prioritisation however it was also shown that choosing between proposals requires a number of implicit scientific judgements. These are required as some primary outcomes are more severe than others and some primary outcomes are surrogates rather than outcomes of value in themselves. Additionally, because the primary outcomes are not comparable across disease areas the analysis presented in Table 2.3 cannot help address other questions which may be of interest to decision makers such as whether sufficient resources being devoted to research.

The aim of this chapter is to show how these limitations can be addressed and the scope of the analysis expanded by using external evidence to link each primary outcome to a more comprehensive measure of health outcome. This more comprehensive measure must be generic enough to compare benefits across disease areas and should capture both quality and quantity of life. QALYs are used throughout this thesis but the methods apply to any other comprehensive measure of health (e.g. disability adjusted life years (DALYs)). In this chapter we illustrate how to link binary primary outcomes to QALYs to calculate the health impact of the research in QALYs. This extension has been demonstrated in (McKenna et al., 2016) and is illustrated by continuing to use the P1 case study from Chapter 2.

After this we explore the implications of prioritising across a set of research proposals when all benefits are quantified in QALYs. There are a number of implications which are illustrated using the full set of 6 retrospective research proposals received by NETSCC.

The budget for funding research is often not the same as the budget that funds general healthcare activities such as staff salaries, drug costs, equipment etc. Therefore, the benefits of a particular research proposal must be compared to the benefits of other research proposals which could have been funded with these resources. This means that research bodies such as NETSCC in the UK cannot fund every research proposal which offers health gains relative to general health expenditure. In order to reflect these trade-offs appropriately, the "bookshelf" approach described by Culyer (2016) is used in this chapter. The bookshelf framework is used here to rank interventions from highest to lowest health impact to identify the "best buys" for decision makers. This framework is also provides a basis to understand whether health outcomes could be improved by expanding the research budget relative to the budget for general health expenditure. The bookshelf also shows the population health implications for charitable and industry contributions to research funding.

As the research budget is limited, funding research as a method to change practice will necessarily divert resources away from research projects which could address genuine uncertainties in the health system. However, in some scenarios, research may be required to change practice. In Section 3.4 we use the framework to show that when practice *cannot* change in absence of research then smaller research studies which report results sooner (and thus can change practice more quickly) will be more valuable to fund than large time consuming studies. Additionally, Section 3.3.4 provides an explicit method to inform an appropriate MCD with reference to the health impact of the primary outcome.

3.2 LINKING PRIMARY OUTCOME TO A COMPREHENSIVE MEASURE OF HEALTH OUTCOME

Health related quality of life and disease related costs

Continuing the P1 case study from Chapter 2, the benefits of research were estimated in terms of functional recoveries gained. Functional recovery was based on the GOSE score which captures both mortality and quality of life effects. In this section, we show how the analysis can be extended to estimate the QALYs associated with functional recovery.

If a patient functionally recovers there is a chance that they will end up in one of the GOSE states from moderate disability to good recovery (GOSE scores of 5 to 8). If a patient does not recover, there is a chance that they will either die (score of 1) or end up in vegetative or severe disability states (scores of 2 to 4). A previous TBI study by Nichol et al., (2015) is used to estimate the relative probabilities of ending up in each of the GOSE states conditional on functional recovery status; this is shown in column 3 of Table 3.1.

Table 3.1: Costs and QALYs associated with functional recovery as defined by GOSE score 5-8

Primary outcome occurs (functional recovery)							
Possible states	**Outcome**	**Probability of being in state**	**Life expectancy (years)**	**Utility weights**	**Disease related costs**	**Expected QALYs**	**Expected costs**
GOSE 5	Moderate disability	42%	16.73	0.7	£27,047	4.92	£11,360
GOSE 6		24%	16.73	0.81	£27,047	3.25	£6,491
GOSE 7	Good recovery	20%	19.23	0.96	£19,575	3.69	£3,915
GOSE 8		14%	19.23	1	£19,575	2.69	£2,741
		100%				14.55	£24,507

Linking survival (Shavelle et al., 2006), quality of life (Fuller et al., 2017) and disease related costs (Nyein et al., 1999; Wood et al., 1999)[18] to each of these GOSE states allows us to estimate the consequences of functional recovery expressed in terms of costs and QALYs. The expected QALYs for each GOSE state are calculated

[18] See NICE guidance (National Institute for health and care excellence, 2014) Head injury: Methods, Evidence & Guidance. Economic evaluation reported in chapter 11: A cost-effectiveness analysis of transporting patients with serious head injury directly from the injury scene to a specialist neurosciences hospital.

by multiplying the life expectancy in each state by the utility weight and the probability of being in that state e.g. for GOSE 5 the expected QALYs are 16.73 x 0.7 x 42% = 4.92 QALYs. Summing the expected QALYs for each of the states classified as functional recovery provides an estimate of the expected QALYs associated with functional recovery, which is 14.55 QALYs. The same process is carried out to estimate the expected disease related costs of functional recovery (£24,507) and the consequences of no functional recovery (shown in Table 3.2).

Table 3.2: Costs and QALYs if functional recovery does not occur as defined by GOSE score 1-4

Primary outcome does not occur (no functional recovery)							
Possible states	**Outcome**	**Probability of being in state**	**Life expectancy (years)**	**Utility weights**	**Disease related costs**	**Expected QALYs**	**Expected costs**
GOSE 1	Dead	29%	0	0	£0	0	£0
GOSE 2	Vegetative	7%	7.11	0.11	£45,450	0.05	£3,182
GOSE 3	Severe disability	41%	12.52	0.41	£154,324	2.1	£63,273
GOSE 4		23%	12.52	0.58	£154,324	1.67	£35,495
		100%				3.82	£101,950

Each functional recovery is expected to result in (14.55 – 3.82 =) 10.7 QALYs gained per person and (£101,950 - £24,507 =) £77,443 in cost savings per person over a lifetime. These costs savings will free up resources in the NHS budget and, therefore, the health impact associated with these resources should also be considered. In order to reflect the health opportunity costs associated with higher costs to the NHS, we use the value of £15,000 per QALY, which has been endorsed by the UK Department of Health for use in health impact assessments (Claxton et al., 2015b; NHS England, 2015). This means that for every £15,000 of NHS resources, the health system can expect to produce one additional QALY. This implies that the cost savings per functional recovery are expected to be worth (£77,443/£15,000 =) 5.16 QALYs elsewhere in the health system. Combining the direct health benefits (10.7 QALYs) and the indirect health benefits through cost savings (5.16 QALYs) means that each additional functional recovery results in a gain of 15.86 QALYs[19].

[19] It should be noted that the method shown here is just one approach to estimating the expected incremental net health benefit associated with functional recovery. A more complex approach would be to use a Markov model to estimate lifetime costs and health outcomes associated with functional

Treatment costs

In addition to disease related costs, treatment costs should be taken into account to reflect the differences in the treatment options. For P1, all patients treated with early PTP incur additional treatment costs of £14.10 per patient, which implies additional costs of (£14.1 x 8,800 =) £124,080 per year. These increased costs are expected to displace (£124,080/£15,000 =) 8.3 QALYs per year for early PTP.

Expected upper bound on the value of research

To understand the benefits of research in terms of a comprehensive measure of outcome we follow the same rapid method as shown in Section 2.3.4. The uncertainty in the primary outcome is estimated by sampling from the uncertain distributions of relative effect and baseline event rate and multiplying by the number of patients per annum whose treatment choice is to be informed by the decision. This creates a distribution of the health consequences of uncertainty, which is expressed in terms of the primary outcome measure. Here we reflect the scale of the health benefits of the primary outcome (functional recovery in this case) by multiplying it by its NHB (15.86 QALYs for each functional recovery). The additional costs are then subtracted to reflect the empirical evidence on health opportunity costs to help understand the NHBs associated with uncertainty. The steps in this calculation are shown in Table 3.3.

recovery (Briggs et al., 2006). Utilising a Markov approach would bring this method closer to "full modelling".

Table 3.3: Rapid estimation of the health benefits of additional evidence in terms of QALYs for a primary outcome of functional recovery

	Functional recoveries per annum			QALYs per annum		Net health benefits of additional evidence	
	Late PTP [K]	Early PTP [L]	Additional treatment costs (=£14.1*8,800) [M]	Late PTP (=K*15.86) [N]	Early PTP (=L*15.86 - M/15,000) [O]	Absolute net effect for intervention in QALYs per annum (=H-D) [P]	Consequences of uncertainty for intervention (=-P if P<0) [Q]
Sample 1	4,745	1,871	£124,080	75,256	29,666	-45,590	45,590
Sample 2	4,923	3,499	£124,080	78,079	55,486	-22,593	22,593
Sample 3	4,569	4,763	£124,080	72,464	75,533	3,069	0
Sample 4	5,103	7,425	£124,080	80,934	117,752	36,818	0
Sample 5	4,836	7,609	£124,080	76,699	120,670	43,971	0
Average	4,835	5,033	£124,080	76,686	79,821	3,135	13,637

Table 3.3 shows that early PTP is expected to be superior after explicitly taking account of the health impact of the primary outcome and treatment costs. Again, there are samples in which early PTP is expected to be less effective than late PTP (samples 1 and 2). Taking the average consequences of uncertainty (column 8) we can estimate an upper bound for the benefits of research 13,637 QALYs per year. Repeating this process for a large number of samples results in an estimate of 9,169 QALYs per year. Extending the yearly consequences of uncertainty over the 15 year time horizon, the maximum value of research is estimated to be 107,000 QALYs gained over the full time horizon (after discounting). It is expected that it will take 5 years for the research to report and so the upper bound on the value of this research is expected to fall to 64,942 QALYs gained[20]. The rapid VOI approach in Chapter 2 was a decision tree in which the payoffs were quantified in terms of the primary outcome. The method presented here is also a decision in which the payoffs are linked to QALYs.

Health costs of using sub optimal treatments in research

As research will often involve allocating patients to each of the alternative treatments; this means that some patients will be allocated to the treatments which are expected to be sub optimal given current evidence (Briggs et al., 2006; Eckermann and Willan, 2008b). The net health loss associated with the suboptimal treatment may be subtle at the individual level but should be considered when estimating the health impacts of research proposals as they can have important consequences. This net health loss may result from allocating participants to treatments which are more costly and/or expected to be less effective. According to current evidence, early PTP is expected to improve the chance of functional recovery with a mean odds ratio of 1.09 and a 95% CI of 0.23 to 5.24. The probability of functional recovery with current practice is 55.1%. Combining the odds ratio for early PTP with the probability of functional recovery with current practice results in an expected probability of functional recovery of 56.5% for early PTP.

[20] It is worth noting that the value of research can be estimated in terms of QALYs by a method which is equivalent to the one demonstrated in this Section. The alternative approach is to calculate the value of research in terms of cost per primary outcome (as shown in Chapter 2) and then divide this by the NHB associated with the primary outcome. This approach assumes that the MCD is appropriate. Methods to determine an appropriate MCD are provided in Section 3.3.4.

As functional recovery is expected to be worth 15.86 QALYs the expected benefits of late PTP and early PTP are (55.1% x 15.86 =) 8.739 and (56.5% x 15.86 =) 8.961 QALYs respectively. Taking account of the additional costs of early PTP means the expected benefit of allocating a patient to early PTP is (8.961 - £14.1/£15,000 =) 8.960. This implies that for every additional patients not allocated to the optimal treatment (early PTP) the health system loses (8.960 - 8.739 =) 0.221 QALYs. Therefore, according to the design of P1, allocating 650 patients to late PTP as part of P1 is expected to be associated with net health losses of (0.221 x 650 =) 143.65 QALYs.

The "NHS support and treatment costs" are reported for each proposal in column 5 of Table 2.3. These are the costs of research which are borne by the general health system e.g. treatment costs, staff time and health losses associated with allocating patients to sub-optimal treatments during research[21]. From the calculation above the health opportunity losses associated with funding the trial are 143.65 QALYs. The total benefits of the proposed trial are the direct benefits of the trial minus the losses associated with allocating patients to sub-optimal treatments, (64,942 – 143.65 =) 64,798.35 QALYs. This means that the value of the trial can be expressed as one QALY gained per (£2,854,000/64,798.35 =) £44 of NETSCC expenditure.

The value of the trial is now expressed in a generic measure of health outcome, which can be compared to other proposed research competing for funding.

For every £15,000 spent on NHS service provision, the health system can expect to produce one QALY. By funding P1, NETSCC only has to spend £44 to produce one QALY. This means that P1 offers excellent value for money to the health system compared to general service provision. However, given that the budget for funding research is fixed and separate from general health care provision, the decision about whether the proposed trial represents good value to NETSCC depends on how it compares to other proposals competing for the same NETSCC funding.

[21] From the information available the methods used by applicants to calculate NHS support and treatment costs appear to only consider the additional financial costs associated with new treatments. A full estimate of the cost the NHS support and treatment costs is 143.65 x £15,000 = £2,154,750 in monetary terms, which is substantially greater than the £490,000 NHS support and treatment costs reported in the proposal documentation.

3.3 INFORMING KEY QUESTIONS IN RESEARCH PRIORITISATION

Table 3.4 presents the expected value of research for each of the six NETSCC proposals. The value of each proposal is expressed in terms of NETSCC cost per QALY gained (column 4) and as QALYs per £15,000 of NETSCC expenditure (column 5). The latter is a simple manipulation the former which can potentially make health impact more explicit. For example funding P1 is expected to be equivalent to paying £44/QALY, this is equivalent to generating (£15,000/£44 =) 341 QALYs per £15,000 of expenditure. The proposals are ordered from top to bottom by their total expected impact on population health.

The method to estimate the benefit of P1 in terms of QALYs has been illustrated in this chapter. Rapid VOI approaches have also been used to estimate the value of funding P2-P6 in terms of QALYs. The methods used are an extension of those illustrated previously. See Chapter 4 for analysis of P2-P5. See Chapter 5 for analysis of P6. Results are presented here primarily to illustrate how a QALY analysis can facilitate research prioritisation decisions.

Table 3.4: Expected value of proposed research study expressed in terms of QALYs for each of the six NETSCC proposals.

Proposal	Upper bound for health impact of research	Cost to NETSCC	NETSCC cost per QALY	QALYs per £15,000 NETSCC expenditure
P1: Brain injury	64,798 QALYs	£2,854,000	£44	341 QALYs
P4: Alzheimer's	967 QALYs	£3,310,883	£3,422	4.4 QALYs
P5: Psychosis	38 QALYs	£601,481	£14,806	1 QALY
P6: Information booklet	10 QALYs	£882,177	£86,836	0.2 QALYs
P3: Rare disease	707 QALYs lost	£855,403	NA	12.4 QALYs lost
P2: Melanoma	4,160 QALYs lost	£2,523,000	NA	24.7 QALYs lost
Total costs		£11,026,654		

Variation in expected value of research proposals

There is wide variation in the expected maximum value of research across the six proposals (column 2 in Table 3.4) with a range from 64,798 QALYs gained (P1) to 4,160 QALYs lost (P2). This means that different research proposals are expected to have very different contributions to population health (as measured in QALYs). The variation between proposals can be explained by a number of factors which combine to result in the ordering shown in Table 3.4 and Figure 3.1:

Incident population

As in Section 2.5 proposals which affect a large number of patients each year such as P4 (100,000/year) and P1 (8,800/year) tend to be associated with a large health impact. These proposals rank highly compared to P5 (1,563/year), P2 (1,137/year) and P3 (26.3/year). The exception to this is P6 which has a large incident population (259,150/year) but ranks below P5 which affects far fewer people. This is explained by the low levels of uncertainty in P6 which will be discussed next.

Degree of uncertainty in relative effects

P6 affects a large number of people (259,150/year) but has a low health impact, ranking below P1, P4 and P5. As discussed in Section 2.5 this is because the new intervention that P6 investigates is cheap and is expected to be beneficial, therefore there is little uncertainty to address. There is also little uncertainty in the clinical decisions investigated by P2 and P3 resulting in low value of research. In these cases the lack of uncertainty is due to large differences in relative treatment costs.

Relative treatment costs

P3 and P2 both investigate the use of reduced doses of highly expensive treatments. The substantial costs savings associated with these reduced doses regimens means that according to current evidence the cheaper alternative interventions are highly cost effective relative to current practice. Therefore, there is very little uncertainty about the optimal choice. Funding P3 and P2 is associated with negative QALYs as there is very little uncertainty and commissioning this research will prolong the use of the highly expensive current practice. Prolonging the use of expensive treatments results in QALYs lost from the opportunity costs of this additional health

expenditure[22]. In practice, there may be barriers to switching to more cost effective alternatives in the absence of commissioning new research. This is discussed in Section 3.4.

Net health impact of primary outcome

The magnitude of the net health consequences of the primary outcome is also significant as it reflects the relative importance of uncertainty in different disease areas. For example, the estimated net health impact of functional recovery in P1 is 15.86 QALYs compared to 1.06 QALYs per relapse avoided in P5. This means that achieving an additional functional recovery has a similar health impact as preventing approximately (15.86/1.06 =) 15 relapses. The estimated value of each research proposal reported in Table 3.4 is a result of the interplay between each of the factors above.

[22] Table 2.3 reports "NHS support and treatment costs" to be cost savings of £62,410,967 and £10,608,500 for P2 and P3 respectively. These alleged savings arise from allocating patients to new interventions which are substantially cheaper than current practice. For both P2 and P3 the new interventions are expected to be cost effective given current evidence, therefore these estimated cost savings are better thought of as cost increases as patients should be immediately switched to the more cost effective alternatives and so commissioning the research involves prolonging the use of the more costly current practice.

Informing key questions in research prioritisation

The results in Table 3.4 can be used to support explicit and evidence based discussion on a broad range of policy questions faced by research funders such as:[23]

- Which proposals represent "best buys" within the research funding budget?
- Are sufficient resources being devoted to research?
- How can charity or industry contributions towards research be considered?
- What change in the primary endpoint is required before practice should change?

3.3.1 WHICH PROPOSALS REPRESENT "BEST BUYS" WITHIN THE RESEARCH FUNDING BUDGET?

In order to rank the priority of a set of research proposals we apply the "bookshelf" approach described by Culyer (2016) and Remme et al., (2017) to the research prioritisation task. This allows us to rank research proposals from highest to lowest health impact and so decision makers can take account of which projects represent "best buys" within the research funding budget. Proposals in Table 3.4 are ranked by the QALYs gained per £15,000 of research expenditure (column 5), or equivalently, ranked by lowest NETSCC cost per QALY (column 4). Proposals with a higher QALY impact per £15,000 expenditure (e.g., P1) have the potential to increase population health more than those with a lower impact per £15,000 expenditure (e.g., P5) and, therefore, should be prioritised first.

The order of funding priority is P1 followed by P4, P5, and then P6. P2 and P3 are expected to result in negative population health per NETSCC expenditure; therefore, these proposals should not be funded at any level of the budget[24]. The set of research proposals that can be funded depends on the size of the budget available for research funding. The relevance of the available budget is illustrated in Figure 3.1. Each of the six proposals is represented by a column, where the height represents the number

[23] Prioritising a set of research proposals using a finite budget is similar to the task of constructing a package of interventions to fund in a health system. Therefore, the analysis here is similar to that described in Ochalek et al., (2018b).

[24] These projects only have value if practice cannot change in the absence of the research (see Section 3.4).

of QALYs per £15,000 of NETSCC expenditure, the width represents the cost of the proposal to NETSCC (also labelled within brackets) and the area of the column reflects the net health impact (i.e. the total payoff).

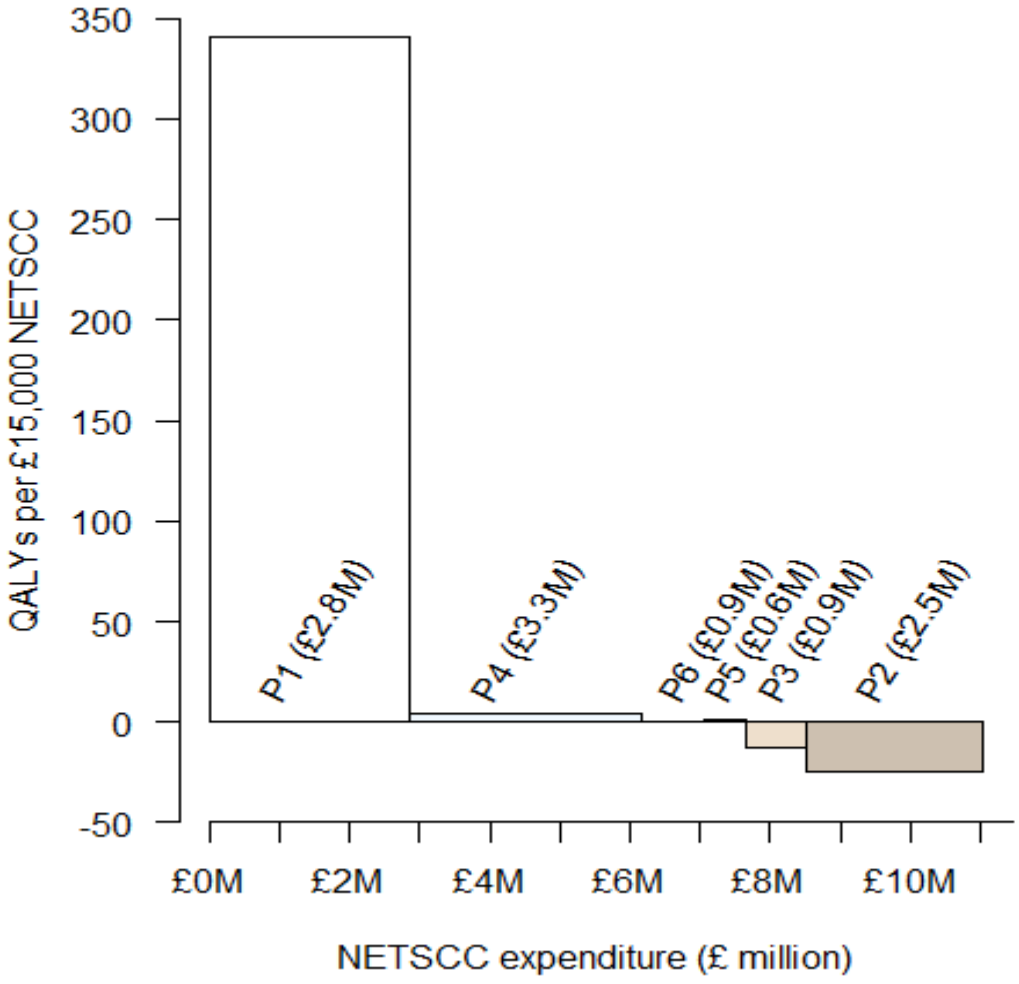


Figure 3.1: The net health impact and budget impact of each of the six NETSCC proposals

The horizontal axis shows the cumulative NETSCC budget. A budget of just under £3 million is required to fund P1, a budget of £6.1 million is required to fund P1 and P4, and so on, up to £7.6 million, at which point there are no worthwhile projects remaining.

How robust is this ordering?

A number of scientific and social value judgements are required in order to estimate the expected health benefits of research. These judgements are unavoidable and subject to uncertainty. Sensitivity analysis is used to assess the impact of changing the value of inputs on the expected value of research. It provides a quick and intuitive method to understand the impact that different judgements may have on the value of research and whether these are worth exploring further. A scenario analysis is presented here which constructs a high value and low value scenario for the expected value of research. This shows the effect of changing a group of inputs simultaneously, where each input is increased or decreased by 10%. Table 3.5 shows the inputs used in the scenario analysis and corresponding results for P1.

Table 3.5: The maximum expected value of research for P1 under high value, base case, and low value scenarios

	High value scenario	**Base case**	**Low value scenario**
Baseline Probability of functional recovery	146 out of 265 at risk	162 out of 294 at risk	178 out of 323 at risk
95% CI for odds ratio for functional recovery	0.21 to 5.76	0.23 to 5.24	0.25 to 4.72
Cost of early PTP per year	£136,488	£124,080	£111,672
Incidence	9,680	8,800	7,920
Time horizon for information	16.5 years	15 years	13.5 years
Duration of research	4.5 years	5 years	5.5 years
Research cost to NHS	£441,000	£490,000	£539,000
Cost per QALY	**£32**	**£44**	**£63**
QALYs per £15,000 NETSCC spend	**469**	**341**	**238**

Between the low value and high value scenarios, the cost per QALY varies between £32 and £63 per indicating that the results are relatively stable for P1. From Table 3.4 the nearest alternative to P1 is P4 with a cost per QALY of £3,422, therefore the P1 remains the top priority for funding in this sensitivity analysis. A similar exercise could be carried out for each of the remaining proposals. Because the estimated

value of proposals P5 and P6 are closer in value (1 and 0.2 QALYs per £15,000 NETSCC expenditure respectively) it is possible that their ordering in priority could change under some scenarios. Such uncertainty is unavoidable in decision making and is an important component of a deliberative approach to research prioritisation. It is important to emphasise that the results of sensitivity analysis should be interpreted in the context of a given a set of research proposals and a given budget constraint. As in the case of P1 sensitivity analysis may result in large variations in health impact but no change in the overall funding priority. In other cases it may be that small variations in health impact under sensitivity analysis result a substantial reordering. This is more likely in cases in which there are a number of research proposals with base case estimates of value which are in close proximity.

3.3.2 ARE SUFFICIENT RESOURCES BEING DEVOTED TO RESEARCH?

Empirical work estimates that approximately £15,000 of general NHS expenditure is required to generate one QALY in the UK (Claxton et al., 2015b; Lomas et al., 2018). Therefore, research projects which generate more than one QALY per £15,000 (such as P1 and P4) compare favourably to direct NHS health care provision. If there are research projects which are expected to result in more than one QALY per £15,000 expenditure that cannot be funded within the current research budget constraints, this indicates that research may be insufficiently funded relative to general health care expenditure. The value of expanding the research budget could be estimated by calculating the value of research of proposals which just missed out on funding due to resource constraints[25] (Culyer, 2016). These are the marginal projects which would have been funded if the research budget was larger. This could form an empirical basis for helping to understand the "marginal productivity" of the research budget i.e. how much research funding is required to produce one QALY. By comparing this to the marginal productivity of the general health care budget (approximately £15,000/QALY) this can inform discussions about the overall size of the research budget relative to general health care expenditure. If more health can be

[25] The methods described in this thesis provide an expected upper bound on the value of research when all uncertainty is resolved. As research projects have a finite sample size, they will only partially resolve the uncertainty. Methods to adjust research value for sample size and other aspects of research design are well developed (A. Briggs et al., 2006; Strong et al., 2015). Without this adjustment, the value of research would be overestimated relative to direct service provision.

gained by expanding the research budget than would be lost by shrinking the general health care budget, then population health may be improved by shifting resources towards research[26].

3.3.3 How can charity or industry contributions towards research be considered?

Publicly funded research projects may receive financial contributions from charities and/or industry. In this thesis we characterise research costs as falling on two types of budget; the research prioritisation body's budget (e.g. NETSCC in the UK) and the general health care budget (e.g. the NHS in the UK). The effect of savings from contributions to research funding will depend on which budget these savings accrue to. To illustrate, the Alzheimer's society has offered to contribute to the costs of funding P4. If £1 million was granted directly to NETSCC to help fund P4, then the cost to NETSCC falls from £3,310,883 to £2,310,883 and, therefore, the NETSCC cost per QALY falls to (£2,310,883/967 =) £2,390 per QALY. It is interesting to note that though the contribution from the Alzheimer's society does improve the relative value of P4 it does not change its overall priority, it remains the second best value for money after P1, but it does free up NETSCC resources to fund other potentially worthwhile research.

If the cost savings were allocated to the NHS to help cover the additional support and treatment costs associated with the research (for example by directly contributing to the costs of drug acquisition), then NHS support and treatment costs fall by £1M. This improves the health benefits of research by (£1M/£15,000 =) 66.7 QALYs due to cost savings which results in total net benefits of research of (967 + 66.7 =) 1,034 QALYs and therefore a NETSCC cost per QALY of (£3,310,883/1,034 =) £3,202 per QALY. In this case contributing to cover NETSCC costs improves the relative value of P4 more than contributing to cover the NHS support and treatment costs.

[26] For large changes in the research budget and/or the general health budget the effect of non-marginal changes must be accounted for (Lomas et al., 2018).

3.3.4 DETERMINING MCD: TREATMENT PRICE AND THE CHANGE IN PRIMARY ENDPOINT REQUIRED BEFORE PRACTICE SHOULD CHANGE

Chapter 2 illustrated how the value of research will depend on the MCD required for practice to change. Informing the appropriate MCD requires an implicit trade-off between the health benefits of the primary outcome and the other outcomes of interest which are not captured by the primary endpoint (the secondary outcomes). In this section we provide a method to help inform an appropriate MCD which uses the explicit linkage of the primary outcome to QALYs.

Consider the comparison of a new intervention to a baseline comparator. In addition to differences in the relative effect of each treatment on the primary outcome there may be a number (n) of relevant secondary outcomes to consider ($i = 1,\ldots n$). The appropriate overall MCD for the new treatment reflects the improvement in the primary outcome required to make up for net losses (if any) associated with the secondary outcomes relative to the baseline comparator. These secondary outcomes may have countervailing effects, for example the new intervention may be more expensive but have a better side effect profile relative to current practice. Here we provide a heuristic for informing overall MCD for a treatment. This heuristic applies in health systems which are concerned with impacts on total population health[27].

When considering effects on population health, the MCD reflects the additional number of primary outcomes which are required to make up for the health losses (if any) associated with secondary outcomes[28] e.g. side effects. The number of additional primary outcomes required to make up for the health loss associated with a secondary outcome will depend on the incremental NHB (INHB) associated with the primary outcome. A larger MCD is required if secondary outcome i is associated with a large net health loss (NHL_i). QALYs can be used to facilitate the required trade-off.

[27] Other approaches to resource allocation which attempt to prove a basis for resource allocation in insurance based systems through utility maximisation using welfare economics (Basu and Sullivan, 2017) also involve trade-offs between primary and secondary outcomes. Therefore heuristics to define the required MCD for a new treatment to enter the insurance program can also be developed for these cases. This should reflect the consumption losses of increased premiums and/or the losses from disinvestment in another treatment. Currently the methods outlined here assume linearity in the trade-off between primary and secondary outcomes which may need to be relaxed in a utility maximisation framework.

[28] The primary outcome may be harmful, e.g. heart attacks, in which case the MCD is defined as the additional number of primary outcomes which must be avoided to make up for the health losses associated with the secondary outcomes.

For an individual patient the change in the probability of the primary outcome required to make up for secondary outcome i is represented by $MCD_{i.}$ The required MCD_i is such that the health gains in the primary outcome cancel out the health losses from the secondary outcome[29]:

$$\mathrm{MCD}_i.\,\mathrm{INHB} - \mathrm{NHL}_i = 0$$

Rearranging gives $MCD_{i.} = NHL_i/INHB$. For example, early PTP is associated with additional treatment costs (ΔC) of £14.1 per person and this is the only relevant secondary outcome (n = 1). Given the marginal productivity of the NHS (k) of £15,000/QALY, these additional treatment costs are associated with a NHL of ($\Delta C/k$ = £14.1/£15,000 =) 0.00094 QALYs per person. From Section 3.2, each additional functional recovery is associated with an additional 15.86 QALYs, therefore only (0.00094/15.86 =) 0.00006 additional functional recoveries are required to make up for the additional costs. This translates to a required MCD of 0.006%, the effect size of the new treatment must be greater than this to make up for the additional costs. The MCD is approximately zero as the additional costs of the new treatment are small relative to the health gains from the primary endpoint. The relationship between MCD and additional treatment costs (ΔC) is summarized in Equation 3.1 below.

$$\mathrm{MCD}_1 = \frac{\mathrm{NHL}_1}{\mathrm{INHB}} = \frac{\Delta C}{k.\mathrm{INHB}} \qquad (3.1)$$

The MCD formula shows that higher relative costs (ΔC) require a larger MCD. Larger relative costs may be due to higher prices or increased dosage. A larger MCD is required in these cases because costs will result in health opportunity costs and larger improvements in the primary endpoint are required to compensate for this foregone health. A larger k implies that health care costs do not displace as much health and so a larger k will reduce the required MCD. If the primary endpoint is associated with large NHBs then this will also reduce the required MCD as small changes in the primary outcome have large health consequences. If a new treatment has a lower price/dose than current practice (and has a similar side effect profile) then ΔC will be negative reflecting the fact that the secondary outcomes associated

[29] It should be noted that this formulation imposes a structural relationship between the MCD the primary outcome and the secondary outcomes. This simple relationship may not be appropriate in all cases but is a necessary simplifying assumption to keep the complexity of the analysis down.

with the new treatment are favorable on net. From Equation 3.1 this implies a negative MCD which will make the adoption of the new treatment more likely (see Chapter 6 for more details). Any decision, requiring minimum changes in primary endpoints makes implicit judgements about ΔC, k and INHB, each of which can be usefully informed with reference to external evidence.

The above analysis only considers price as a relevant difference between the treatments, however the MCD implicitly proxies the health effects of all other differences between treatments. The appropriate overall MCD is the sum of each of the individual effects of all relevant secondary outcomes (i = 1,…n) as shown:

$$\text{MCD} = \frac{NHL_1}{INHE} + \cdots + \frac{NHL_n}{INHE} \qquad (3.2)$$

3.4 FUNDING RESEARCH TO CHANGE CLINICAL PRACTICE

There are two distinct ways to improve health outcomes. The first is to implement the findings of existing evidence (implementation value), while the second is to fund research to reduce uncertainty about the health impact of alternative treatments (information value). For some clinical decisions, the methods described here may indicate that current practice is not the most cost-effective and there is very little or no uncertainty about this judgement. As there is no uncertainty there is no value in further research. However, in special circumstances research may be necessary to change current practice. This scenario can occur in situations in which key inputs such as relative treatment effects have been informed using indirect sources of evidence such as meta-epidemiological evidence or expert elicitation. In these situations clinicians, legal and/or professional bodies may demand a minimum amount of direct empirical evidence from RCTs or well conducted observational research before a change in practice is warranted. If practice cannot change in absence of research then the value of the research is the implementation value plus the information value. If the information value is zero the effect of the trial will be dominated by its effects on practice. As gathering information is no longer the objective smaller trials will have larger health impacts as they will report quicker and so can change practice quicker.

For example P2 describes research to investigate the effect of reducing the use of expensive immunomodulating anti-programmed death receptor 1 (PD1) antibodies in patients with advanced (unresectable stage III, IV) melanoma who are due to start anti PD1 as first line treatment. The primary outcome in the trial is the effect on progression free survival (PFS), with a non-inferiority margin defined for reduced doses of PD1 antibodies. The proposed trial was designed to compare continuous treatment until disease progression (TTP) to treatment for 12 months (TF12) and treatment for 6 months (TF6). Both of the new interventions TF12 and TF6 come with substantial expected cost savings. No relevant empirical evidence (RCTs or observations studies) was reported which investigated the use of TF12 or TF6 in this population. In absence of direct evidence, meta-epidemiological evidence was used as a starting point to understand the range of outcomes that may be plausible with TF12 or TF6. Due to the significant health opportunity costs resulting from the high prices of PDI antibodies, the TF6 dose reduction is expected to be superior to current

practice and there is no uncertainty in this decision. Due to prolonging the use of TTP during the research and the absence of any uncertainty in the decision, funding P2 is expected to reduce population health (as reported in Table 3.4 and Figure 3.1).

Though TF6 appears to be the most cost effective option, there is no direct evidence on the use of TF6 in this population and the current (high cost) practice is mandated by NICE. Therefore practice may be difficult to change in absence of direct evidence. As TF6 is more cost-effective than current practice, health is foregone by delaying the switch to this treatment. Table 3.6 shows that the value of changing practice declines over time as the number of patients who can potentially be provided the cost effective alternative gets smaller[30]. After 10 years it is assumed that the treatment will no longer be relevant due to innovation and the value of changing practice falls to zero.

Table 3.6: Benefits of changing practice over time for P2 in terms of QALYs.

Year	Discounted vale of changing practice per year (QALYs)	Total value of changing practice over time (QALYs)
1	11,463	98,423
2	11,069	86,960
3	10,688	75,891
4	10,321	65,203
5	9,966	54,882
6	9,623	44,917
7	9,292	35,294
8	8,972	26,002
9	8,664	17,029
10	8,366	8,366
11	0	0

As the value of changing practice declines over time, any trial which is commissioned with the express purpose of providing a minimum standard of direct empirical evidence should be as small as possible so as to minimise this delay and report quickly[31].

[30] This table is calculated by (continuously) discounting the implementation value per year (11,665) over the 10 year time horizon (column 2) and calculating the cumulative sum from year 10 back to year 1 (column 3).

[31] Unbalanced research designs which preferentially enrol individuals to the expected optimal treatment will also be favoured.

The research design as described in the proposal requires a large number of patients (n = 2,025) and so will facilitate changing practice only after it reports in 6 years. From Table 3.6, this change in practice from TTP to TF6 after 6 years is expected to result in an additional 44,917 QALYs through cost savings. Considering the health losses associated with delaying the switch to TF6, a smaller research design (say, n = 300) which would report within approximately 2 years may be more appropriate[32]. In addition to being cheaper for NETSCC to fund, by facilitating changing practice at 2 years the small trial is worth approximately 86,960 QALYs meaning it dominates the original design. The policy implication is that if practice *cannot* change in absence of research then a smaller design than P2 would result in greater health impact.

The requirement of minimum levels of empirical evidence (i.e. the unwillingness to change practice based on indirect methods such as meta-epidemiological evidence or expert elicitation), such that it exists, may be due to professional discomfort and/or ethical obligations to patients to not expose them to untested interventions. As the research budget is limited, using the research budget as a method to change practice will necessarily divert resources away from research projects which could address genuine uncertainties in the health system. Therefore, in addition to the ethical concerns, the health losses associated useful research foregone must also be considered.

In many cases the other mechanisms and policies such as incentives and sanctions will be more appropriate to affect implementation of treatments which are expected to be cost effective. Indeed, the analysis of P2 strongly suggests that immunomodulating anti PD1 antibodies should not have been approved for use with a TTP dosing schedule at the current price. A more appropriate action from the reimbursement agency would have been to reject the treatment at the offered price and treatment schedule. This action would have maintained the incentive for the manufacturer to either reduce the price, carry out research on the less intensive treatment regimes, or both carry out research and reduce the price (Claxton et al., 2012; Griffin et al., 2011; Rothery et al., 2017).

[32] This smaller trial must enrol only 300 individuals. This is approximately (300/2,025 =) 15% the size of the original design. Therefore the cost to NETSCC is assumed to be (£2,522,710 x 15% =) £378,407, it is expected to take approximately (6 x 15% =) 10.8 months to report. This assumes a constant rate of marginal costs and enrolment, more sophisticated estimates of cost and duration are possible.

3.5 DISCUSSION

This chapter has shown that quantification of health benefits in terms of QALYs is a feasible extension to the methods introduced in Chapter 2. The P1 case study was used to illustrate how to link binary primary outcomes to costs and QALYs using the method described in (McKenna et al., 2016). A limitation of this method is that it can only be used for binary primary outcomes and so is insufficient to address all of the NETSCC proposals, Chapter 4 extends these methods to link continuous and survival outcomes to costs and QALYs.

This chapter also demonstrated how quantifying the benefits of research in terms of a generic measure of health outcome (e.g. QALYs) can facilitate the consideration of a number of questions of relevance to research prioritisation bodies. A central contribution of this thesis is to address research prioritisation questions through a novel application of the "bookshelf" approach. This captures the resource constraints faced by research prioritisation bodies and provides a framework to inform the key questions in research prioritisation.

The bookshelf approach allows for identification of "best buys" by explicit ranking the available research projects by a comparable health outcome. Decision makers may have a range of considerations other than health maximisation such as reducing health inequalities, or maintaining equity in funding across disease areas. However, quantitative estimates of research benefit can help policy makers understand whether the trade-offs involved are worth making and provides a framework to communicate their decisions to stakeholders (Ochalek et al., 2018b).

The bookshelf also facilitates calculation of the marginal productivity of the research budget. The marginal productivity of the research budget is the amount of additional research funding which is required to produce one QALY (or some other measure of benefit). Recent years have seen a flowering of marginal productivity research which aims to empirically estimate the health effect of direct service provision in health care (e.g. general NHS activity). This has been carried out in the UK (Claxton et al., 2015b; Lomas et al., 2018), Spain (Vallejo-Torres et al., 2018), Australia (Edney et al., 2018) and estimates have been produced for 98 low-income and middle-countries (Ochalek et al., 2018a). Estimates of the marginal productivity of the general health budget can be used make decisions about which treatments to fund as it facilitates

judgements about whether gains in health from the treatment are expected to be greater than health gains that would have been possible if the additional resources required for the treatment had, been used in other healthcare services. As the budget for direct service provision (e.g. NHS) and the health research budget (e.g. NETSCC), in principal, compete for funding, estimates of their respective marginal productivities can also be used to decide whether more resources should be devoted to research rather than to direct patient care (Claxton et al., 2007; Remme et al., 2017).

The consequences of charity or industry contributions towards research have also been explored in this chapter. A framework is provided for taking account of these contributions when prioritising research. From the perspective of a research prioritisation body, these contributions should be subtracted from the original costs of the research as they do not fall on the research budget. From a the perspective of a charitable organisation such as the Alzheimer's society, any contribution towards research in will come with opportunity costs i.e. the money given to research could have been used in another charitable Alzheimer's related activity such as in providing psychological support to families. This resource allocation decision faced by charities requires a comparison of "information value" and "implementation value" as discussed in Chapter 2. Therefore VOI methods can also be used by charitable bodies to aid decision making.

Section 3.3.4 provides a method to help inform an appropriate MCD which uses the explicit linkage of the primary outcome to QALYs. This represents an important development of the MCD concept from the work of Claxton et al., (2015a) and McKenna et al., (2016) as it provides a consistent link between MCD, the opportunity costs of health expenditure, price and the net health benefits associated with the primary outcome.

In addition to providing quantitative estimates of the value of research, applying VOI methods to the research prioritisation task also provides a consistent framework for decision making. In Section 3.4 this framework is used to conceptualise the implications of using research to change practice. The approach assumes that funding a trial will have an effect on the use of treatments in the health system meaning the value of the research is entirely "implementation value" (Fenwick et al.,

2008). This provides an upper bound for the value of funding this research to change practice as it represents the value of practice immediately changing to the most cost effective treatment option after the trial reports. In reality, it takes time for practice to respond to research. Grimm et al., (2017) extending Willan and Eckermann, (2009) provide an extended framework which explicitly models the effect of study results on the diffusion of adoption decisions through the health system. This takes account of the counterfactual case in which the study was not carried out. This approach provides a more realistic characterisation of the decision however the additional modelling requirements for each proposal may be prohibitive for resource constrained decision makers. Future research should investigate the effects of considering diffusion and, should they prove significant, develop rapid methods which could incorporate these concerns without a large analytical burden.

The methods described in this thesis to understand the value of research primarily focus on the value of research in reducing uncertainty in relative treatment effects e.g. the odds ratio[33]. However, there are a number of other inputs into decision making which may be uncertain. For example, when linking primary outcomes to costs and QALYs there may be uncertainty about the magnitude of health gains and cost savings associated with a primary outcome. The sensitivity analysis such as that carried out in Section 3.3 partially addresses this issue. Sensitivity analysis can demonstrate how the value of research changes with different assumptions about how the primary outcome relates to costs and QALYs, however, it does not provide estimates of the value of reducing uncertainty about this relationship. This is important as understanding the value of reducing different aspects of decision uncertainty can be used to guide the prioritisation of research which is appropriate to resolve the most relevant uncertainties.

[33] For binary endpoints, uncertainty in the baseline probability of the primary outcome is also considered.

Chapter 4

Generalising methods for rapid assessment of need for evidence

4.1 INTRODUCTION

Chapters 2 and 3 have illustrated how to rapidly estimate the value of comparative effectiveness research for a binary primary outcomes using the methods introduced by Claxton et al., (2015a) and McKenna et al., (2016). These methods can be used to rapidly estimate the value of research in terms of cost per primary outcome or cost per generic health outcome (e.g. QALY). It has also been shown how these methods can assist decision makers in addressing the key tasks of research prioritisation. However, these methods are directly applicable to only two of the six research proposals provided by NETSCC (P1 and P6). The remaining four proposals have features which are outside of the scope of the methods in their existing form. The current methods are limited in two respects; first, they address only comparative effectiveness research i.e. research which identifies which interventions work best for improving health. Second, the methods can only be applied to binary primary outcomes. The aim of this chapter is to extend these methods of Claxton et al., (2015a) and McKenna et al., (2016) to provide decision makers with toolbox of methods which can be rapidly applied to feasibility/pilot studies and a wider range of primary outcomes types.

P5 is for a feasibility study which cannot be analysed as comparative effectiveness research. Though there are subtle differences in the definitions of "feasibility studies" and "pilot studies"[34], both involve carrying out a small initial study to determine whether a larger comparative effectiveness research project (a "full trial") is possible. These study types are an important component of modern health research and so methods are required to understand their value and avoid waste (Chalmers et

[34] From the NIHR website (2017): 'Feasibility Studies are pieces of research done before a main study in order to answer the question "Can this study be done?" They are used to estimate important parameters that are needed to design the main study. For instance: standard deviation of the outcome measure, willingness of participants to be randomised etc. Pilot studies are a smaller version of the main study used to test whether the components of the main study can all work together. It is focused on the processes of the main study, for example to ensure that recruitment, randomisation, treatment, and follow-up assessments all run smoothly. It resembles the main study in many respects, including an assessment of the primary outcome.'

al., 2014; Morgan et al., 2018). If funding for feasibility/pilot studies and full trials come from the same research budget then quantitative estimates of the health benefits of feasibility/pilot studies are required in order to understand their value relative to full trial proposals. P4 utilises a continuous primary outcome, P2 utilises a survival (time to event) primary outcome and P3 utilises a binary primary outcome for which the costs of treatment depend on whether the outcome occurs or not.

In this chapter we first outline a rapid method to link a binary primary outcome to costs and QALYs when the costs of treatment depend on whether the outcome occurs or not. This is an extension of the methods demonstrated in Chapter 3 and is illustrated using the P3 case study. Second, we outline a rapid method to calculate the value of research in terms of both cost per primary outcome and cost per QALY for continuous primary outcomes using P4 as a case study. Third, we do the same for survival primary outcomes using P2 as a case study. Fourth, we demonstrate a rapid method to estimate the value of a given feasibility study using P5 as a case study throughout.

The analysis for each case study was carried out using the RANE tool (formally introduced in Chapter 5). The full set of inputs required to reproduce the analysis for all case studies is available in the Appendix A1-A6.

4.2 RESEARCH PRIORITISATION WITH DIVERSE PRIMARY OUTCOMES

4.2.1 RESEARCH PRIORITISATION WHEN TREATMENT COSTS DEPEND ON THE BINARY OUTCOME

The methods from Chapter 3 show how to rapidly calculate the benefits of research for binary primary outcomes in terms of QALYs. The method used assumed that treatment costs were independent of the primary outcome. For example, in the P1 case study, the additional costs associated with early treatment (£14.1 per person) are not affected by whether functional recovery occurs or not i.e., they are independent of the primary outcome. However, for P3, this assumption is not appropriate. This is because treatment with the baseline therapy is continuous unless the primary outcome occurs, at which point the treatment may be discontinued. Therefore, if the primary outcome occurs the treatment costs may no longer be incurred. Rapid methods to address this more complex case are described and demonstrated here using P3 as a case study.

4.2.1.1 *Overview of P3*

This proposal is for a non-randomised, open label study to examine the safety of withdrawing Eculizumab (Soliris) in patients with atypical haemolytic uraemic syndrome (aHUS). aHUS is a disease typically causing acute kidney injury, but also damage to other organs. It is a rare disease (0.4 cases/million/year in the UK) but it has a profound impact on patients, with 50% of patients dying or developing renal failure within 1 year of diagnosis (Fremeaux-Bacchi et al., 2013), and associated with a high risk of recurrence after transplantation (Noris and Remuzzi, 2010). Eculizumab has been shown to be effective in the treatment of aHUS, both for the induction and maintenance of remission. Lifelong treatment with Eculizumab is currently recommended because of risk of relapse. However, there is limited safety data on long-term use and Eculizumab is associated with significant treatment costs (£340,000/patient/year). The proposed trial aims to answer whether individuals with aHUS receiving Eculizumab treatment can safely withdraw from treatment following supervised withdrawal.

Summary of proposal 3

Research question: Can individuals with Atypical haemolytic uraemic syndrome (aHUS) receiving Eculizumab treatment safely withdraw from treatment?

Intervention: Supervised withdrawal of Eculizumab infusions every 2 weeks (high cost) replacing it with weekly blood tests (month 1), fortnightly blood tests (months 2-6) then monthly thereafter and home urinalysis monitoring (low cost). Reintroduction of Eculizumab if relapse occurs.

Control: Continuous treatment with Eculizumab.

Primary outcome: Number of serious withdrawal attributable adverse events (SAE) during a 2 year follow up (death, meningococcal infection or irreversible organ injury).

Proposed study: Non-randomised, open label study (withdrawal, n= 30; continuous treatment, n =20)

Duration of proposed study: 4 years.

Costs of proposed study to NETSCC: £855,403

NHS support and treatment costs: £10,608,500 cost savings

Box 4.1: Summary of proposal 3

4.2.1.2 *Health benefits of further research in terms of primary outcome*

The primary outcome in P3 is a composite of serious adverse events (SAE) composed of; death, meningococcal infection or irreversible organ injury. The benefits of research can be quantified in terms of SAE avoided using the same methods as described in Chapter 2. From the proposal the probability of SAE with continuous treatment is between 5% and 6%. As shown in Chapter 2, the probability of SAE with the new treatment is determined by the baseline probability combined with an estimate of the relative effect. There currently exists no evidence to inform this and so in the absence of direct evidence meta-epidemiological data is used as a starting point to understand uncertainty. As the primary outcome (SAE) is harmful it is expected that the odds ratio is likely to be between (95% CI) 4.39 and 0.19 (see Section 2.4.4 for further details).

The health consequences of uncertainty can be understood in terms of SAEs in the same manner as described in Chapter 2: by drawing random samples from the distribution of the relative effect (CI around the odds ratio) and combining these samples with the baseline event rate. This is combined with; incidence (26.3 individuals per year), the time the information is expected to be valuable for (10 years), the time it will take the research to report (4 years) and the cost of the research (£855,403). All inputs used are provided in Appendix A6. With these inputs, the upper bound for the health benefit of the proposed research is estimated to be 1 SAE avoided over the full time horizon. The proposed research is expected to cost the research funder £855,403 meaning the maximum value of the proposed research is estimated to be (£855,403/1 =) £570,510 per SAE avoided.

4.2.1.3 *Linking the primary outcome to a comprehensive measure of health outcome*

In order to translate this outcome into costs and QALYs to aid research prioritisation, the substantial cost differences between withdrawal and continuous treatment must be reflected. As the SAE includes death, treatment costs will no longer be incurred for those individuals. Therefore, treatment costs will depend on the primary outcome. The methods developed previously have assumed that treatment costs do not depend on the primary outcome and so an extension is required. This extension is

very similar to the approach described in Chapter 3 and is illustrated here in three steps. In step 1, the primary outcome (SAE) is linked to costs and QALYs. In step 2, expected treatment costs are linked to the primary outcome. In step 3, these are combined to link uncertainty in relative effects to costs and QALYs. The only difference between this method and the one outlined in Chapter 3 is that the treatment costs depend on whether the primary outcome occurs or not.

4.2.1.4 *Step 1: Link primary outcome to costs and QALYs*

If the primary outcome occurs (i.e. a patient has a SAE), there is a chance of the individual entering one of three different health states; death, meningococcal infection or irreversible organ injury. Using the same approach as in Chapter 3, in this step we estimate the health consequences of the primary outcome by attaching the clinical states to cost and quality of life outcomes.

Each of the three health states (death, meningococcal infection and irreversible organ injury) have different disease related costs and health consequences. For illustrative purposes, we assume that there is an equal chance (33.3%) of each of the events occurring conditional on the individual experiencing a SAE[35]. If the individual dies, we assume that this happens near the start of the trial and there are no additional disease related costs associated with this. If meningococcal infection occurs, we assume that individuals will live for the full additional life expectancy of 35.47 years *(National Institute for Health and Care Excellence*, 2015) but will have the lower health state utility of 0.2 from Christensen et al., (2013). According to the same source they will also incur disease related costs associated with time spent in hospital (£2936.2) and a follow up appointment (£279.98). If irreversible organ injury occurs we assume this will require kidney transplant costing £17,000 in the first year and £5,000 in subsequent years (*NHS England*, 2013). Individuals are assumed to survive for their remaining 35.47 years of life, resulting in a total cost of £115,503 after discounting. We assume that utility in this health state will be 0.59, which is the average UK utility in diseases of kidney and ureters (Sullivan et al., 2011). Table 4.1 illustrates how these values are combined to help understand the health impact of the primary outcome in terms of costs and QALYs.

[35] This is an assumption made to quickly illustrate the P3 case study. This judgement could be informed with reference to published studies which report the proportion of patients who experience death, meningococcal infection and irreversible organ injury.

Table 4.1: Expected costs and QALYs associated with the primary outcome; serious adverse event. This is a composite outcome of death, meningococcal infection and irreversible organ injury.

Primary outcome occurs (serious adverse event)						
Possible states	**Probability of being in state**	**Life expectancy (years)**	**Utility weights**	**Disease related costs**	**Expected QALYs**	**Expected costs**
Death	33.3%	0	0	£0	0	£0
Meningococcal infection	33.3%	35.5	0.2	£3,216	2.36	£1,071
Irreversible organ injury	33.3%	35.5	0.59	£115,503	6.97	£38,462
	100%				9.33	£39,533

If the primary outcome does not occur we assume that individuals will live for the full additional life expectancy of 35.47 years. We also assume again that utility in this health state will be 0.59 which is the average UK utility in diseases of kidney and ureters (Sullivan et al., 2011). Other than the treatment costs, we assume no other disease related costs. This is illustrated in Table 4.2.

Table 4.2: Expected costs and QALYs when the primary outcome does not occur and the patient's illness proceeds naturally.

Primary outcome does not occur (no serious adverse event)						
Possible states	**Probability of being in state**	**Life expectancy (years)**	**Utility weights**	**Disease related costs**	**Expected QALYs**	**Expected costs**
Normal illness	100%	35.47	0.59	£0	20.93	£0
	100%				20.93	£0

Comparing these two possibilities, each additional SAE is expected to result in (20.93 – 9.33 =) 11.6 QALYs lost per person and £39,533 in additional costs. Taking account of the opportunity costs of health expenditure implies that the cost increases per SAE are expected to be worth (£39,533 /£15,000 =) 2.64 elsewhere in the health system (Claxton et al., 2015b). Combining the direct health loses (11.6 QALYs) and the indirect health loses through cost increases (2.64 QALYs) means that each additional SAE results in a loss of 14.24 QALYs.

4.2.1.5 *Step 2: Link treatment costs to primary outcome*

The therapies considered in this trial are: continuous treatment (baseline) and withdrawal. In contrast to the method described in Chapter 3, for each option we require the treatment costs when the primary outcome occurs and the treatment costs when it does not. For the continuous treatment, we assume that if the primary outcome does not occur, individuals will continue receiving Eculizumab for the expected additional lifetime (35.47 years). With yearly costs of £340,000 this results in total expected costs of £7,316,623 per person after discounting[36]. As discussed above, if the primary outcome occurs, there is a chance of the individual entering one of three health states; death, meningococcal infection or irreversible organ injury. These health states are associated with different treatment costs and these should be reflected in the analysis. For continuous treatment, if the primary outcome occurs we assume treatment will cease for those who die but will continue for all those who survive. The relationship between continuous treatment costs and the primary outcome is illustrated in Table 4.3.

Table 4.3: Expected treatment costs associated with continuous treatment. Treatment costs depend on whether the primary outcome (SAE) occurs or not.

Continuous treatment costs (baseline)				
Primary outcome	**Possible states**	**Probability of being in state**	**Treatment costs (per year)**	**Expected treatment costs (discounted)**
No SAE	Normal illness		£340,000	**£7,316,623**
SAE occurs	Death	33.3%	£0	£0
	Meningococcal infection	33.3%	£340,000	£7,316,623
	Irreversible organ injury	33.3%	£340,000	£7,316,623
	Expected costs if SAE occurs			**£4,877,749**

SAE: serious adverse event.

If the primary outcome occurs, continuous treatment with Eculizumab is expected to cost (£0 + £7,316,623 + £7,316,623)/3 = £4,877,749 per individual. If SAE does not occur then the expected treatment costs are £7,316,623. This means that for

[36] It should be noted that this assumes that this treatment will continue to be used for the remainder of the patient's lifetime. However, new treatments may enter which increase competition or change best practice. There are a number of uncertainties involved in long term forecasts and so it is difficult to know in advance how these factors would affect treatment costs.

continuous treatment, treatment costs are lower when the primary outcome occurs due to the fact that treatment costs will not be incurred if the patient dies.

For the withdrawal treatment option, if the primary outcome does not occur, we assume that each individual will receive the equivalent of one year of Eculizumab during the tapering off period[37]. Furthermore, the proposal states that after withdrawal approximately 50% will successfully withdraw from treatment for the remainder of their lifetime with the remaining 50% returning to treatment[38].

With yearly costs of £340,000 this results in total expected costs of £340,000 + 0.5 x (35.47 x £340,000) = £6,369,900 per individual and £3,965,432 per individual after discounting. Again, if the primary outcome occurs treatment will cease for those who die. Continuous treatment will be reinstated in those who experience meningococcal infection or organ injury (£7,316,623). The relationship between continuous treatment costs and the primary outcome is illustrated in Table 4.4.

Table 4.4: Expected treatment costs associated with withdrawing treatment. Treatment costs depend on whether a serious adverse event (SAE) occurs or not.

Withdrawal treatment costs (new treatment)				
Primary outcome	**Possible states**	**Probability of being in state**	**Treatment costs (per year)**	**Expected treatment costs (discounted)**
No SAE	Normal illness		£340,000	**£3,965,432**
SAE occurs	Death	33.3%	£0	£0
	Meningococcal infection	33.3%	£340,000	£7,316,623
	Irreversible organ injury	33.3%	£340,000	£7,316,623
	Expected costs if SAE occurs			**£4,877,749**

SAE: serious adverse event.

[37] To inform decision making this should be justified with reference to empirical evidence and/or expert opinion.

[38] Though not discussed in the proposal this judgement is likely to be uncertain. A more complex analysis would take account of the uncertainty in the number of patients who can successfully withdraw. Including this uncertainty would increase the overall decision uncertainty and increase the value of research.

If the primary outcome occurs, the withdrawal of Eculizumab is expected to cost (£0 + £7,316,623 + £7,316,623)/3 = £4,877,749 per individual.

In summary, if the primary outcome occurs both continuous treatment and withdrawal are expected to cost £4,877,749 per individual. If the primary outcome does not occur continuous treatment is expected to cost £7,316,623 per individual and withdrawal is expected to cost £3,965,432. From this it is clear that there are substantial cost savings associated with withdrawal.

4.2.1.6 *Step 3: Linking uncertainty in relative effects to costs and QALYs*

The uncertainty in the expected number of primary outcomes per year can be characterised in the same manner as described in Chapter 2: by drawing random samples from the distribution of the relative effect (CI around the odds ratio) and combining these samples with the baseline event rate. To get an indication of the health consequences of uncertainty in terms of a more comprehensive measure of outcome, the health consequences of the primary outcome and the treatment costs must be linked to the probability of SAE. A simplified example is shown for illustrative purposes. Table 4.5 displays five samples for the probability of SAE with each treatment.

Table 4.5: Probability of primary outcome with continuous treatment and withdrawal of treatment

	Probability of serious adverse event	
	Continuous treatment [A]	Withdrawal of treatment [B]
Sample 1	5.0%	1.0%
Sample 2	5.1%	2.7%
Sample 3	5.5%	5.0%
Sample 4	5.6%	18.8%
Sample 5	5.6%	20.7%
Average	5.4%	9.6%

Taking the continuous treatment as an example, treatment costs are calculated by multiplying the probability of SAE by the treatment costs when SAE occurs

(£4,877,749) and the probability of SAE not occurring by the treatment costs when SAE does not occur (£7,316,623). This is illustrated in Table 4.6.

Table 4.6: Expected treatment costs for continuous treatment. Treatment costs depend on the probability of the primary outcome occurring.

	Probability of serious adverse event (SAE)	Treatment costs for continuous treatment	
	Continuous treatment	SAE occurs (=A*£4.9M)	SAE does not occur (=(1-A)*£7.3M)
	[A]	[C]	[D]
Sample 1	5.0%	£243,887	£6,950,792
Sample 2	5.1%	£248,765	£6,943,475
Sample 3	5.5%	£268,276	£6,914,209
Sample 4	5.6%	£273,154	£6,906,892
Sample 5	5.6%	£273,154	£6,906,892
Average	5.4%	£261,447	£6,924,452

There are health opportunity costs associated with the increased treatment costs and these can be calculated for each sample by dividing by the opportunity cost of health expenditure (£15,000) (see Table 4.7, column E). From step 1, each additional SAE results in a loss of 14.24 QALYs. The expected NHBs for each sample can be calculated by multiplying the probability of SAE by -14.24 and then subtracting the health opportunity costs of treatment (see Table 4.7, column F).

Table 4.7: QALY effects of continuous treatment, taking account of the opportunity cost of health expenditure

	Probability of SAE with continuous treatment	Treatment costs for continuous treatment		Opportunity costs of treatment	QALYs per person
		SAE occurs (=A*£4.9M)	SAE does not occur (=(1-A)*£7.3M)	(=(C+D)/15000)	(=A*-14.24-E)
	[A]	[C]	[D]	[E]	[F]
Sample 1	5.0%	£243,887	£6,950,792	479.6	-480.4
Sample 2	5.1%	£248,765	£6,943,475	479.5	-480.2
Sample 3	5.5%	£268,276	£6,914,209	478.8	-479.6
Sample 4	5.6%	£273,154	£6,906,892	478.7	-479.5
Sample 5	5.6%	£273,154	£6,906,892	478.7	-479.5
Average	5.4%	£261,447	£6,924,452	479.1	-479.8

The same process can be carried out for withdrawal and the expected QALYs per person calculated for each sample. The results of this are shown in Table 4.8.

Table 4.8: Net QALY effects of withdrawal treatment, taking account of the opportunity cost of health expenditure

	Probability of SAE for withdrawal	Treatment costs for withdrawal treatment		Opportunity costs of treatment	QALYs per person
		SAE occurs (=B*£4.9M)	SAE does not occur (=(1-B)*£3.9M)	(=(G+H)/15000)	(=A*-14.24 - I)
	[B]	[G]	[H]	[I]	[J]
Sample 1	1.0%	£48,295	£3,926,170	265.0	-265.1
Sample 2	2.7%	£132,604	£3,857,630	266.0	-266.4
Sample 3	5.0%	£245,346	£3,765,975	267.4	-268.1
Sample 4	18.8%	£918,374	£3,218,828	275.8	-278.5
Sample 5	20.7%	£1,007,821	£3,146,110	276.9	-279.9
Average	9.6%	£470,488	£3,582,943	270.2	-271.6

As we now have samples for the NHB of continuous treatment (Table 4.7 column F) and withdrawal (Table 4.8 column J) the expected health consequences of uncertainty can be calculated in the same manner as described in Table 3.3. Comparing column F and J of Table 4.7 and Table 4.8 respectively shows that for

each sample treatment withdrawal is superior to continuous treatment. Therefore there is no uncertainty about the superior treatment meaning that the value of research is zero. This is driven by the large differences in costs between the treatment options. The treatment savings from withdrawal primarily arise from the assumption that for 50% of patients the high costs of Eculizumab over a lifetime are avoided. These savings are expected to result in benefits across the health system and so the analysis suggests that accepting the uncertainty and switching treatment to withdrawal results in the largest population health benefits[39].

However, just as for any other, this analysis does not proscribe social choice. There may be ethical arguments to distribute research resources in a manner which is not health maximising. The aim of this analysis is to help inform these discussions by estimating the health consequences of decisions and showing the health foregone which is inevitable in decision making. Decision makers may still choose to fund this study over other research proposals with larger expected health impacts. By making the ethical trade-offs clear this analysis can add to the transparency and accountability of this process.

[39] Furthermore, funding the trial may delay the recommended change in practice and this is also associated with health opportunity costs. From the proposal the health costs are £10.6M which translates to (£10.6M / £15,000 =) 707 QALYs in health opportunity costs. As the value of research is zero, the overall impact of the project is a loss of 707 QALYs.

4.2.2 RESEARCH PRIORITISATION FOR CONTINUOUS PRIMARY OUTCOMES

Rapid methods to address continuous outcomes are described and demonstrated here in using P4 as a case study.

4.2.2.1 *Overview of P4*

This proposal is for a four arm complex adaptive RCT to examine whether Exenatide, Telmisartan or their combination can slow neurodegeneration in Alzheimer's disease (AD). The only drugs approved for the management of AD (cholinesterase inhibitors and memantine) are for symptomatic relief and show limited clinical effects. At least 200 new drugs have advanced to phase 2 trials in the past 30 years but none have demonstrated disease modification (Schneider et al., 2014). Exenatide is a glucagon like peptide-1 (GLP-1) agonist that crosses the blood brain barrier and facilitates insulin signalling. Its use is based on the hypothesis that AD is linked to glucose and insulin signalling, and animal models provide empirical evidence to support the hypothesis. Telmisartan is an angiotensin II receptor blocker (ARB) which has been associated with delayed disease progression and preserved cognitive function in cognitive disorders. The NHS has effective mechanisms for diagnosing AD, but current drug therapies offer only modest improvement in symptoms and do not affect decline. Though P4 is a complex adaptive trial, a partial analysis is carried out in which the trial is assumed to be a regular 4 arm RCT (i.e., without accounting for the adaptive nature of the complex trial), which is expected to complete and report in 6 years.

Summary of proposal 4

Research question: Can Exenatide, Telmisartan or their combination slow neurodegeneration in Alzheimer's disease (AD)?

Interventions: Exenatide prolonged release 2mg once weekly, Telmisartan 40mg once daily and combination of Exenatide 2mg once weekly and Telmisartan 40mg once daily

Control: Placebo

Primary outcome: Slowed 2 year decline in Mini Mental State Examination (MMSE) of 3.1 points (vs 4.5 points expected for placebo)

Proposed study: Four arm complex adaptive RCT, n = 920 randomised 1:1:1:1

Duration of proposed study: 6 years for first stage with option to extend trial to add treatments

Costs of proposed study to NETSCC: £3,310,883

NHS support and treatment costs: £1,297,789

Box 4.2: Summary of proposal 4

4.2.2.2 *Informing uncertainty for mean difference*

In order to estimate the value of additional research for continuous outcomes an explicit summary of current uncertainty about the expected mean difference in MMSE score for the new interventions (Exenatide, Telmisartan and their combination) is required. However, meta-epidemiological evidence currently exists only for binary and survival outcomes[40]. To inform uncertainty in the primary outcome for P4 we apply an approach similar to Bennette et al., (2016). This approach calibrates current uncertainty in mean difference using a judgement about the probability of treatment success (alternative hypothesis) and the probability of treatment failure (null hypothesis). For P4, treatment success is defined in the proposal as an increase on the MMSE scale of 1.4 points. Treatment failure is defined as a relative decrease on the MMSE scale. Combining judgements about the probability of treatment success and the probability of treatment failure implies a distribution of expected mean difference in MMSE score for the new interventions.

First we demonstrate how a judgement about the probability of treatment success can be formed based on current evidence; we then show how this can be used to inform a judgement about the uncertainty in mean difference.

[40] Meta-epidemiological analysis to help inform continuous outcomes may be feasible. This could be carried out by converting the results of historical continuous outcome trials to standardised mean differences (SMDs). The global mean and heterogeneity for the set of SMDs could then be calculated and the predictive distribution could then be used in trials that are exchangeable by converting the SMD to the natural outcome (this requires the standard deviation of the natural outcome).

Probability of treatment success (the alternative hypothesis)

The proposal states that a change of 1.4 MMSE points constitutes disease modification and that this has not been achieved in the 200 previous phase 2 trials carried out over the previous decade (Schneider et al., 2014). Using this information it is possible to statistically inform the probability of success (i.e., a change of 1.4 MMSE points) in a new trial. This judgement can be informed using Laplace's rule of succession (Laplace, 1829). This is a method used in quality control to calculate the probability of an event occurring given that the event has not occurred after a large number of trials. The probability of the event is given by:

$$\text{Probability of event} = (\text{Observed events} + 1) / (\text{Number of trials} + 2)$$

In P4, zero events (treatment successes) have been observed from 200 previous trials. This implies that the probability of success in the two hundred and first trial is approximately (0 + 1) / (200 +2) = 0.5%. This approach assumes that that each previous trial investigated a different drug and the trials were sufficiently powered to detect treatment success[41].

Characterising uncertainty in mean difference

The probability that the new treatment is less effective than the current treatment is assumed to be 50%, i.e., there is equipoise and we are equally uncertain about the effectiveness of the alternative treatments. This represents the probability that the null hypothesis is true. From the calculation above, there is a 0.5% chance that a new treatment will result in an increase of more than 1.4 MMSE points over no treatment. Using these values we assume the outcome is approximately normally distributed and find the distribution which fits these values. A normal distribution which has a 50% chance of negative values and a 0.5% chance of values greater than 1.4 has a mean of zero and a standard deviation of 0.54 (see Box 4.3 for fitting method). This

[41] There are a number of assumptions made in the approach taken here. First, the method assumes that each of the drugs tested are independent. If groups of drugs are from a similar class (e.g. similar structure and/or biological pathway) then the independence assumption will not hold. Second, the approach assumes that the previous studies had 100% statistical power, which is unlikely in any research study. The impact of both of these assumptions is to underestimate the probability of success in future trials. In principal, it is possible to extend this method to take account of these limitations but this is beyond the scope of this thesis.

distribution reflects current uncertainty about the expected mean difference in MMSE score for a new treatment.

Fitting a normal distribution to reflect uncertainty in mean difference

Here we show how to find the mean (μ) and standard deviation (σ) of a normal distribution which is consistent with the null (treatment failure) and alternative hypothesis (treatment success).

From equipoise, we assume that the probability that the new treatment is worse than the current treatment is 50%. This implies that the difference in MMSE score for the new treatment has a 50% chance of being below zero. This is the null hypothesis. As the normal distribution is symmetrical this means that the mean of the required normal distribution (μ) must be zero.

From the fact that there have been zero treatment successes in 200 previous trials we calculated that there is a 0.5% chance that a new treatment will result in an increase of more than 1.4 MMSE points over no treatment. This implies that the difference in MMSE score for the new treatment (△MMSE) has a 0.5% chance of being greater than 1.4. This is the alternative hypothesis and is represented by the formula:

$$P(\Delta MMSE > 1.4) = 0.005$$

From the rules of standard normal distributions, this implies:

$$\frac{1.4 - \mu}{\sigma} = \Phi^{-1}(1 - 0.005) = 2.576$$

As we know that the mean is zero, we can substitute zero for μ to calculate σ:

$$\frac{1.4}{\sigma} = 2.576$$

$$\sigma = \frac{2.576}{1.4} = 0.54$$

This same logic can be used more generally to fit normal distributions if plausible estimates of the probability of null and alternative hypotheses can be made.

Box 4.3: Fitting a normal distribution to reflect uncertainty in mean difference

4.2.2.3 *Health benefits of further research in terms of primary outcome*

For continuous outcomes, uncertainty about the benefits of a new treatment is represented by the standard deviation or CI around the mean difference in effect. This represents the plausible change in the outcome compared to the baseline treatment, where zero represents no difference. In a similar manner as for binary outcomes, the consequences of uncertainty for continuous outcomes are derived by taking random samples from the distribution of mean difference. A simplified version of the P4 analysis is used to illustrate how we can get from mean difference with uncertainty to an estimate of the value of changing practice and the health consequences of uncertainty[42]. Assuming a normal distribution, consider a random sample of five possible values for the mean difference in MMSE for Exenatide (treatment) compared with no treatment (-1.08, -0.52, 0, 0.54, 1.08) which is the control. By definition, the baseline change in MMSE with no treatment is zero. The five random samples from the range of relative treatment effect and baseline change are shown in Table 4.9.

Table 4.9: Mean difference in primary outcome (MMSE) with Exenatide and no treatment

	Relative treatment effect Mean difference in MMSE (Exenatide vs no treatment) [A]	Change in MMSE with no treatment [B]
Sample 1	-1.41	0
Sample 2	-0.54	0
Sample 3	0	0
Sample 4	0.54	0
Sample 5	1.41	0
Average	0	0

[42] Only two treatment options are considered here: no treatment and Exenatide. In P4 there is also Telmisartan and a combination of Exenatide and Telmisartan.

The health impact of additional research

From Table 4.9, the average mean difference in MMSE for Exenatide is zero and so there is no additional health to be gained by changing practice from no treatment to Exenatide. The expected health consequences of the uncertainty depend on the chance that Exenatide is more effective, in addition to the size of this effect and the number of patients eligible for treatment. Given the information in Table 4.9, we can estimate the chance that Exenatide is more effective than no treatment. This is the probability of observing a mean difference of greater than 0, which is 2 out of 5 samples, i.e. 40%. To understand the scale of health that would be lost if no treatment were used instead of Exenatide in these samples, we calculate the number of additional MMSE points that would be observed with no treatment and compare it to the number that would be expected with Exenatide given an incident population of 100,000. The steps in this calculation are shown in Table 4.10.

Table 4.10: Rapid estimation of the health benefits of additional evidence for a continuous primary outcome, Mini-Mental State Exam (MMSE)

Sampled realisation of uncertainty	Exenatide mean difference	No treatment mean difference	Additional MMSE points per 100,000 eligible patients		Health benefits of additional evidence	
			Exenatide (=A*100,000)	No treatment (=B*100,000)	Absolute effect of Exenatide (=C-D)	Consequences of uncertainty for intervention (=F if A>0)
	[A]	[B]	[C]	[D]	[F]	[G]
Sample 1	-1.41	0	-141,000	0	-141,000	0
Sample 2	-0.54	0	-54,000	0	-54,000	0
Sample 3	0	0	0	0	0	0
Sample 4	0.54	0	54,000	0	54,000	54,000
Sample 5	1.41	0	141,000	0	141,000	141,000
Average	0	0	0	0	0	39,000

Combining the probability of error with the loss of health we can estimate the expected health consequences of uncertainty, 0.2 x 54,000 + 0.2 x 141,000 = 39,000 MMSE points per year. The above is a highly stylized example involving only two treatments and 5 samples. The analysis of P4 (using RANE) considers all 4 treatment options and is based on 50,000 samples (see Appendix A2 for full list of inputs).

This results in an estimate of the health consequences of uncertainty of 47,609 MMSE points per year. This provides an estimate of the value of research but does not yet account for the differences in costs across the treatments; this is addressed in the next section.

Other aspects of outcome to consider

Although change in MMSE score may be the most appropriate primary outcome, it is not necessarily the only relevant outcome for assessing the value of the proposed treatments. As for binary outcomes, specifying a minimum clinical difference (MCD) required to change clinical practice is one way to incorporate concerns about increased costs and/or potential adverse events. Since each of the active interventions in P4 (Exenatide and Telmisartan) are associated with increased treatment costs, there may be a need to demonstrate an improvement in average MMSE before practice should change. The research proposal specifies a required average change of 1.4 MMSE points per person for each of the active treatments. A single MCD for every treatment is hard to justify as the treatments differ substantially in price. A MCD of 1.4 is used here for illustrative purposes but this assumption is explored in the next Section (4.2.2.4). This MCD implies that if any of the treatments demonstrate a 1.4 point increase in MMSE at 2 years relative to placebo then that treatment would be considered superior to placebo and should be implemented. Requiring a 1.4 point increase implies that the additional costs are equivalent to 140,000 additional MMSE points per year[43]. Table 4.11 illustrates how the MCD is used to calculate the net MMSE points provided by Exenatide under different realisations of uncertainty.

[43] The trade-off implied by requiring an absolute change in a continuous outcome is calculated by multiplying the required absolute change (e.g. 1.4 points) by the incident population (e.g. 100,000) = 100,000 x 1.4 = 140,000 points.

Table 4.11: Rapid estimation of the health benefits of additional evidence for a primary outcome of Mini-Mental State Exam (MMSE) taking account of MCD = 140,000 MMSE points per annum

Gross MMSE points per annum for an incidence of 100,000 eligible patients			Net health benefits of additional evidence	
Exenatide (=A*100,000)	No treatment (=B*100,000)	Net MMSE points per annum for Exenatide (=C-140,000)	Absolute net effect for Exenatide in number of deaths per annum (=H-D)	Consequences of uncertainty (=I if I>0)
[C]	[D]	[H]	[I]	[J]
-141,000	0	-281,000	-281,000	0
-54,000	0	-194,000	-194,000	0
0	0	-140,000	-140,000	0
54,000	0	-86,000	-86,000	0
141,000	0	1,000	1,000	1000
0	0	-140,000	-140,000	200

Only in sample 5 does Exenatide provide a sufficiently large increase in MMSE to provide net health gains. Combining the probability of error with the loss of health we can estimate the expected health consequences of uncertainty, 0.2 x 1,000 = 200 MMSE points per year. As before, the above is a highly stylized example using 5 samples to compare only Exenatide and no treatment. The full analysis of P4 reported in Table 2.3 is based on 50,000 samples and considers four treatment options: no treatment, Telmisartan, Exenatide and the combination of Telmisartan and Exenatide[44]. An MCD of 1.4 for all active treatments results in an estimate of the health consequences of uncertainty of 266 MMSE points lost per year. When this is extended over the full time horizon (20 years) and the trial duration is taken account of the upper bound for the value of research is estimated to be 2,390 net MMSE points.

[44] This analysis is carried out using RANE with full inputs in the Appendix A2.

4.2.2.4 *Informing an appropriate MCD: what change in the primary endpoint is required before practice should change?*

A MCD of 1.4 was used above for all active treatments however as Exenatide is more expensive than Telmisartan, Exenatide might be required to demonstrate a larger improvement in MMSE than Telmisartan before it can be considered superior.

As discussed in Section 3.3.4 the MCD for each intervention should represent the change in the primary outcome that is required to make up for relevant secondary outcomes associated with each intervention. The appropriate MCD is determined by the same logic for continuous outcomes as for binary outcomes. Equation 3.2 applies to continuous outcomes but in this case $INHB$ is understood as the net health effect of a unit change in the continuous primary outcome measure. Treatment specific MCDs can be estimated for P4 by calculating the improvement in MMSE required to compensate for the additional costs associated with each of the active treatments.

As will be shown in Section 4.2.2.6 a one unit increase in MMSE is approximately associated with an additional 0.01242 QALYs and a (£14 x 12 =) £168 reduction in costs. The opportunity cost of health expenditure (£15,000/QALY) can be used to express the benefit of a unit change in MMSE score in net health terms. This is calculated by adding the QALY gain per additional MMSE point and adding the expected health benefits of any cost savings. In the present case this is 0.01242 + £168/£15,000 = 0.0236 QALYs per additional MMSE point.

Assuming that additional costs are the only relevant secondary outcome, the appropriate MCD is that which offsets the treatment costs for each new intervention. Assuming a 1 year treatment duration, these treatment costs are; (£73.36 x 12 =) £880.32 for Exenatide, (£14.83 x 12 =) £177.96 for Telmisartan and (£88.19 x 12 =) £1,058.28 for the combination treatment.

From Equation 3.1 the appropriate MCDs for each of the treatments are therefore (£880.32/(0.0236 x £15,000) =) 2.5 for Exenatide, (£177.96/(0.0236 x £15,000) =) 0.5 for Telmisartan and (£1,058.28/(0.0236 x £15,000) =) 3 for the combination treatment. Repeating the full analysis using these MCD values means the upper bound for the value of research is estimated to be 45,764 MMSE points. This is much larger than the previous estimate of 2,390 MMSE points because the cheapest

treatment, Telmisartan, is no longer being penalised by the same amount as the other, more expensive, treatments. Because it is no longer being penalised the analysis suggests there is a 17% chance that Telmisartan provides positive NHB and so there is more value in carrying out research to investigate this. This result illustrates how sensitive results can be with different values of MCD and the importance of considering the MCD for each treatment carefully.

4.2.2.5 *Linking the primary outcome to a comprehensive measure of health outcome*

Here we demonstrate a method to link continuous outcomes to costs and QALY, again using P4 as a worked example. In step 1, changes in the primary outcome are translated to QALYs; in step 2, the primary outcome is linked to disease related costs; in step 3, these are combined along with treatment costs to link uncertainty in relative effects to costs and QALYs.

4.2.2.6 *Step 1: Link changes in primary outcome to quality of life*

One method to understand the health impact of changes in the primary outcome is to link these changes to QALYs. QALYs are calculated by multiplying a measure of health related quality of life for a health state by the time spent in that health state. One questionnaire used to measure health related quality of life and calculate QALYs is called the EQ-5D (Briggs et al., 2006; Dolan, 1997). The resulting EQ-5D score is often referred to as health state utility. The relationship between changes in the primary outcome and changes in utility will depend on the severity of the disease and the units and range of the primary outcome measure. “Mapping” studies which use statistical methods to estimate the effect of a unit change in a clinical outcome on utility provide this link. These are commonly used in economic modelling and a database of mapping studies is available from Dakin et al., (2018). For P4, evidence on the relationship between MMSE and health utility is based on a study by Jönsson et al., (2006). This study was used in a NICE appraisal of drugs for Alzheimer’s disease (Bond et al., 2012). Jönsson et al., (2006) measured both MMSE and EQ-5D

for Alzheimer's patients across a range of disease severity. The results are reproduced below in Table 4.12.

Table 4.12: Relationship between Mini-Mental State Exam (MMSE) and health related utility in Alzheimer's population.

MMSE	Utility
0–9	0.33
10–14	0.49
15–20	0.5
21–25	0.64
26–30	0.69

Taking the midpoint of the MMSE score and estimating the relationship between MMSE and utility using simple linear regression implies that a one unit increase in MMSE results in a 0.01242 increase in utility[45]. Using this result, the additional health utility associated with changes from baseline can be converted to incremental QALYs by multiplying it by a judgement about the treatment effect duration. Estimates of treatment effect duration exist for few outcomes in the literature so in practice this will require expert opinion to inform this. Sensitivity analysis should also be carried out to investigate the impact of different values on results. For the purposes of the analysis it was assumed that the treatment effect will last for 1 year and that patients stay on treatment for the duration of treatment effect. This implies that a one unit increase in MMSE will result in an additional (0.01242 x 1 =) 0.01242 QALYs.

4.2.2.7 *Step 2: Link changes in primary outcome to disease related costs*

Changes in the primary outcome may also be expected to result in changes in disease related costs. Therefore, in the same manner as for health utility, changes in the continuous outcome must be approximately linked to changes in disease related costs. From the NICE appraisal of drugs for Alzheimer's disease (Bond et al., 2012), a study by Wolstenholme et al., (2002) was identified which estimates that a one point increase in MMSE is associated with a £56 decrease in 4 monthly costs.

[45] A more flexible model relating changes in MMSE to changes in utility (e.g. a spline) would fit the data better. However, incorporating any non-linearities into the model relating MMSE and utility would substantially increase the complexity of the model.

Assuming linearity, this implies a one unit increase in MMSE is associated with a (£56/4 =) £14 decrease in monthly costs.

4.2.2.8 *Step 3: Link uncertainty in relative effects to costs and QALYs*

In addition to disease related costs, treatment costs must also be taken into account to reflect the differences in treatment options. Here it is assumed that individuals are treated for the duration of treatment benefit and so treatment costs per person are calculated by multiplying monthly cost by treatment effect duration (in months). For P4, there are substantial costs associated with each of the active treatments; Exenatide costs £73.36 and Telmisartan costs £14.83 per month respectively (Medicines Complete, 2018). The combination treatment therefore costs (£73.36 + £14.83 =) £88.19 per month. As the incident population is 100,000 one month of treatment for Exenatide, Telmisartan and their combination is expected to cost £7.3 million, £1.4 million and £8.8 million respectively.

To understand the impact of each treatment on costs and QALYs, each sample for mean difference is translated to QALYs, linked to cost savings/increases in disease related costs and the treatment costs are recorded. The opportunity costs of health expenditure can then be used to translate the increased costs into health foregone to calculate NHBs. The steps in this calculation are shown in Table 4.13.

Table 4.13: Net QALY effects of Exenatide, taking account of the opportunity cost of health expenditure

	Exenatide mean difference [A]	Treatment effect duration in months [K]	Exenatide treatment costs per annum (=K*£73*100,000) [L]	Disease related costs per annum (=A*-£14*100,000) [M]	Total health opportunity costs per annum (=(L+M)/15,000) [N]	QALYs per annum (=A*0.01-N) [O]
Sample 1	-1.41	12	£88,032,000	£1,974,000	6,000	-6000
Sample 2	-0.54	12	£88,032,000	£756,000	5,919	-5919
Sample 3	0	12	£88,032,000	£0	5,869	-5869
Sample 4	0.54	12	£88,032,000	-£756,000	5,818	-5818
Sample 5	1.41	12	£88,032,000	-£1,974,000	5,737	-5737
Average	0	12	£88,032,000	£0	5,869	-5,869

By definition the baseline change in MMSE with no treatment is zero and the additional treatment costs are also zero. Therefore the baseline option will be associated with zero additional QALYs per year for every sample. We now have samples for the NHB of no treatment (zero) and Exenatide (column O) so the expected health consequences of uncertainty can be calculated as shown in Section 3.2 in Chapter 3. As column O does not contain any samples in which the NHB of Exenatide is positive there is no uncertainty about the superior treatment meaning there is no value in further research. This is due to the high treatment costs (£88M per year) of Exenatide, with gains in MMSE insufficient to make up for the health opportunity costs of these additional treatment costs.

The full analysis of P4 reported in Table 3.4 is based on 50,000 samples and considers all four treatment options: no treatment, Telmisartan, Exenatide and the combination of Telmisartan and Exenatide. The full analysis with RANE (full inputs in Appendix A2) results in an estimate of the health consequences of uncertainty of 117 QALYs per year and 967 QALYs over full time horizon. The value of research is positive in this case as Telmisartan monotherapy is considerably cheaper than Exenatide and there is a non-zero probability that Telmisartan is superior to no treatment in some simulations. The analysis suggest that it is very unlikely that Exenatide will provide sufficient benefit to compensate for its additional costs. This means that the expected value of the research could be increased by replacing the Exenatide arms with other promising therapies which are cheaper than Exenatide.

These conclusions will be sensitive to the assumptions made in the analysis. An important assumption is the duration of treatment effect. The above analysis is based on assuming a one year treatment effect duration. Extending this to two years increases the value of research from 967 to 1,985 QALYs. The value of research increases as each additional MMSE point is associated with a larger impact on population health.

4.2.3 RESEARCH PRIORITISATION FOR SURVIVAL PRIMARY OUTCOMES

Rapid methods to address survival primary outcomes are described and demonstrated here using P2 as a case study.

4.2.3.1 *Overview of P2*

P2 describes a three arm RCT investigating treatment with immunomodulating anti-programmed death receptor 1 (PD1) antibodies for patients with advanced (unresectable stage III, IV) melanoma who are due to start anti PD1 as first line treatment. The proposed trial will compare continuous treatment until disease progression (TTP) to treatment for 12 months only (TF12) and treatment for 6 months only (TF6). The anti PD1 antibodies are nivolumab and pembrolizumab which have been recently approved for use by NICE. The treatments are licensed for continuous use until disease progression (TTP) and so this is the UK standard of care. Prior to this guidance, the standard of care was dabrafenib, ipilimumab and vemurafenib. The proposal states that two years of treatment with nivolumab costs £140,000 for one patient which imposes significant costs on the NHS (we assume the same costs for pembrolizumab). There is no biological evidence that justifies the intensive treatment schedule and there are reports that patients who have discontinued treatment due to toxicity have maintained good disease response. The trial investigates whether acceptable rates of 2 year progression free survival (PFS) can be maintained with more conservative treatment schedules. The primary analysis of 2 year PFS is planned to report after 6 years. There is also an additional long term follow up study which is planned to report after 10.3 years.

Summary of proposal 2

Research question: Can individuals with advanced melanoma receiving 1st line anti-programmed death receptor 1 (PD1) antibody therapy (nivolumab or pembrolizumab) achieve and maintain as good an outcome if they receive 6 or 12 months total treatment duration compared with standard treatment duration of until disease progression?

Interventions: 6 months of anti-PD1 therapy; 12 months of anti-PD1 therapy (or until disease progression if sooner)

Control: Anti-PD1 therapy until disease progression

Primary outcome: 2-year PFS

Proposed study: Three-arm non-inferiority RCT in 1,068 participants (6-month intervention arm, n = 361; 12-month intervention arm, n = 361; control arm, n = 361)

Duration of proposed study: Primary analysis at 6 years (recruitment period of 5 years, followed by 12 months follow-up); long-term follow-up for 4 years taking the duration of the study to 10.3 years

Costs of proposed study to NETSCC: £2,522,710

NHS support and treatment costs: Saving of £62,410,967

Box 4.4: Summary of proposal 2

4.2.3.2 *Health benefits of further research in terms of primary outcome*

Survival outcomes describe the length of time spent in a particular health state. The health state in question may be beneficial such as remission in which case longer durations spent in the state are desirable. Or the state may be harmful, such as time in a state of relapse. For each health state there is a probability of leaving the state in each time period (e.g. each month). Consider the stylized example in Figure 4.1, each month there is assumed to be a 99% chance that patients will remain in the origin state (e.g. remission) and a 1% chance they will move to some post origin state (e.g. relapse or death).

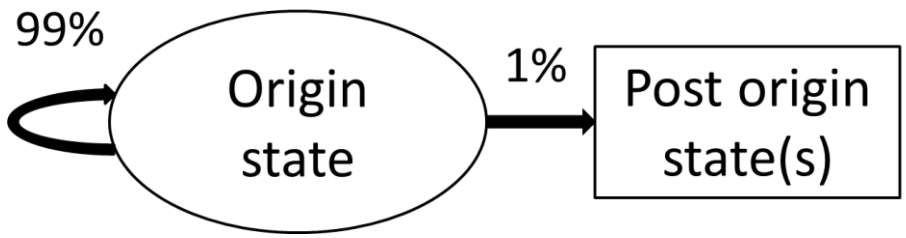


Figure 4.1: Example illustrating the transition between different health states in which hypothetical patients transition from the origin state to the health state at a rate of 1% per month

If the post origin states are not desirable (e.g. disease progression, death) then it is beneficial to remain in the origin state (e.g. remission) for longer periods. The chance of moving out of the origin state is governed by a hazard. This determines the probability of leaving the state each month. A smaller hazard is associated with a lower probability of leaving the state and vice versa. The relative effect of new interventions on the baseline hazard is represented using hazard ratios. A hazard ratio below 1 implies that individuals are less likely to move out of the origin state with the new intervention. Uncertainty about the effect of a new intervention is represented by the CI surrounding the hazard ratio. Similar to binary and continuous outcomes, the consequences of uncertainty for survival outcomes are derived by taking random samples from the distribution of the hazard ratio. This is combined with the baseline hazard and this is used to calculate the expected number of months spent in the origin state for each sample.

A simplified version of the P2 analysis is used to illustrate how we can use a hazard ratio with uncertainty to rapidly estimate the value of changing practice and the health consequences of uncertainty. In the case of P2 the origin state is pre-progression melanoma and the post origin states are disease progression and death. As the post origin states are not desirable, it is beneficial to remain pre-progression for longer periods. Time spent in the origin state is the PFS time and this is the primary outcome in P2. Therefore, in this section the benefits of research will be quantified in terms of months of PFS. P2 aims to compare TTP, TF12 and TF6; however, for simplicity we illustrate the methods comparing only TTP and TF6.

Baseline hazard

The first step is to define the baseline hazard for current practice. From the sample size calculation it is expected that 51% of individuals receiving the baseline treatment will not have exited the pre-progression state by two years. Assuming constant hazards this implies a baseline hazard of 0.03 for TTP (see Box 4.5)[46].

Estimating the baseline hazard i.e. baseline rate

The sample size calculation for P2 states that PFS at two years (24 months) is expected to be 51% for the baseline treatment (TTP). If a constant hazard is assumed; then this implies a (1 – 0.51 =) 0.49 probability of having left the progression free state by 24 months. To calculate the hazard over one month we use the formula from (Briggs et al., 2006):

$$r = -\log(1 - P)/t$$

Where P represents the probability of leaving the state (0.49) and t represents the time period (24 months). This results in a baseline monthly hazard of:

$$r = -\log(1 - 0.49)/24 \approx 0.03$$

Box 4.5: Estimating the baseline hazard for P2

[46] The constant baseline hazard assumption can be relaxed by modelling the baseline hazard using a using a more flexible model such as a Weibull. This requires two inputs; a lambda and a gamma parameter. The lambda parameter determines the initial hazard and the gamma parameter describes how this hazard changes over time. These inputs may come from a published survival analysis or expert elicitation exercise. It is also assumed that there is no uncertainty in the baseline hazard.

Relative effect; hazard ratio

Section 2.4.4 described a meta-epidemiological approach to characterizing uncertainty in relative effects in the absence of existing data. This was based on an empirical analysis by Djulbegovic et al., (2012) who analyzed the results of 743 publically funded RCTs pooling log odds ratios and log hazard ratios. The results of individual RCTs are unpredictable and tend to fall within an approximate 95% interval from 0.19 to 4.39. In the absence of other information, this approach represents a reasonable starting point to inform the magnitude of effect and uncertainty for the relative effectiveness of TF6 relative to TTP. However, the meta-epidemiological analysis currently does not take account of the fact that TF6 is effectively a reduced dose of TTP. Therefore there are good reasons to expect that TF6 will be less effective than current treatment and so will be associated with a higher hazard of transitioning out of pre-progression state to either progression or death; this implies a hazard ratio > 1. As it is very unlikely that treatment for 6 months will be superior to standard care we assume a range for the hazard ratio from 1.1 to 4.39. Further meta-epidemiological research and/or expert elicitation is required to better characterize uncertainty in this case.

Value of additional research

Let us now consider a random sample of five possible values for the hazard ratio for progression with the new intervention (0.9, 1.9, 2.2, 3.5 and 4.3). These five random samples are shown in Table 4.14.

Table 4.14: Samples of hazard ratio and baseline hazard for treatment and control

Sampled realisation of uncertainty	Relative treatment effect TF6 hazard ratio for leaving pre-progression state [A]	Baseline hazard Hazard of leaving pre-progression state with TTP [B]
Sample 1	0.9	0.03
Sample 2	1.9	0.03
Sample 3	2.2	0.03
Sample 4	3.5	0.03
Sample 5	4.3	0.03
Average	2.56	0.03

As the average hazard ratio for TF6 is greater than 1 this implies that on average individuals will leave the pre-progression state faster with the new treatment. Therefore the current evidence does not support changing practice as patients benefit when they remain in the pre-progression state longer. The expected health consequences of the uncertainty depend on the probability that TF6 is more effective, in addition to the size of this effect and the number of patients eligible for treatment. Given the information in Table 4.14 we can estimate the chance that current treatment (TTP) is less effective than TF6. This is the probability of observing a hazard ratio less than 1, which is 1 out of 5 samples, i.e. 10%. To understand the scale of health that would be lost if practice remains as TTP, we calculate the number of months of PFS that would be observed with TTP and compare it to the number that would be expected with TF6. Assuming constant hazards the expected number of months spent in the pre-progression state is 1/(hazard of leaving the state). For TTP is simply 1/(column B). For TF6 the hazard of leaving the state is calculated by multiplying the baseline hazard (column B) by the hazard ratio (column A). The NICE budget impact statement for the appraisal of nivolumab and pembrolizumab estimates that approximately 1,137 individuals per year will meet the criteria required for use of these drugs and so the number of months spent in the pre-progression state must be multiplied by 1,137 to calculate the population impact. The steps in this calculation are shown in Table 4.15.

Table 4.15: Rapid estimation of the health benefits of additional evidence for a survival primary outcome, months in pre-progression state. TTP, treatment to progression. TF6, treat for six months.

Sampled realisation of uncertainty	TF6 hazard ratio for leaving pre-progression state	Hazard of leaving pre-progression state with TTP	Months in pre progression state per annum		Absolute effect of TF6 (=D-C)	Consequences of uncertainty for TTP (=F if A<0)
			TTP (=1/B*1,137)	TF6 (=1/(B*A)*1,137)		
	[A]	[B]	[C]	[D]	[F]	[G]
Sample 1	0.9	0.03	37,900	42,111	4,211	4,211
Sample 2	1.9	0.03	37,900	19,947	-17,953	0
Sample 3	2.2	0.03	37,900	17,227	-20,673	0
Sample 4	3.5	0.03	37,900	10,829	-27,071	0
Sample 5	4.3	0.03	37,900	8,814	-29,086	0
Average	2.56	0.03	37,900	19,786	-18,114	842

Combining the probability of error with the loss of health we can estimate the expected health consequences of uncertainty in the same manner as described for other outcomes, 0.2 x 4,211 = 842 months in the pre-progression state per year.

Other aspects of outcome to consider

In the sample size calculation in the research proposal a hazard ratio of 1.25 is considered acceptable for the new interventions (TF6 and TF12). This implies that if any of the treatments demonstrate a hazard ratio below 1.25 then that treatment would be considered superior to current practice. To state this MCD in absolute health terms, assuming a constant baseline hazard of 0.03 this hazard ratio implies that a reduction in 6.7 months of PFS per individual is acceptable for the new interventions[47]. At a population level this implies that the additional costs are equivalent to 7,618 months of PFS each year[48].

As discussed in Box 4.5 this baseline hazard estimate is based on two year expected PFS of 51% in the control arm (TTP). The MCD in absolute health terms will depend on this value of baseline hazard and so sensitivity analysis is useful to understand how this can change with different values. A two year PFS of 40% and 60% implies baseline hazards of 0.04 and 0.02 respectively. In this case, the hazard ratio of 1.25 implies that a reduction of 5.2 months and 9.4 months PFS per individual is acceptable for the new treatment. Alternative formulations for the baseline hazard will also affect these results.

As in P4 a single MCD for both treatments (TF6 and TF12) is hard to justify as the treatments differ substantially in costs, however this MCD is used here for illustrative purposes. The steps involved in taking account of the MCD to calculate *net* PFS from *gross* PFS are the same used in Section 2.3.4.1; the *gross* benefits of each treatment are calculated in terms of the primary outcome then the absolute

[47] The relationship between months of an outcome (e.g. months of PFS) and a hazard ratio (e.g. 1.25) given a baseline hazard (e.g. 0.03) is given by: 1/(hazard x hazard ratio) – 1/hazard = 1/(0.03 x 1.25) – 1/(0.03) = -6.7 months.

[48] The trade-off implied by requiring an absolute change in a survival outcome is calculated by multiplying the acceptable absolute change (e.g. -6.7 months) by the incident population (e.g. 1,137) = 1,137 x -6.7 = 7,818 months.

MCD is added to the new treatment to calculate the *net* benefits of the new treatment. This is illustrated in Table 4.16.

Table 4.16: Rapid estimation of the health benefits of additional evidence for a primary outcome of progression free survival (PFS) taking account of MCD = 6.7 months per person. TTP, treatment to progression. TF6, treat for six months.

Gross PFS per annum (in months)		MCD: acceptable decrease of 6.7 months per person for TF6 (=1,137*6.7)	Net PFS per annum for TF6 (=C+H)	Net health benefits of additional evidence	
TTP (=1/B*1,137)	TF6 (=1/(B*A)*1,137)			Net absolute effect of TF6 (=I-C)	Consequences of uncertainty (=J if J>0)
[C]	[D]	[H]	[I]	[J]	[K]
37,900	42,111	7,618	49,729	11,829	11,829
37,900	19,947	7,618	27,565	-10,335	0
37,900	17,227	7,618	24,845	-13,055	0
37,900	10,829	7,618	18,447	-19,453	0
37,900	8,814	7,618	16,432	-21,468	0
37,900	19,786	7,618	27,404	-10,496	2,366

As the expected net PFS for the new treatment is negative (bottom row, column 5), current evidence suggest that TTP should be chosen above TF6. Only in sample 1 does TF6 provide a sufficiently large increase in months of PFS to provide net health gains. Combining the probability of error with the loss of health we can estimate the expected health consequences of uncertainty, 0.2 x 11,829 = 2,366 months of PFS per year.

The full analysis of P2 reported in Table 2.3 is based on 50,000 samples and considers three treatment options: TTP, TF6 and TF12[49]. Given an absolute MCD of 6.7 for all active treatments the upper bound for the value of research is estimated to be 544 net months of PFS. This is much lower than the one year estimate from Table 4.16; this is due to the small sample size used.

4.2.3.3 *Informing an appropriate MCD: what change in the primary endpoint is required before practice should change?*

The MCD in this case should represent the change in the primary outcome (PFS) that is acceptable given the total per person cost savings associated with TF12 and TF6. The appropriate MCD is determined by the same logic for survival outcomes as for binary and continuous outcomes. Equation 3.2 applies to survival outcomes but in this case $INHB$ is understood as the net health effect of a unit change in the survival primary outcome.

As will be shown in Section 4.2.3.4 an additional month of PFS is associated with approximately (0.79/12 =) 0.0658 QALYs and £100 in additional costs. The net health effect of an additional month of PFS ($INHB$) is calculated by adding the QALY gain and subtracting the health opportunity costs of any additional costs. In the present case this is 0.0658 - £100/£15,000 = 0.059 QALYs per additional month of PFS.

Assuming that additional costs are the only relevant secondary outcome, the appropriate MCD is that which offsets the treatment costs for each new intervention. Given that per person costs associated with current treatment are estimated to be £215,760 the cost savings of TF12 and TF6 are expected to be (£215,760 - £72,504 =) £143,256 and (£215,760 - £36,252 =) £179,508 for TF12 and TF6 respectively.

[49] This analysis is carried out using RANE with full inputs in Appendix A3.

The MCD should represent the reduction in the primary outcome which is acceptable given the cost savings.

From Equation 3.1 the appropriate per person MCDs for each of the treatments are therefore (£143,256/(0.059 x £15,000) =) 162 months of PFS for TF12 and (£179,508/(0.059 x £15,000) =) 203 months of PFS for TF6. This implies that given the cost savings, reductions in PFS of 162 months and 203 months for PF12 and PF6 are acceptable. As discussed previously, the research proposal specifies a MCD of 6.7 month reduction in PFS for both of the interventions. The analysis here suggests that this MCD is too small to reflect the large cost savings associated with the interventions.

As such large reductions in PFS are impossible at the individual level this suggests that the new treatments will dominate current practice due to the large cost savings. Therefore there is a zero probability that current practice will be shown to be the optimal treatment in a trial and so there is no value in including TTP in research. It also suggests that current practice should immediately switch to one of either TF6 or TF12 whether or not the trial is carried out[50].

There may be reasons to dispute these treatment specific MCD values. For example if the correct estimate of k was £30,000 rather than £15,000 the appropriate per person MCDs for each of the treatments would be (£143,256/(0.059 x £30,000) =) 81 months of PFS for TF12 and (£179,508/(0.059 x £30,000) =) 101 months of PFS for TF6. Treatment costs may also differ from those included in the formula (£143,256 for TF12 and £179,508 for TF6). Finally, the net health benefits of additional months of survival may be greater or less than 0.059 QALYs. Sensitivity analysis should be conducted to investigate the impact of changing these assumptions.

4.2.3.4 *Linking the primary outcome to a comprehensive measure of health outcome*

Decision making can be supported by translating months of PFS into generic health outcomes along with the relevant costs and cost savings. Here we demonstrate this method to link survival outcomes to costs and QALYs using P2 as a worked

[50] There may be reasons that it is not desirable and/or feasible to change practice in absence of new research. This was discussed in Section 3.4.

example. In step 1, changes in the primary outcome are translated to QALYs; in step 2, the primary outcome is linked to disease related costs; in step 3, treatment costs are incorporated[51].

4.2.3.5 *Step 1: Link changes in primary outcome to quality of life*

To understand the magnitude of health consequences associated with spending additional months in the origin state, a health state utility is required. For P2 the utility for the pre progression health state is estimated to be 0.79 from the CHECKMATE-006 trial (Robert et al., 2015). To illustrate, this implies that if a treatment is associated with a 30 months PFS this will result in ((0.79 x 30)/12 =) 1.98 QALYs.

4.2.3.6 *Step 2: Link changes in primary outcome to disease related costs*

In the same manner as for health utility, to understand the magnitude of disease related costs associated with spending additional months in the origin state, monthly disease related costs are required. From the NICE appraisal of pembrolizumab *(National Institute for Health and Care Excellence*, 2015), a study by Johnston et al., (2012) was identified which estimates that monthly costs in the pre-progression state (not including treatment costs) are £100. For example, this implies that if a treatment is associated with a 30 months PFS this will result in (£100 x 30) =) £3,000 in disease related costs.

4.2.3.7 *Step 3: Link uncertainty in relative effects to costs and QALYs*

In addition to disease related costs, treatment costs must also be taken into account to reflect the differences in treatment options. Treatment may continue (with the attendant costs) until the individual leaves the origin state or it may halt after a set period of time. In this approach, treatment is always assumed to stop after leaving the origin state.

Monthly costs for each treatment in P2 are expected to be £6,042. The expected PFS for the baseline treatment is (1/rate = 1/0.03 =) 35.7 months. At £6,042 per month this implies expected cost per patient of (35.7 x £6,042 =) £215,760. The maximum

[51] Continuous discounting can be implemented to take account of the fact that treatment costs and benefits are expected to occur at different times. This is not shown here to avoid obscuring the demonstration.

duration of treatment for the new interventions (TF12 and TF6) are capped at 12 and 6 months, resulting in maximum per patient costs of (12 x £6,042 =) £72,504 and (6 x £6,042 =) £36,252 for TF12 and TF6 respectively. These are the maximum per patient treatment costs as treatment is assumed to stop if the patient leaves the pre-progression state.

Calculating NHBs

To understand the impact of uncertainty on costs and QALYs, each sample for PFS in Table 4.15 is translated to QALYs, linked to cost savings/increases in disease related costs and the treatment costs are recorded. The opportunity costs of health expenditure can then be used to translate the increased costs into health foregone to calculate NHBs.

For example, if TF6 is expected to be associated with 15 months PFS in a given sample this would translate to ((0.79 x 15)/12 =) 11.85 QALYs. This duration of PFS is also associated with (£100 x 15 =) £1,500 in disease related costs. Additional treatment costs are expected to be (6 x £6,042 =) £36,252 meaning net costs associated with 6 months treatment this simulation are (£36,252 + £1,500 =) £37,752. Combining costs and QALYs using the opportunity cost of health expenditure of £15,000/QALY, NHBs are estimated to be (11.85 - £37,752/£15,000 = 11.85 – 2.52 =) 9.33 QALYs. In the same manner as shown in Section 3.2 this is repeated and combined with the incident population to describe the distribution of health consequences of uncertainty for each of the treatments and thus the value of additional research.

Section 4.2.3.3 illustrated that due to the large treatment costs associated with TTP it is very unlikely to provide greater net health than TF6. This is because the health opportunity costs of using TTP relative to TF6 far outweigh any health gains from additional PFS. This means that the value of research comparing TF6 to TTP is zero. The full analysis includes three treatments; TTP, TF6 and TF12. This analysis also shows that the value of research is zero[52] as TF6 dominates both TTP and TF12 due to cost savings. As discussed in Section 4.2.1.6, this analysis does not proscribe

[52] Funding the trial may delay the recommended change in practice and this is also associated with health opportunity costs. This is discussed in detail in Section 3.4.

social choice; rather it provides a framework for thinking about research commissioning and explores the health consequences of different decisions.

4.3 How can the health impact of feasibility/pilot studies be considered?

The methods from Chapter 2 and 3 show how to rapidly calculate the benefits of research for comparative effectiveness research. However, P5 is for a feasibility study which can be understood as a preliminary step before full comparative effectiveness research is commissioned. The health benefit of feasibility/pilot studies primarily derives from the full trials they potentially lead to. Therefore to understand the value of a feasibility/pilot study, some judgement must be made about (i) the value of the full trial it will potentially facilitate and (ii) the probability that the full trial is actually feasible. An implication of this is that proposals for feasibility/pilot studies must include judgements about future research. Here we demonstrate a rapid method to estimate the value of a given feasibility/pilot study[53]. We describe the information required to carry out the analysis and how to compare the health impact of this type of study to other research competing for funding. P5 is used as a case study throughout.

Understanding the value of a feasibility/pilot study requires two steps. First, the value of the potential full trial must be estimated, using either the rapid methods described in this thesis or a full economic model. Second, this "potential VOI" must be adjusted for the fact that the full trial will not take place if the feasibility/pilot study shows that it is not possible. The final section will discuss the relevant considerations when prioritizing feasibility/pilot studies relative to comparative effectiveness research.

[53] The methods described here provide a method to understand the value of a given feasibility/pilot study proposal, they do not provide a framework for understanding whether a feasibility/pilot study or a full trial should be carried out as this would require additional inputs and complexity. This is expanded upon in the discussion.

4.3.1 Overview of P5

The proposal states that there is mounting evidence that antipsychotics (APs) are poorly tolerated by adolescents, causing significant safety concerns. There is also evidence that talking therapies (such as cognitive behavioural therapy or family intervention) can help reduce symptoms and prevent relapse. The NICE guideline for treatment of psychosis and schizophrenia in adolescents (*National Institute for Health and Care Excellence*, 2014) suggests that treatment options should include the possibility of choice between talking therapies, APs or both. However, there is uncertainty around the effectiveness and safety of these interventions, with no clear evidence base for adolescents. There are challenges and uncertainties associated with running a full RCT. To address this, the applicants propose a feasibility study to inform whether a future full, clinical and cost-effectiveness trial is possible.

Summary of proposal 5

Research question: Is it possible to carry out a full RCT to assess whether psychological interventions (PI) for first episode psychosis (FEP) in children and young people (CYP) is non-inferior to APs in managing psychiatric symptoms. Also, to assess whether combination therapy of PI and AP is superior to monotherapies in managing psychiatric symptoms.

Intervention: PI (up to 30 sessions of cognitive behavior therapy delivered to CYP at home over 6 months + 6 sessions of family intervention with parents / carers); Combined treatment of PI + APs.

Control: APs for 6 months. APs will be chosen based on NICE guidance.

Primary outcome: Feasibility of conducting a full RCT. For subsequent RCT.

Proposed study: 3-arm pilot feasibility RCT (n = 90) to inform whether a future full, clinical and cost-effectiveness trial is possible.

Duration of proposed (feasibility) study: 2 years

Costs of proposed (feasibility) study to NETSCC: £601,481

NHS support and treatment costs for feasibility study: £150,000

Box 4.6: Summary of proposal 5

4.3.2 ESTIMATING THE HEALTH BENEFITS OF FEASIBILITY RESEARCH

4.3.2.1 *Step 1: Estimate potential VOI for full trial*

Health consequences of the current uncertainty

P5 aims to compare three treatments for use in FEP, the current treatment (APs) and two new interventions PI and AP + PI. As the primary outcome (relapse) is binary the upper bound for the consequences of uncertainty can be estimated using the rapid methods illustrated in Chapters 2 and 3. The analysis was carried out using the RANE tool with full details on evidence used and judgements made in the Appendix A1.

There currently exists no evidence to inform the relative effect for the new interventions and so in the absence of direct evidence meta-epidemiological data is used as a starting point to understand uncertainty. As the primary outcome (relapse) is harmful it is expected that the odds ratio for PI and AP + PI is likely to be between (95% CI) 4.39 and 0.19 (see Section 2.4.4 for further details)[54].

From this judgement about the current evidence there is considerable uncertainty about the optimal treatment. Using the methods illustrated in Chapter 2 this uncertainty in relative effects is combined with an estimate of baseline probability of relapse (Craig et al., 2004) and incidence (Kirkbride et al., 2013) to understand the consequences of uncertainty in absolute health terms. From this the health consequences of this uncertainty are estimated to be 181 relapses per year.

Health impact of the full trial

Extending the yearly consequences of uncertainty over the 15 year time horizon, the maximum value of research is estimated to be 2,114 relapses avoided over the full time horizon. This is illustrated in Figure 4.2 and represents the maximum value of research.

[54] Applying the same prior distribution to the odds ratio for both PI and PI in combination with APs may not be justifiable on scientific grounds as there may be substantive evidence and/or clinical rationale to suggest that the combination of treatments will work noticeably better than any of the treatments alone (Mills et al., 2012; Thorlund et al., 2017). This is discussed in Section 5.6.

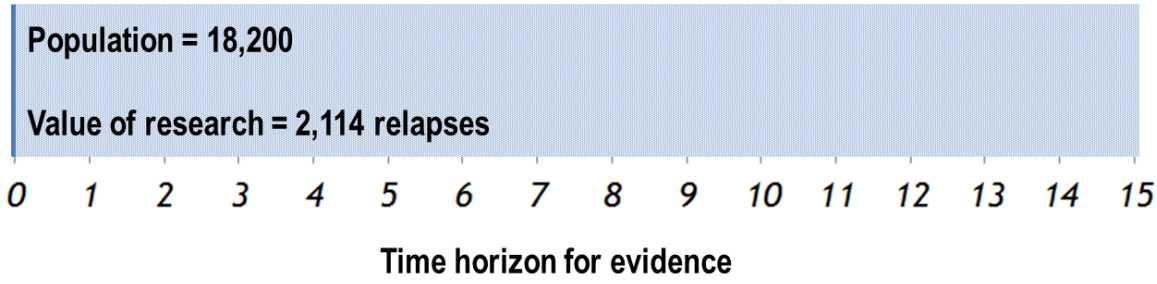


Figure 4.2: Consequences of uncertainty for P5 over the full 15 year time horizon

This must be adjusted for the fact that the longer research takes to report the lower will be its value. This is because as time moves forward medical knowledge and technology moves with it. Therefore the longer research takes to report the greater is the chance that the information gained will be irrelevant. This is important to consider when estimating the value of feasibility studies as it will take time for both the feasibility study to report and for the full trial to report (should the full trial be deemed feasible). From the proposal, the feasibility study is expected to take two years to report. An estimate of the time required for the potential full trial to be commissioned and report is also required but was not provided in the research proposal. Morgan and colleagues (2018) carried out a survey and analysed the trials funded by the NIHR Research for patient benefit (RfPB) programme to understand the costs and benefits associated with feasibility studies. They found that the average time from the reporting of a feasibility study to the reporting of the associated full trial report is approximately six years. This provides an empirically based starting point for analysis.

If the full research is deemed possible by the feasibility trial, the earliest it will report is the sum of the time taken for both the feasibility study and the full trial to report. In the case study this is (2 + 6 =) 8 years, and so the upper bound on the value of the full trial is expected to be 850 relapses avoided (see Figure 4.3).

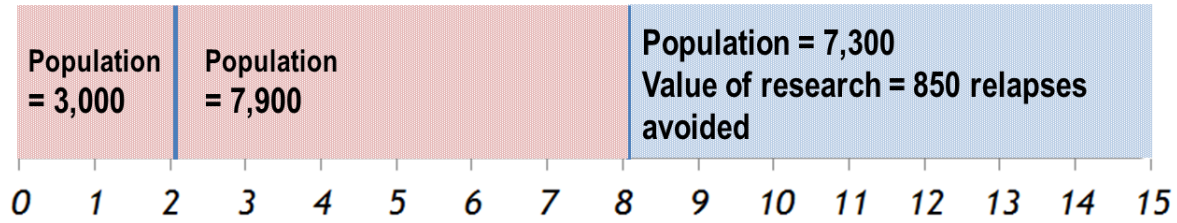


Figure 4.3: Upper bound for value of research given that the feasibility study is expected to report in two years and the full research is expected to take an additional 6 years

In some cases the value of research may fall to zero after the time to report has been taken account of. In these cases there is zero value in the feasibility study. In this scenario shorter full trials may be still provide research value and research funding applicants should take this into account.

4.3.2.2 *Step 2: Value of feasibility research*

The above value of research (850 relapses avoided) represents an upper bound on the value of the feasibility study assuming it leads to a full study. However the motivation for the feasibility study is that there is uncertainty about whether the full trial is possible. If the feasibility study shows that the full trial is not possible, the research budget spent on the feasibility study will have no impact on health outcomes. For this reason, the likelihood of a feasibility study leading to the full trial is an important determinant of its value. The value of the feasibility study is calculated by multiplying the value of the proposed full research by the probability of this research occurring. Morgan and colleagues (2018) estimate that there is a 64% chance that a feasibility study will lead to a full trial. Taking account of this, the upper bound on the value of this project falls to (850 x 64% =) 544 relapses avoided.

In the same manner as the expected health impact of the project must be adjusted for the fact that the full trial may not occur, the expected costs of the project must also be adjusted for. The feasibility study costs (of £601,481) will be incurred regardless of whether the full trial is commissioned. However, the full trial costs will only be incurred if the full trial is commissioned. Again, from Morgan and colleagues (2018)

full trials funded by the NIHR RfPB programme cost £1,163,996 on average and if the full trial is feasible, it is expected to take place 10 months after the feasibility trial reports i.e. (24 + 10)/12 = 2.8 years. Discounting to present value the expected cost of the full trial is approximately £1 million[55]. The total expected NETSCC research costs are therefore (£601,481 + £1,000,000 x 64% =) £1,241,481[56]. The expected cost per outcome of the research project is the expected cost to the research funder divided by the expected health benefit, (£1,241,481/544 =) £2,282 per relapse avoided in this case.

If the primary outcome is linked to costs and QALYs, the research costs imposed on the NHS can also be taken account of. As before, the NHS costs associated with the feasibility trial are always incurred but the NHS costs for the full trial are conditional. Prior to taking account of the costs imposed in the NHS, P5 is estimated to be worth a maximum of 115 QALYs in total (inputs listed in Appendix). Assuming the NHS costs associated with the full trial are £490,000 (after discounting) in the case study, the expected costs of research imposed on the NHS are estimated to be (£150,000 + £490,000 x 64% =) £463,600. The health opportunity costs associated with the above health system resources can then be subtracted from the maximum value of the research. The opportunity costs are estimated to be (£463,600/£15,000 =) -31 QALYs. After these costs have been subtracted the maximum value of this research falls from 115 to 84 QALYs. Therefore, the expected upper bound on the value of funding the feasibility trial is (£1,241,481/84 =) £14,780 per QALY gained[57].

4.3.2.3 *How to value pilot/feasibility studies relative to other research proposals?*

[55] £1,163,996/(1.035)^2.8 = £1,057,106

[56] This assumes that if it is possible, the costs of funding the full research will fall on NETSCC (or a similar UK body).

[57] Table 3.4 reports that the value of P1 is £14,806/QALY rather than £14,780 per QALY this slight discrepancy is due to rounding.

In understanding value (£14,780 per QALY) of a feasibility[58] study, the costs and health benefits of the feasibility study cannot be separated from the costs and health benefits of the potential full trial. This is because the health benefits of a feasibility study depend on the health benefits of the full trial, as shown above.

Though the *value* of the project is determined by the total expected research costs (£1,241,481) only the up-front feasibility study costs (£601,481) are borne at the time of feasibility study commissioning. This means that in applying the bookshelf approach to research funding, commissioning a feasibility study will consume £601,481 of the research budget and so this is the appropriate width of the "book" representing P5 in Figure 3.1. The height of the P5 book is determined by the expected value of the project (£14,780 per QALY).

If the feasibility study is "successful" and the full research is found to be feasible, then the value of this full research should be assessed on its own merits relative to the other projects available for funding.

[58] This reasoning also applies to pilot studies.

4.3.3 DISCUSSION

This chapter has extended the models outlined in Chapters 2 and 3 to provide a set of methods which can be applied to a wide set of research proposals. The extensions were motivated by the set of six historical proposals that were considered as part of the NETSCC programme (introduced in Chapter 2). Methods were developed in two directions: first, to address trials with primary outcomes other than simple binary outcomes; second, to extend the types of research design to include feasibility/pilot studies. Limitations and possible extensions of these methods are discussed in turn below.

4.3.3.1 *Research prioritisation with diverse primary outcomes*

This chapter extended the methods of Claxton et al., (2015a) and McKenna et al., (2016) to rapidly estimate the value of research for binary, continuous and survival primary outcomes. The "rapid approach" to VOI analysis is to provide decision makers with models which are practical, built around primary outcomes and are quick to implement. Therefore in developing rapid models for each of these primary outcomes the aim was; to keep the models general enough to apply to a number of disease areas, require a minimum number of inputs and be capable of reflecting the salient difference between treatments (see Chapter 7 for further discussion). There are a number of limitations and possible extensions to the methods described in this chapter. These are discussed individually for each type of primary outcome; binary, continuous and survival.

Binary, multinomial and composite primary outcomes

The most important limitations in linking primary outcomes to QALYs arise when the binary primary outcome is a result of an inappropriate dichotomisation i.e. when a multifaceted outcome has been inappropriately forced into two categories. Applying models which link binary primary outcomes to QALYs assumes that the dichotomisation of outcomes is appropriate. This means that it is possible and appropriate to analyse patient outcomes in terms of an event which either occurs or does not occur. For example in the P1 case study described in Chapters 2 and 3, the primary outcome was a binary concept called "functional recovery" which was deemed to have occurred if the patient had a GOSE score of 4 or above and was

deemed to not have occurred if the GOSE score was 3 or below. Splitting the GOSE scale into two parts in this way follows recent clinical literature in this field (Nichol et al., (2015) which reports the probability of functional recovery as its primary outcome. Linking this primary outcome to QALYs as shown in Chapter 3 is one means to understand the health impact of the primary outcome. However, as the GOSE scale has eight categories dichotomisation may not sufficiently reflect the different health consequences of these states. McKenna et al., (2016) use the GOSE score as a primary outcome but do not split it into a binary outcome. Rather McKenna and colleagues analyse the GOSE score as a multinomial (multi category) outcome. In this case prior studies are used to estimate the individual probability of entering each of the GOSE states considered and how this probability differs between treatment and control. This is more realistic than the binary approach but correspondingly increases the number of required inputs and hence the complexity of the analysis. Another important barrier to using a multinomial outcome (as opposed to dichotomisation) in research prioritisation is the difficulty in explicitly quantifying the uncertainty for the relative effect of new treatments in absence of previous studies. Meta-epidemiological methods are available as a starting point for binary outcomes (as discussed in Section 2.4.4) but further research is required to provide useful methods for multinomial outcomes.

Composite outcomes in the medical literature result from creating a binary outcome from a group of outcomes. In this case the composite outcome usually consists of a group of harmful outcomes (e.g. heart attack, death and stroke). A rationale for using composite endpoints is that they can increase the number of events in a trial and so will increase the overall statistical power, however the interpretability of results reported in terms of composite outcomes have been challenged (Montori et al., 2005). The P3 case study described in Section 4.2.1 is an example of a composite primary outcome in which the binary primary outcome is a composite of three serious adverse events; death, meningococcal infection and irreversible organ injury. In order to understand the health impact of the outcome (in terms of QALYs) an assumption was required about the relative frequency of the serious adverse events conditional on the composite event occurring. Informing this judgement is difficult and depending on the context this dichotomisation may or may not be appropriate to capture the effects of different treatments on health outcomes. This process

illustrates the implicit assumptions required to understand the health impact of composite endpoints and so can be used more generally to aid their interpretation.

Continuous outcomes model

The methods for analysis of continuous outcomes outlined in Section 4.2.2 are based around changes in mean difference. They assume that the average mean difference for a treatment (the relative effect) is constant over time and ceases after the treatment effect duration has elapsed. It also assumes that the treatment effect duration is the same for each of the treatments considered. In specific cases, subject matter knowledge may indicate that these assumptions are unrealistic. For example, the effects of a pharmacological intervention in chronic diseases may be expected to build up slowly over time, plateau for an extended period and then taper off gradually. This means that the models provided in this thesis can provide only an approximation to the disease process, though the practical implications of this are unlikely to be dramatic.

A potentially more important assumption is that the analysis is currently unconstrained by the minimum and maximum range of the primary outcome scale. This is likely to be important in situations in which the baseline outcome is close to the maximum or minimum value of the scale and large changes in mean difference are plausible. In this case theoretically impossible scale values may influence results. For the P4 case study the primary outcome (the MMSE scale) ranges from 0 to 30 (Folstein et al., 1975), the patients in the proposed trial have mild to moderate AD and so the baseline MMSE will range from 10 – 26. It is unlikely that changes in mean difference sufficient to result in theoretically impossible MMSE scores (below 0 and above 30) will occur. This is because the standard errors for the relative treatment effects are estimated to be 0.54 meaning that changes greater than +/- 1.08 MMSE points are unlikely. A model which takes account of the maximum and minimum range of continuous outcomes is possible but would require users to specify the expected baseline score on the primary outcome scale (including uncertainty) in addition to the scale's theoretical minimum and maximum.

Another potentially worthwhile extension to the rapid model used for continuous outcomes is to allow for a relationship between the continuous outcome and expected survival. This would be important in scenarios in which changes in the

continuous outcome are predictive of survival. To implement this extension, users would be required to specify the change in survival associated with a unit change on the primary outcome in addition to the expected baseline score on the scale (including uncertainty) (Ibrahim et al., 2010).

Survival outcomes model

The most important simplification made when linking survival outcomes to QALYs (outlined in Section 4.2.3) is that only the costs and health gains associated with time spent in the origin state are considered. This approach assumes that all relevant differences between treatment effects are captured by the time spent in the origin state. More technically it assumes that the outcomes which occur after leaving the origin state are fixed, in that they do not depend on (i) the particular treatment used or (ii) the time spent in the origin state. For example, in the P2 case study it is assumed that post-progression survival, quality of life and costs are not affected by the time spent in the pre-progression state. Post progression outcomes are also not affected by any side-effects of the treatment choices. To relax this assumption would require the user to explicitly specify the costs and health outcomes in the post origin state in addition to the relationship between time spent in the origin state and time spent in the post origin state. This is the approach taken by Bennette et al., (2016) and Carlson et al., (2018). These authors constructed a customised Markov decision model for each research proposal submitted to a US based oncology research prioritisation body. Compared to the simple survival model outlined in this chapter, the fully customised approach can create a more realistic model of the disease process but is more complex and time consuming. In particular, the customised approach presents challenges to the development of software which can rapidly construct a decision model and calculate VOI outputs (see Chapter 5). It is important to recognise, however, that the more intensive methods described by Bennette and Carlson are not competing alternatives to those described in this chapter, rather, these methods form a continuum of options from which analysts and decision makers should choose the most suited to their context.

4.3.3.2 *Feasibility/pilot studies*

The method described in this chapter can be used to compare the value of feasibility/pilot studies to more directly informative comparative effectiveness research such as RCTs when both types of research compete for scarce funding. In order to apply this method, information on the potential definitive trial is required (such as trial duration, cost etc.). This information was absent from the P5 application received by NETSCC. Only information on the feasibility study itself was included, with no information provided about the potential full research that the feasibility study may lead to. Though information on the definitive trial may be limited, resources exist to empirically inform the necessary judgements (Morgan et al., 2018). Though it provides a useful starting point for analysis, the method described in this section contains a number of simplifying assumptions. As elsewhere in this thesis, the methods proposed aim to balance both complexity and realism with demands on analyst resources.

It is possible to model feasibility studies with varying degrees of complexity. For example, a feasibility study may be conceptualised as a diagnostic test which is used to determine whether the full research is feasible or not. The method for assessing the value of feasibility studies described in the previous section assumes that the full research is either feasible, in which case it reports in full (true positive), or it is not feasible, in which case this is discovered with certainty by the feasibility study (true negative). It is assumed (for simplicity) that the feasibility study is perfectly predictive of whether the full research will successfully report or not, but in reality there is always the possibility that the feasibility study will incorrectly declare the full research impossible (false negative) or incorrectly declare the full research possible (false positive).

A more sophisticated approach still would be to explicitly link the recruitment rate, the feasibility of randomisation and effect sizes observed in a feasibility trial to the decision to carry out the full research. This could be extended to inform decisions about whether it is better to fund a feasibility study, conduct an internal pilot phase within the trial or to attempt to run the full research and accept the possibility that it

may fail[59]. The additional inputs, complexity and computational burden of these methods mean they are outside the scope of the present thesis, though they are potentially an important area of future research.

[59] This more complex approach would nest uncertainties about recruitment rate, randomisation etc. into the overall research design space (Conti and Claxton, 2009).

Chapter 5

Rapid Assessment of Need for Evidence (RANE) tool

5.1 INTRODUCTION

The methods described in Chapters 2 to 4 provide a means to rapidly estimate the value of research for a range of primary outcomes for both comparative effectiveness research and feasibility/pilot studies. In order to reduce the demands on analyst resources these methods were designed to capture the most important aspects of the decision without requiring a large number of inputs. In addition to searching for appropriate inputs, constructing decision models and performing VOI analysis demands a large amount of analyst resources and expertise (Myers et al., 2012).

The rapid approach to VOI analysis is to provide decision makers with models which are practical, built around primary outcomes and are quick to implement. The aim of this chapter is to introduce and test a tool which has been designed to reduce the technical barriers to implementation of VOI methods. This tool is called Rapid Assessment of Need for Evidence (RANE) and has been developed as part of this PhD. RANE is open source, hosted by the University of York and is freely available for use by anyone at https://shiny.york.ac.uk/rane/ (Full code available at https://github.com/david-glynn). This website embeds the methods described in Chapters 2 to 4 and so allows users to quickly carry out VOI calculations to help inform research prioritisation. This is important as reducing the technical barriers to VOI analysis can facilitate its use more widely in the health system thus improving the transparency and accountability of research decision making.

In this chapter we first informally review the software currently available for research prioritisation using VOI. Second, we provide an overview of the RANE tool and its capabilities. Third, we use a case study (P6) in a step by step illustration of how to use the RANE tool. Finally we assess the generalisability of the tool by investigating whether it could be applied to a new set of NETSCC research funding proposals.

5.2 CURRENT SOFTWARE FOR RESEARCH PRIORITISATION USING VOI

Chapters 2 to 4 demonstrated methods which can be used to aid research prioritisation. For this approach to be practical for resource constrained applicants and decision makers these methods must not present a large technical barrier. Online tools have begun to emerge to facilitate VOI analysis for the purposes of research prioritisation. A prominent example is Sheffield Accelerated Value of Information (SAVI) available at http://savi.shef.ac.uk/SAVI/ which takes the results of a full economic model and uses these to calculate the value of further research (Strong et al., 2014). The limitation of this approach is that a full economic model is required before this the value of research prioritisation can be estimated. As discussed, these models require time, technical skill and large amounts of information to construct. A web tool which aims to address this limitation is Value of Information for Cardiovascular Trials and Other Comparative Research (VICTOR) available at https://uwchoice.shinyapps.io/victor/ which provides a simple user interface to build a disease model for cardiovascular outcomes. This model draws on survival models based on US, English and Danish data (Basu et al., 2018a)[60]. VOI calculations are then carried out given the user inputs and the background disease models. This approach addresses both the time and technical barriers posed by full economic modelling; however it is limited as it is currently only applicable to cardiovascular outcomes and currently only quantifies the benefits of research in terms of life expectancy (i.e. VOI cannot be calculated in terms of a generic outcome such as QALYs).

To overcome the limitations of the existing tools we have developed the RANE tool. Unlike SAVI, this tool is based on the method described in Chapters 2 to 4 and so does not require a full economic model. Unlike VICTOR, it is based on a primary outcome so it is not limited to only cardiovascular research and can be linked to QALYs to facilitate decision making across disease areas.

[60] VICTOR implements a Markov model with six states: no events, non-cardiovascular death, cardiovascular death, myocardial infarction, stroke and bleeding. The latter three states are attached to life expectancy payoffs. See link for further details: https://sop.washington.edu/wp-content/uploads/VICTOR_LEcalculations.pdf

In the next section we illustrate the how to use RANE by applying it to the P6 case study (this is one of the six research applications received by NETSCC, as discussed in Section 2.4.1).

5.3 OVERVIEW OF RANE TOOL

RANE is written in Shiny (Chang et al., 2017) which provides user friendly interface for the R statistical programming language (R Core Team, 2017). RANE currently supports binary, continuous and survival primary outcomes. The tool can be used to estimate the potential value of full research (i.e. RCTs or well conducted observational studies) and feasibility/pilot studies. After entering the required information and clicking the "run analysis" button, the tool automatically calculates, presents and interprets a VOI analysis (based on 50,000 simulations) in approximately 2 seconds. After becoming familiar with the tool, the analysis of new proposals is expected to take approximately between one and four days per research proposal where the majority of this time is spent searching the literature for relevant inputs. This is an approximation, thorough piloting is required to establish time required in practice.

5.4 STEP BY STEP APPLICATION OF THE RANE TOOL TO A RESEARCH APPLICATION

5.4.1 OVERVIEW OF P6

This proposal is for a two arm RCT to examine the benefits of an education booklet containing practical information on the management of common problems during end of life care for family carers of advanced cancer patients with estimated prognosis of 8-24 weeks. Most cancer patients want to die at home (Higginson and Sen-Gupta, 2000) and a 201 Macmillan survey of 1,019 UK adults living with cancer cited in the proposal found 73% would prefer to die at home if concerns e.g. access to pain relief, support for family carers were addressed, however currently approximately 30% achieve their wish (Office for National Statistics, 2014). The educational booklet was designed to provide carers with knowledge and to facilitate patients' wishes to die at home. It was co-compiled with family carers, assessed by the Plain English Campaign.

Summary of proposal 6

Research question: Does an educational booklet facilitate family carers in supporting end of life care for patients with advanced cancer and allowing them to die at home (if this is their preferred place)?

Intervention: Practical booklet with standardised structured advice from a health professional on its use (with usual care)

Control: Usual care

Primary outcome: Death at home (if this is their preferred place), assessed and documented at recruitment and post-death

Proposed study: Two arm RCT in 679 patients (intervention arm, n = 453; control arm, n = 227)

Duration of proposed study: Primary outcome at 3 years

Costs of proposed study to NETSCC: £882,177

NHS support and treatment costs: £4,104

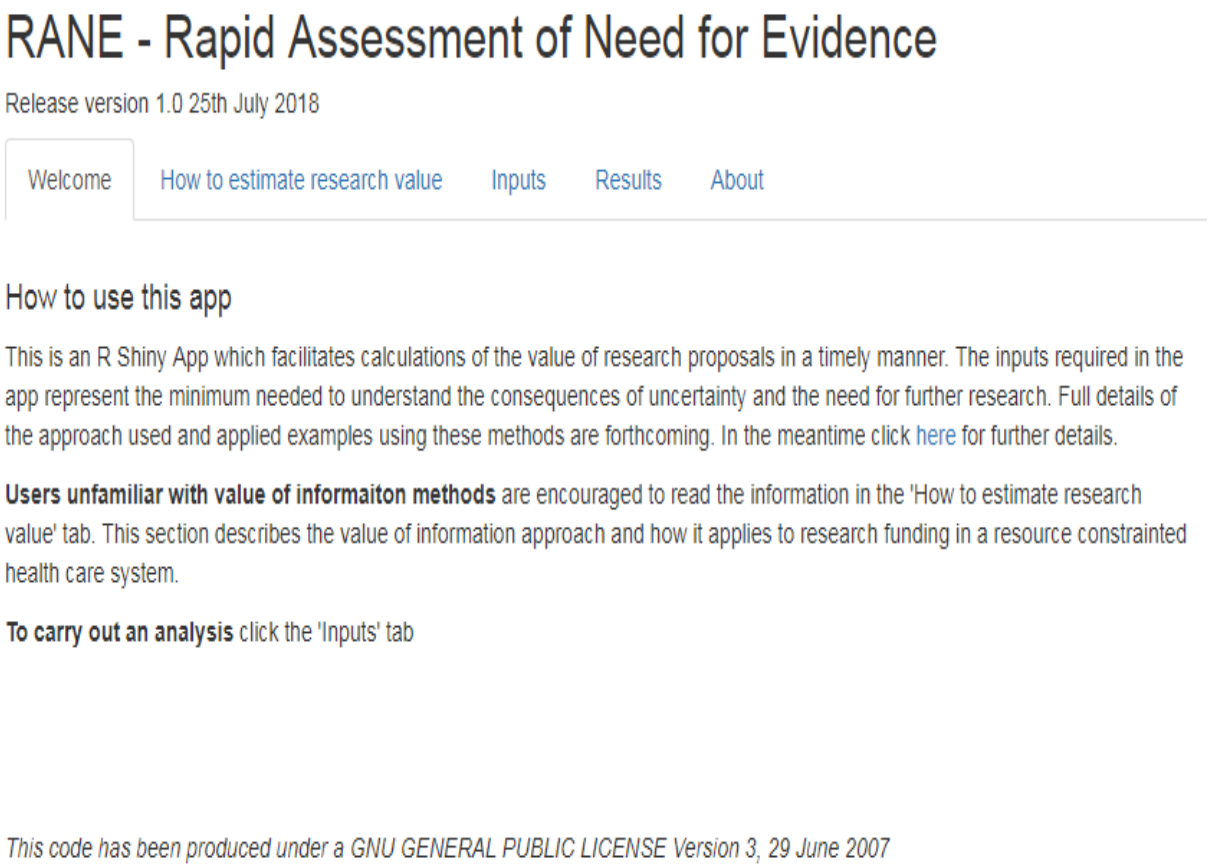
Box 5.1: Summary of proposal 6

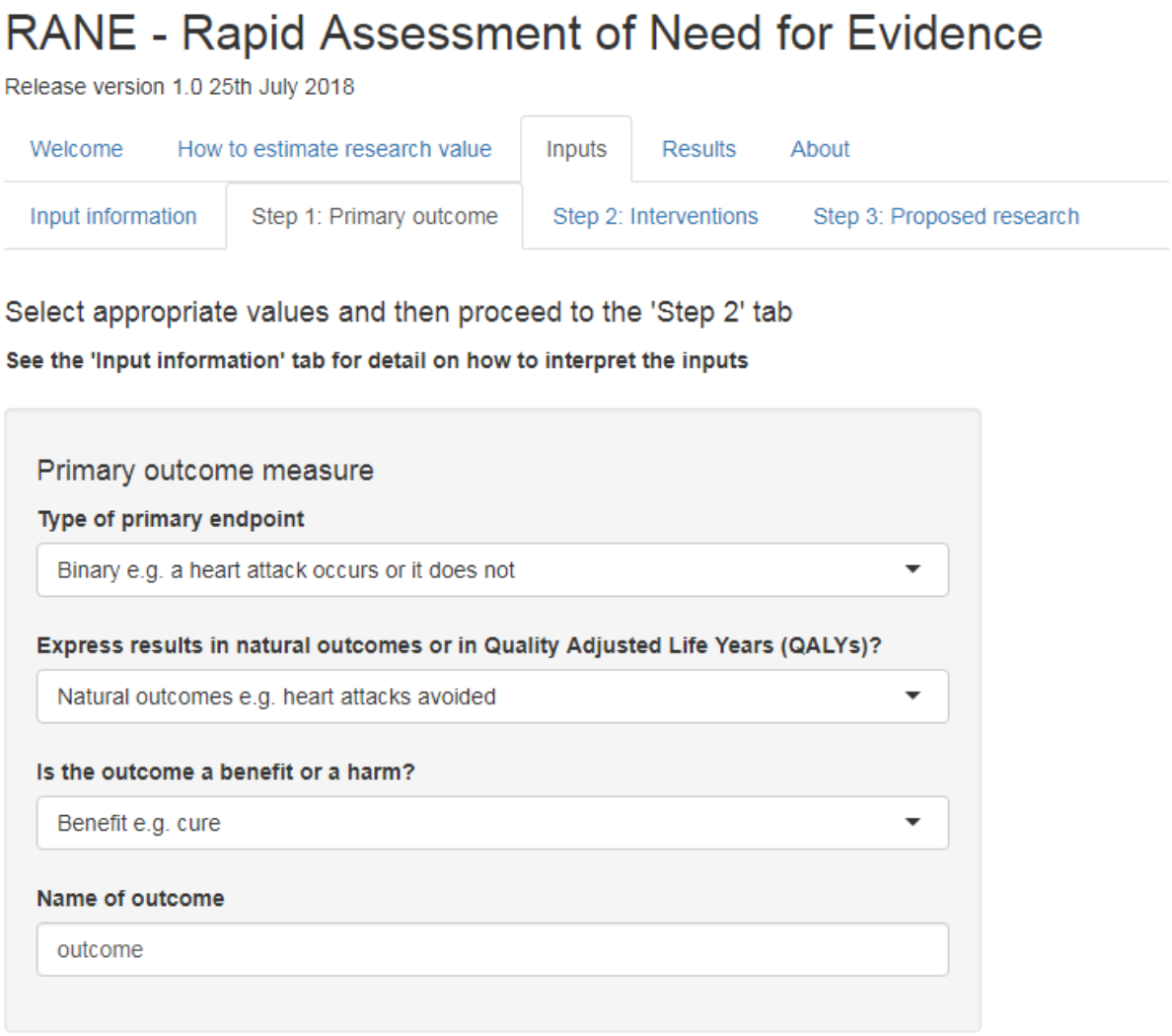
5.4.2 APPLYING THE RANE TOOL TO A RESEARCH FUNDING APPLICATION

There are three steps involved in applying RANE to a new application. In this section we illustrate the inputs and judgements required for each step to inform the value of research in terms of costs and quality adjusted life years (QALYs) using P6 as a case study. Depending on the type of analysis (e.g. type of primary outcome, type of research, QALY outcomes etc.) the required inputs will differ slightly.

Step 1: Primary outcome

After clicking the link to RANE (https://shiny.york.ac.uk/rane/) the user is taken to the "Welcome" tab. For information on VOI methods and the methods underlying the tool the user can choose the "How to estimate research value" tab. To begin the analysis the user should choose the "Inputs" tab and will be taken to "Step 1: Primary outcome" subtab.

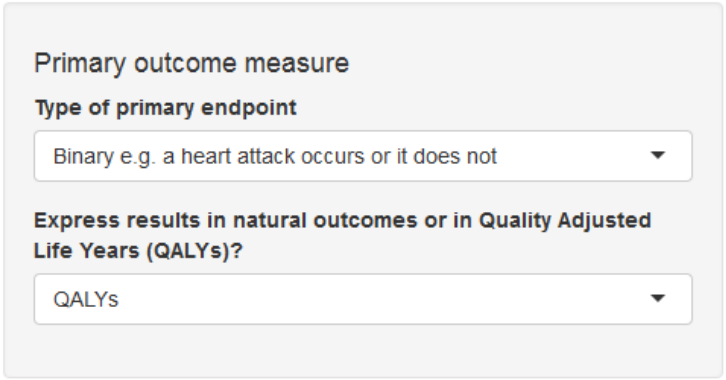




The first choices are to select the type of primary outcome (binary, continuous or survival) and whether the results are to be presented in terms of natural outcomes (i.e. in clinical units such as number of additional functional recoveries) or in QALYs. As calculating results in terms of QALYs involves additional steps we illustrate this analysis here.

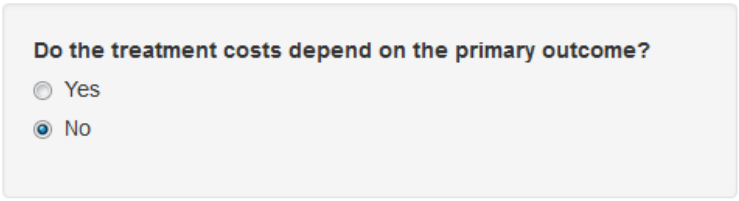
Choosing the type of primary outcome

As in any analysis judgements are required about the appropriate analysis given time, expertise and resource constraints (Brennan et al., 2006). For P6, the primary outcome is death at home (when this is the preferred place). This is a binary outcome i.e. there are only two possible outcomes; either the patient dies at home or they are assumed to die in hospital. The inputs into RANE required for a QALY analysis of a binary outcome are shown:

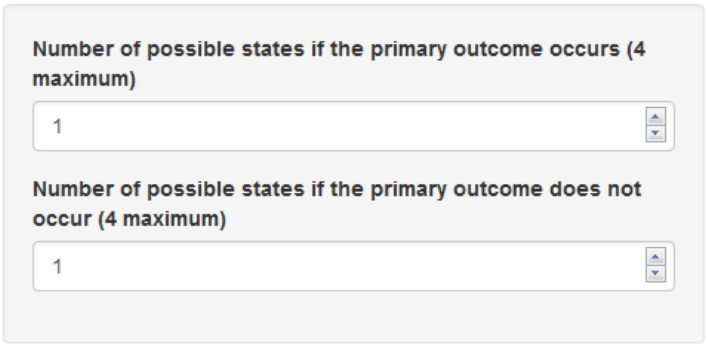


Linking primary outcome to costs and QALYs

If QALY analysis is chosen a set of judgements is required to link the primary outcome to this generic measure of health outcome. As the costs of the treatment (the booklet) are not expected to depend on the primary outcome (whether the patient dies at home) in P6, "No" should be left selected.



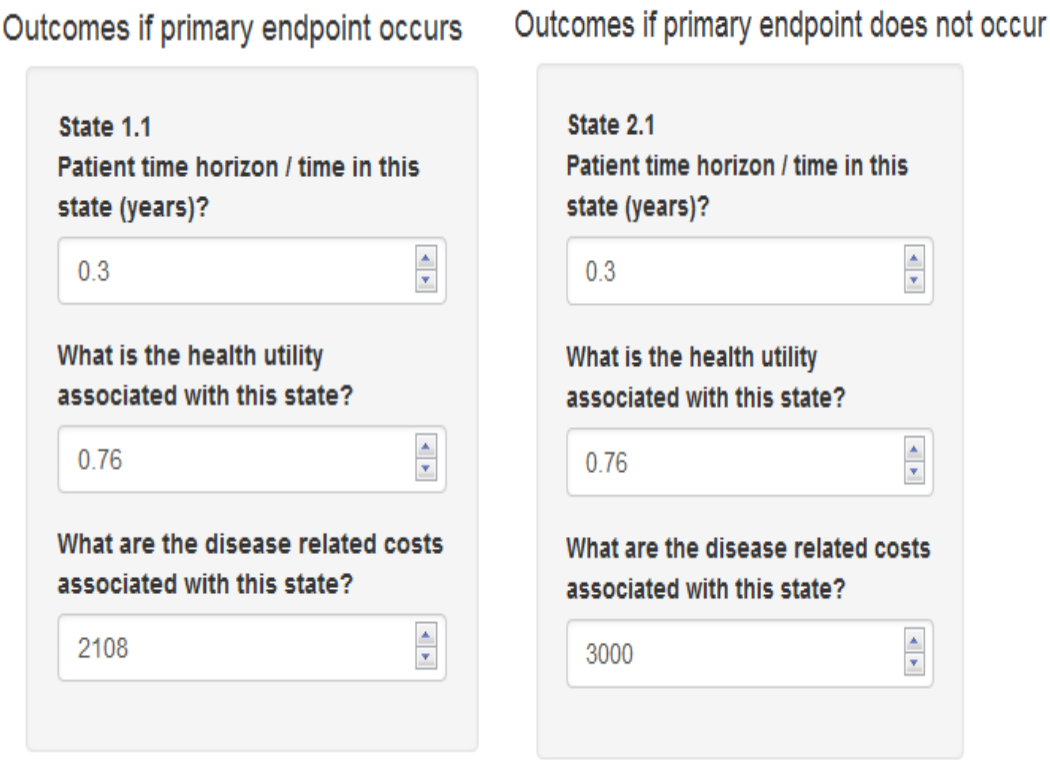
As death at home is not a composite endpoint and there is no scale of severity, we only consider the costs and health effects when the endpoint occurs and when it does not. This is contrasted with P1 in Section 3.2 in which those experiencing the primary outcome (functional recovery) may enter one of four possible health states. As there is only one possible state when the patient dies in their preferred place and one possible state when the patient dies elsewhere, the inputs into RANE are shown below[61].



To understand the costs and health consequences for patients who die at home compared to those who do not; judgments are required for the patient time horizon, health state utility and disease related costs associated with both states. From proposal, those included in the trial have a prognosis for survival of 8-24 weeks. The

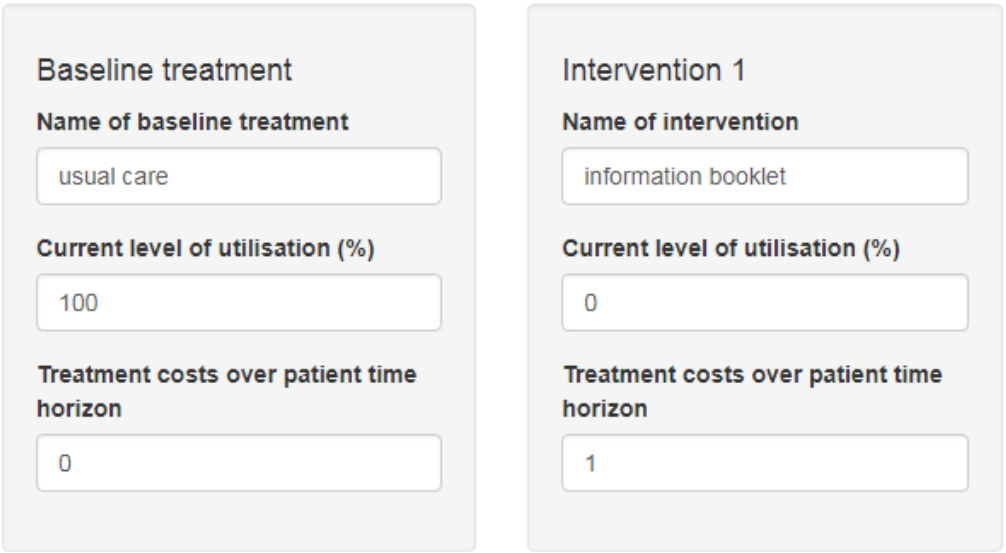
[61] In this analysis, patients are assumed to either die at home or in hospital. If death in a hospice is considered and there are differences in the expected costs and health outcomes associated with dying in a hospice compared to dying in hospital then this could be reflected by choosing 2 possible states associated with the primary endpoint not occurring; one for dying in a hospital and one for dying at a hospice.

midpoint in this range is 16 weeks which is equivalent to (16/52 =) 0.3 years. It is assumed that dying at home or at hospital does not affect survival or health related quality of life and so the patient time horizon for both states is 0.3 years. It is also assumed that the health state utility is not affected by whether the patient dies at home or at hospital. As the patients requiring care will be heterogeneous we use a review of health state utilities at the end of life to estimate an average utility of 0.76 (Dixon et al., 2009). To understand the costs associated with the primary outcome we assume that if patients do not die at home they will die at hospital. From a review carried out by the End of Life Care Programme (2012), the average cost of a hospital stay ending in death is estimated to be approximately £3,000. This same report provided a range of £1,415 to £2,800 for the cost of dying in the community. Taking the midpoint of this range, a death a home is expected to cost £2,108. These judgements are entered as shown below, completing step 1. The user now must click the subtab "Step 2: Interventions".



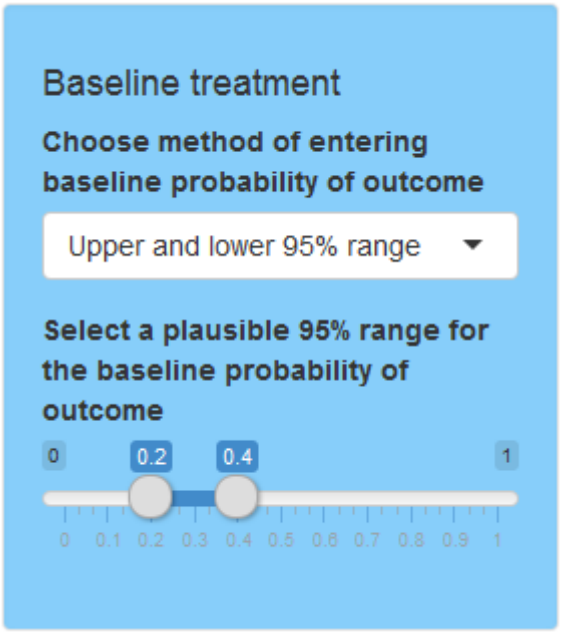
Step 2: Interventions

In this section the user must select the number of treatments and provide an explicit quantification of the current evidence for these treatments. In order to estimate the value of implementing the results of current evidence, the user can also input the current utilisation of the interventions, though this is optional. The first step is to select the number of treatments considered. For P6 there are two treatments; usual care which is the baseline treatment and the information booklet which is the new intervention. No information was provided on the cost of the booklet, it is assumed that additional copies cost £1 to produce. It is also assumed that the booklet is not currently in use so utilisation of the new treatment is 0%. These judgements are entered as shown.



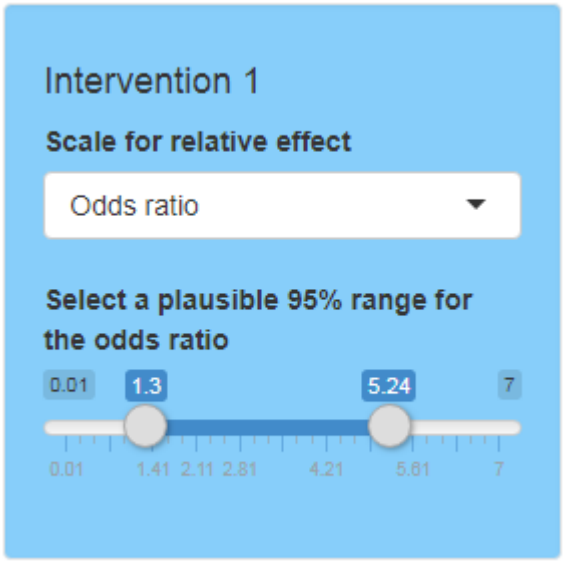
Baseline event rate

A judgement about the probability of dying at home with the baseline treatment is required to understand the value of research. From proposal, of those who wish to die at home approximately 30% achieve their wish (Office for National Statistics, 2014). No more information is provided about this study and so we assume +/- 10% uncertainty for this estimate resulting in a 95% CI from 20% to 40%. This is entered as a range in the tool as shown below.



Relative treatment effect

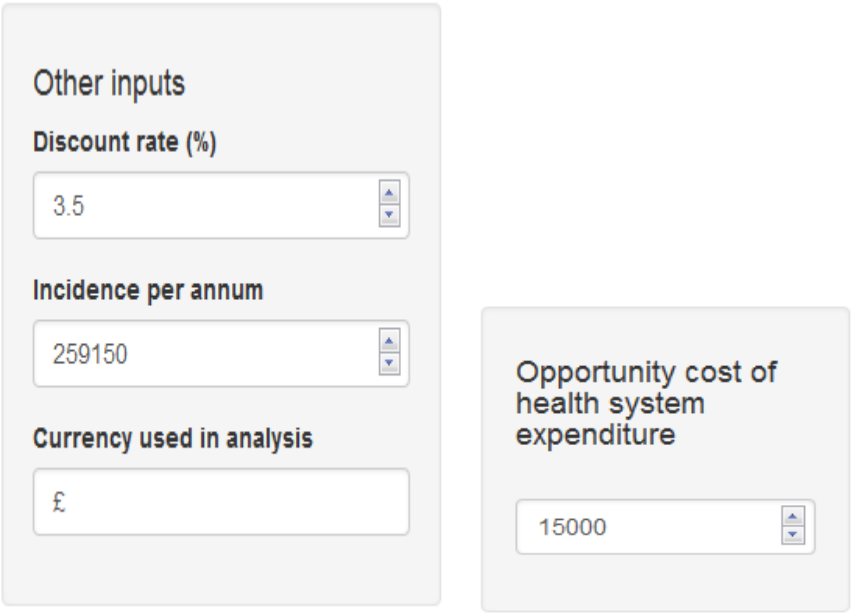
Using the meta-epidemiological approach described in Section 2.4.4 as a starting point to inform the uncertainty about dying at home reflects the fact that before carrying out the research we don't know what the odds ratio for the effect of the new treatment will be relative to usual care. Current evidence suggests that because death in home (when it is preferred) is a binary beneficial outcome the odds ratio for the effect is likely to fall between 0.23 and 5.24. However, it is reasonable to expect that providing the information booklet will be unlikely to reduce the chance of dying in home, which implies that odds ratios < 1 are very unlikely. Therefore we assume a 95% range for the odds ratio from 1.3 to 5.24. Further meta-epidemiological research and/or expert elicitation is required to formally characterize uncertainty in cases such as this. These judgements are entered as shown below, completing step 2. The user now must click the subtab "Step 3: Proposed research".



Step 3: Proposed research

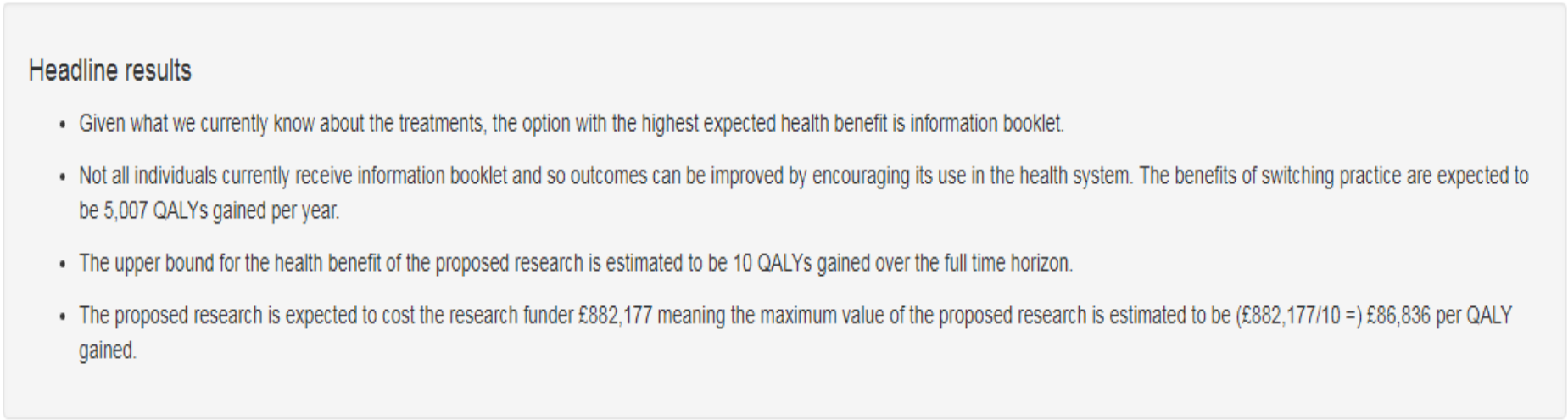
In this step the user selects the inputs describing the type, duration and costs of the research in addition to the time the information is expected to be valuable, the incidence and the opportunity cost of health expenditure. The research proposed is a 3 year RCT which aims to address the research question directly i.e. it is for a full research study, not a feasibility/pilot study. The cost to the research funder (NETSCC in this case) is £882,177 and the estimated excess health system costs are £4,014. As practice in this area appears to move relatively slowly, it is anticipated that the new information will be valuable for a time span of 15 years.

The discount rate is not reported in the proposal, however guidance from the UK Treasury suggests the use of a discount rate of 3.5% per annum (*HMT Green Book*, 2013). To calculate the incident population the Palliative Care Funding Review (T. Hughes-Hallett et al., 2011) estimates that 355,000 patients need palliative care each year and a 2010 Macmillan survey cited in the report estimates 73% of these patients would prefer to die at home. Therefore the population of interest is approximately (355,000 x 0.73 =) 259,150 patients per annum. As in previous sections the opportunity cost of health care expenditure for the NHS is £15,000/QALY (Claxton et al., 2015b; NHS England, 2015).



Results

After these values have been entered, the user must click “Run analysis” and proceed to the “Results” tab where the results will be presented (in approximately 2 seconds). The headline results summarise the value of further research with a more detailed breakdown of results reported below this. The headline results for P6 are shown below.



These results indicate that there is a lot of value in implementing information booklet, 5,007 QALYs per year and very little in carrying out research, 10 QALYs over 15 years. The cost per QALY from funding this research is £86,836/QALY which compares poorly to the other proposals in Section 2.5 i.e. it provides a relatively low number of QALYs for its funding costs. However, there may be other reasons to consider funding this proposal over the other proposals considered for funding. The first reason is that there may be aspects of benefit which have not been captured. In the analysis the benefits of dying at home are captured only by cost savings associated with not dying in hospital. There are likely to be patient and family relevant benefits from dying at home and the analysis here does not reflect this. Taking this into account would increase the impact of the primary outcome and so would increase the value of the research. A second reason is that it may be difficult to distribute the booklet in absence of a trial demonstrating its effectiveness. As discussed in detail in Section 3.4 using the research budget as a method to change practice will necessarily divert resources away from research projects which could address genuine uncertainties in the health system. Because the research booklet is not costly and is unlikely to result in additional patients dying in hospital other mechanisms to change implementation are likely to be more appropriate. Such considerations should be discussed as part of the deliberative process of research prioritisation.

5.5 IS THE RANE TOOL GENERALIZABLE?

In addition to avoiding a large technical or resource barrier, another requirement of rapid methods for research prioritisation is that they are generally applicable to the diverse range of research proposals received by a research funding agency such as NETSCC. The rapid methods (and RANE tool) were designed to address an initial set of six proposals provided by NETSCC (P1-P6). To test their generalisability; we investigate whether their scope is sufficient to address an additional set of NETSCC proposals.

5.5.1 ADDITIONAL SET OF NETSCC PROPOSALS

Six additional retrospective proposals were received from NETSCC, however permission from the authors was only granted for three of the six new proposals. These three are listed below and Table 5.1 summarises the research design, comparisons and primary outcome for each.

- Proposal 7 (P7): Trial of treatments for high risk non muscle invasive bladder cancer.
- Proposal 8 (P8): Trial investigating the timing of labor inducement in order to limit risk in hypertensive pregnancy.
- Proposal 9 (P9): Trial of a diagnostic test to indicate the use of adjuvant radiotherapy for ductal carcinoma.

Table 5.1 Summary of new proposals received from NETSCC P7-P9

Proposal	Research design	Comparisons	Primary outcome
P7: Bladder cancer	2 arm RCT	· BCG treatment · Hyperthermic chemotherapy	Survival: time free from recurrence, progression or death at 2 years.
P8: Hypertensive pregnancy	2 arm RCT	· Planned delivery at 38 to 41 weeks · Monitoring to 40 weeks	(Mother) Binary: composite of poor maternal outcomes. (Baby) Binary: neonatal care unit admission >= 4 hours.
P9: Test directed radiotherapy	2 arm RCT	· Standard treatment with adjuvant radiotherapy treatment · Test directed decision making based on Oncotype DX DCIS score	Survival: recurrence free interval of carcinoma or invasive breast cancer.

Each of the three proposals are for RCTs in which the largest number of treatment arms is three. The RANE tool can currently analyse a maximum of four arms, therefore, the methods are sufficient to address these research designs. The rapid methods (and the RANE tool) are currently capable of analysing binary, continuous and survival outcomes. The primary outcomes for each of the new proposals are either binary (P8) or survival (P7, P9) and the so the rapid methods are applicable to all three. P8, however, aims to investigate an intervention in pregnancy and so involves two primary outcomes, one for mother and one for baby. The research funding decision can be usefully informed by separately presenting the analysis for each of the primary outcomes. However, rapid methods do not currently exist to simultaneously analyse both primary outcomes jointly. P9 is designed to investigate the value of a clinical test i.e. a diagnostic technology. It is possible to use the RANE tool to understand the value of this research by comparing the outcomes in the control and treatment arms. However, because the value of diagnostic tests cannot be separated from the treatments used in response to their results, methods are currently in development to better understand their mechanisms of value (Phelps and Mushlin, 1988; Soares et al., 2018). Future iterations of the RANE tool should reflect these developments where possible.

5.6 DISCUSSION

An important barrier to the use of VOI methods is the time and expertise required to construct decision models and carry out VOI analysis. This chapter describes freely available software which can be used for research prioritisation across disease areas and jurisdictions. The tool provides a non-technical interface which rapidly calculates and interprets VOI outputs. The reduced time and technical barriers can facilitate the wider use of VOI by both research funders (such as NETSCC in the UK) and research applicants (such as academic organisations applying for public funding for research).

As the RANE tool is based on the methods described in Chapters 2 to 4, it has the same limitations as these methods (discussed in Chapters 2, 3 and 4). These limitations reflect trade-offs between complexity and simplicity which are required to keep the number of required inputs to a minimum. One route to move beyond this trade-off is to build a tool which is more disease specific. This is the approach taken by the developers of the VICTOR tool (Basu et al., 2018b). VICTOR implements a relatively complex decision model but the user is not required to search for a large number of inputs. This is possible as the VICTOR tool is specific to cardiovascular outcomes and so a general disease process and a number of the required parameters such as life expectancy and all-cause mortality are pre specified within the tool. This represents an important and useful direction for research but its use in research prioritisation is currently limited by the specificity of the tool to a specific disease area (cardiovascular disease) and the potential opacity of the pre specified inputs to users.

Section 5.5.1 tested the generalisability of the RANE tool to a new set of NETSCC proposals. It was found that though there are is room for further research the tool is sufficient to provide useful analysis for each of the proposals in the new set P7-P9 supporting its claim to generalisability. However this is based on only three proposals, more intensive piloting of the tool is required to understand how generalizable the tool is and to guide further development where necessary.

The methods described in this thesis (and embedded in the RANE tool) provide an expected upper bound on the value of research when all uncertainty is resolved. As research projects have a finite sample size, they will only ever partially resolve the

total uncertainty. Methods to adjust research value for sample size and other aspects of research design are well developed (McKenna and Claxton, 2011; Strong et al., 2015) and further work is required to incorporate this into RANE.

The RANE tool facilitates calculation of VOI outputs for up to four treatment options (inclusive of the baseline treatment). However, it currently forces the assumption that the relative effects for each intervention are uncorrelated with each other. This is important as there may be good scientific reasons to expect that there are relationships between some of the treatment effects which should be captured by the analysis. These relationships may be due to (i) combination treatments or (ii) correlations between baseline and interventions. There are good reasons to expect that there are correlations between combination treatments and their component parts. For example in case study P4 (introduced in Section 4.2.2) three interventions were compared to placebo; Exenatide 2mg, Telmisartan 40mg and the combination of Exenatide 2mg and Telmisartan 40mg. For combinations of treatments, there may be substantive evidence and/or clinical rationale to suggest that the combination of treatments will work noticeably better than any of the treatments alone. This information should be incorporated into the analysis (Mills et al., 2012; Thorlund et al., 2017).

For binary outcomes there may also be correlations between the baseline outcome and the relative effects (Riley, 2009). This is relevant in cases where an intervention has a greater relative effect in patients who are high or low risk and can be important in understanding subgroups.

Both of these relationships between treatments discussed here can potentially influence VOI estimates and therefore will affect the value of research. An extension which would allow analysts to address both of these statistical issues (and many others) would be to permit users to upload a matrix of simulated results directly into the RANE tool. This output matrix can be produced either directly by a statistical software or may require additional modelling of results through Cholesky decomposition (Briggs et al., 2006). This approach maintains all relevant correlations between parameters captured by the analysis and can be carried out for data analysis or for an expert elicitation exercise. This would allow the RANE tool to

provide estimates of the value of research for a wide range of statistical models and so is an important extension for future research.

Chapter 6

Informing early access and research decisions without full economic modelling

6.1 INTRODUCTION

Chapters 2 to 4 have provided practical methods to facilitate transparent and accountable research prioritisation by public bodies such as NETSCC in the UK and PCORI in the USA. By addressing binary, continuous and survival primary outcomes in addition to RCTs and feasibility/pilot studies, the methods have been developed to be generally applicable to a range of research applications. Rapid methods to calculate the benefits of research in terms of costs and QALYs have been proposed in order to inform research prioritisation across proposals. Chapter 5 introduced an online tool which can implement each of the models above in order to reduce the time and technical barriers to carrying out the analysis.

However, not all health research is funded by public prioritisation bodies such as NETSCC. Approximately 30% of applied biomedical research capital is allocated by the public sector with 60% allocated by the private sector (pharmaceutical companies and medical device manufacturers) and 10% allocated by private and non-profit organisations (Røttingen et al., 2013). Previous chapters have focused on methods to improve accountability in public sector research prioritisation. In this chapter we include private and charity sector research funding into the framework and show how research prioritisation decisions can depend on which sector bears the research costs. Price is an important consideration for the private sector and so the relationship between price and approval of treatments will be explored.

Additionally, the previous chapters have implicitly assumed that decisions to recommend an intervention for widespread use and decisions to carry out additional research on that same intervention are independent. However, there are cases in which research cannot be carried out at the same time as the treatment is available for widespread use. The implications of this constraint are also explored in this chapter. By considering decisions about technology approval in addition to research decisions, this chapter moves away from a HTA funding panel setting into a more

comprehensive HTA decision making context in which approval and research decisions are made simultaneously.

6.2 EARLY ACCESS AND RESEARCH DECISIONS

In the current policy environment there is pressure on public payers to grant early access to new, high cost treatments of uncertain benefit (Gyawali and Kesselheim, 2018; McCabe et al., 2016; Prasad, 2017). Some payers have responded by developing accelerated access schemes such as the cancer drugs fund and the accelerated access collaborative (Grieve et al., 2016; NICE, 2016) in the United Kingdom and the food and drug administration (FDA) accelerated approval program in the United States (Gyawali and Kesselheim, 2018; Johnson et al., 2011). Industry representatives argue that these arrangements facilitate faster access to vital medicines (Svensson et al., 2013). However, these policy responses have come under criticism and have been subject to intense discussion across social media and mainstream media outlets (Barczyk, 2018; Edwards, 2018).

Deciding which treatments to provide early (market) access to and which treatments to research further requires trade-offs. This is true regardless of whether the health system is privately or collectively funded and whether the health budget is fixed or elastic (Basu and Sullivan, 2017). Treatments which are more clinically effective may also come with additional out of pocket costs and/or a higher risk of side effects compared to current treatment options. Trade-offs must also be made when funding research as research budgets are limited and choosing to fund research in one clinical area, means that research funding is not available for another clinical area. Further, resources spent on research could have been spent on direct provision of care (see Section 3.3.2). Research decisions can also involve trade-offs between current patients and future patients in the case of early access decisions. The benefit of granting early access to new treatments which appear to be effective based on current evidence is that potentially worthwhile treatments can be quickly provided to patients without undue delay (Claxton et al., 2016; Eckermann and Willan, 2008a). A potential cost of granting early access is that some types of research are not possible after widespread access has been granted (discussed further in Section 6.3.4). This means that even in cases in which the public sector does not bear any

research costs, trade-offs must be made; whether to provide immediate access to potentially beneficial treatments or to delay access to do additional research.

VOI methods provide a means to estimate the health benefits of reducing uncertainty. This provides an explicit, evidence-based basis to inform trade-offs between the value of early approval and the value of further research to future patients. To provide a coherent framework for decision making which takes account of these benefits and costs, a series of assessments have been outlined to structure deliberation (Claxton et al., 2016; Walker et al., 2012). Historically, decision makers have usually considered only binary decision options, either approve or reject the new treatment. However, decision options exist that facilitate early access to new treatments while taking account of the health consequences of using sub-optimal treatment options. These include the conditional coverage options such as "Only in Research" (OIR) and "Approval with Research" (AWR). The former only allows the use of new treatment in a research setting. The latter approves the treatment for widespread use on the condition that additional evidence is collected (McKenna et al., 2015). The decision about the appropriate guidance is determined through a series of seven assessments which are listed below and described in detail in (Claxton et al., 2012).

1. Is the technology worthwhile?
2. Are there significant irrecoverable costs?
3. Does more research seem worthwhile?
4. Is the research possible with approval?
5. Will other sources of uncertainty resolve over time?
6. Are the benefits of research greater than the costs?
7. Are the benefits of approval greater than the costs?

In previous work, these assessments were informed with reference to a full economic model which reported results in terms of costs and QALYs. The barriers to developing full models were discussed in Chapter 2. In this chapter, we demonstrate how to inform the seven required assessments without a full economic model by extending the methods for binary outcomes described in Chapter 2.

There are two primary aims of this chapter. First, we demonstrate how the assessments required for research and early approval decisions can be informed

without a full economic model. Second, we show how early access and research decisions depend on whether costs of additional research are borne by the public sector or some other entity (e.g. from the private or charity sector). It should be noted that this analysis takes a health system perspective in which only (public sector) healthcare costs and health care benefits are taken into account[62] (NICE, 2013; Willan and Eckermann, 2012). The methods are illustrated using the P1 case study described in detail in Chapter 2.

[62] This more narrow perspective is taken here for two reasons: first, this is the perspective recommended by NICE; secondly including other sectors, such as the pharmaceutical industry, would substantially increase the complexity of the analysis. This is beyond the scope of the current thesis but is an important area of further research.

6.3 RESEARCH AND APPROVAL CHECKLIST

6.3.1 POINT 1 - IS THE TECHNOLOGY EXPECTED TO BE WORTHWHILE?

As described in Chapter 2, there are two treatments considered for use in P1; current practice which is late PTP and a new treatment, early PTP. The first step is to assess whether implementing, the new treatment is expected to provide an overall gain or loss to population the health system serves. In the approach outlined here, the starting point to answer this question is the current evidence on the primary outcome, which in the case of P1 is the probability of functional recovery. The relative effect of the new treatment (early PTP) on functional recovery is described by an odds ratio which is 1.09 with 95% CI from 0.23 to 5.24. To understand the population health consequences of implementing the new treatment requires consideration of the absolute number of functional recoveries expected with both the new treatment and current practice. This requires an estimate of the probability of functional recovery with current practice and an estimate of the number of patients which are expected to be affected by the decision each year[63]. As described in Section 2.4 of Chapter 2, the probability of functional recovery with current practice is expected to be 55.1% with uncertainty, which combined with an annual incidence of 8,800 is expected to result in (8,800 x 55.1% =) 4,849 functional recoveries each year with late PTP.

An estimate of the expected probability of functional recovery with the new treatment can be calculated by combining samples from the odds ratio and samples from the baseline uncertainty as shown in Chapter 2. This results in an expected probability of functional recovery of 56.5% for early PTP. Therefore current evidence suggests that the new intervention is expected to result in (8,800 x 56.5% =) 4,972 functional recoveries each year. This implies that implementing the new treatment is expected to result in a *gross* increase of (4,972 - 4,849 =) 123 functional recoveries per year as shown in Figure 6.1.

This indicated that there may be health benefits from providing immediate access to the new treatment. However there may be important differences between the treatments which are not captured by the primary outcome. This may include differences in side effect profile, relative price, out of pocket expenditures etc. These

[63] The treatment of traumatic brain injury considered in P1 is an acute condition and so there is no significant prevalent population. The method described here can theoretically be extended to take account of a prevalent population.

differences which are not captured by the primary outcome will be referred to here as secondary outcomes.

If on balance, the secondary outcomes of the new treatment are unfavourable then the new treatment may be expected to improve the primary outcome to such a degree that it makes up for the unfavourable secondary outcomes. This is the MCD required before practice should change (described in Chapter 2). For binary outcomes the MCD can be expressed in terms of a required odds ratio (OR), risk ratio (RR) or risk difference (RD). As decision making is concerned with absolute differences in outcomes between treatments, MCD is most naturally expressed as a RD i.e. the difference in absolute probability of the primary outcome between two treatments. For a given baseline probability, a required OR can be converted to RD and vice versa. Naturally, as an OR of 1 indicates no difference between treatments, this corresponds to a RD of 0%. However, the relationship between changes in OR and changes in RD is highly non-linear and so exact closed form solutions are challenging. Methods to translate between an MCD expressed in RD and OR are provided the Appendix. These functions require simulation methods which are provided as R functions. Figure 6.1 illustrates that each value of OR (bottom axis) maps onto a RD (top axis). Unless stated otherwise, MCD will be expressed in terms of RD in this chapter.

The function of the MCD is to trade off in the gains (or losses) in the primary outcome against the losses (or gains) associated with the secondary outcomes. Using the primary endpoint as a unit of account, the *gross* expected gain from the new treatment can be adjusted using the MCD to estimate the *net* health benefits (NHBs) of early access in terms of the primary endpoint.

For P1, the new treatment (early PTP) is expected to be more effective on the primary outcome but also slightly more costly and so a positive MCD may be required before it should be implemented. If a MCD of 0.5% (equivalent to a required OR of 1.02) is required to account of the additional costs associated with the new intervention this means an expected improvement of at least (8,800 x 0.5% =) 44 additional functional recoveries are required each year before the new intervention can be considered worthwhile. Taking account of this MCD, the NHB of access is therefore equivalent to (123 – 44 =) 79 functional recoveries each year,

providing an estimate of the net health benefits of early access each year. This is illustrated in Figure 6.1.

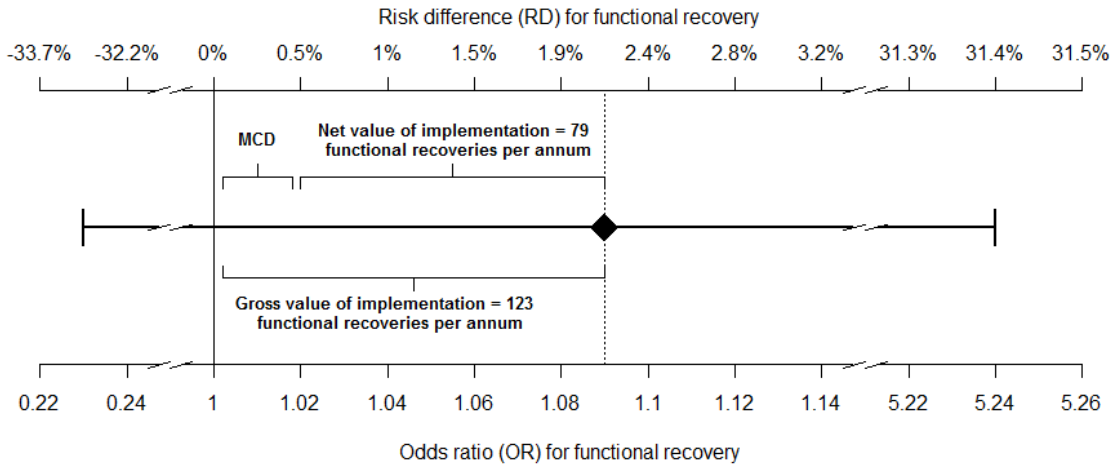


Figure 6.1: The current evidence for the effect of early PTP on functional recovery is shown on both the OR (bottom) and RD (top) scale. For the OR scale the new treatment has a mean of 1.09 and 95% CI from 0.23 to 5.24, for the RD scale the new treatment has a mean of 2.1% and 95% CI from -33% to 31.4%. In gross terms early PTP is expected to provide 123 additional functional recoveries each year. After adjusting for the MCD (which is 1.02 on the OR scale and 0.5% on the RD scale) the net functional recoveries associated with approval are 79 each year. CI, confidence/credible interval. OR, odds ratio. RD, risk difference.

As this treatment decision will have consequences beyond one year, this should be taken into account to understand the benefits of early access. If the treatments are expected to be in use for approximately 15 years with a discount rate of 3.5%, the total discounted population over time is therefore 101,353 patients[64]. The gross gain in functional recoveries with the new treatment is (101,353 x 56.5% - 101,353 x 55.1% =) 1,419 functional recoveries. With a MCD of 0.5% this would be a net gain of (1,419 - 101,353 x 0.5% =) 912 functional recoveries.

As relative price, among other factors, is a component of MCD (see Section 3.3.4), it is useful to illustrate how the benefit of approving the new treatment changes over a range of MCD values. Figure 6.2 below illustrates the net functional recoveries associated with the current treatment (equivalent to a Reject decision) and implementing the new treatment (equivalent to an Approve decision).

[64] $\sum_{t=1}^{15} 8800/(1+0.035)^t$ = 101,353

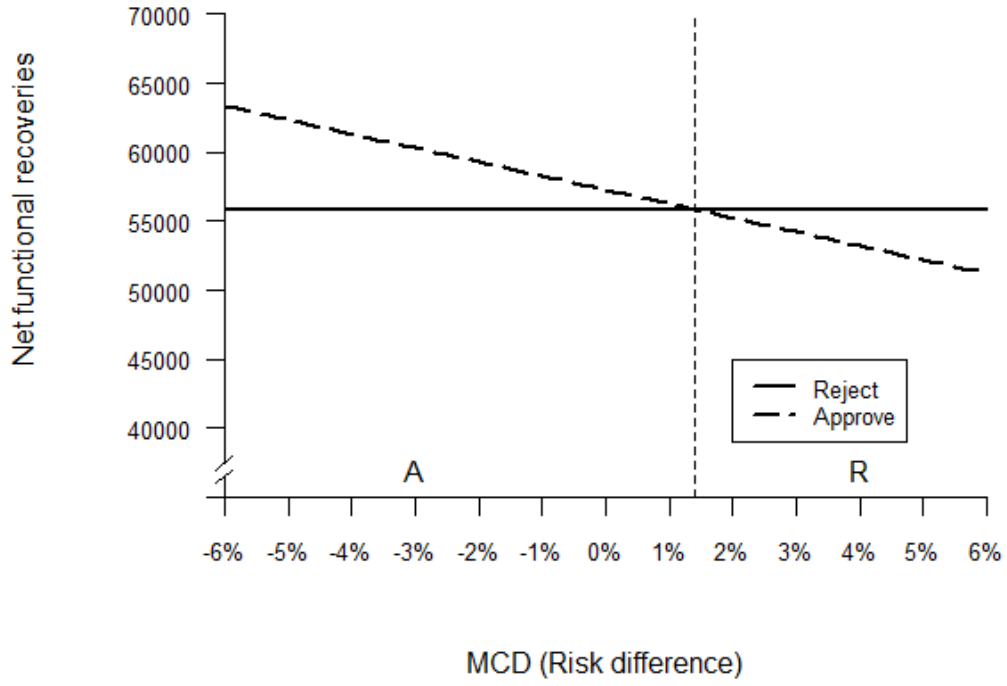


Figure 6.2: Net functional recoveries from approving or rejecting early PTP for different possible MCD values. If the appropriate MCD is below 1.4% then there is a net gain to early approval, otherwise there is expected to be a net loss and the new technology should be rejected. MCD, minimum clinical difference. A, approve. R, reject.

Figure 6.2 illustrates that the number of functional recoveries expected from rejecting early PTP is the expected functional recoveries from keeping late PTP which is (101,353 x 55.1% =) 55,846 functional recoveries in total. It also illustrates that given differences between the treatments if the appropriate MCD is below the Approve | Reject threshold of 1.4% then there is a net gain in approving early PTP. This threshold is simply the risk difference between the alternative treatments with current evidence, in the case of P1 this is 56.5% - 55.1% = 1.4%. If the appropriate MCD is above 1.4% (due to the new treatment having a much larger price for example) there is expected to be a net health loss and the new technology should be rejected. It is not necessary to have perfect knowledge about the exact MCD to understand whether a technology expected to be worthwhile or not, what is required is a judgement about whether the MCD is likely to be below the Approve | Reject threshold. It should be noted that in Figure 6.2 the MCD may be negative, for a treatment to justify a negative MCD it must demonstrate substantial improvements over the comparator e.g. lower relative price or reduced side effects (see Section 3.3.4 for a method of informing an appropriate MCD).

6.3.2 POINT 2 - ARE THERE SIGNIFICANT IRRECOVERABLE COSTS?

An assessment of whether there are irrecoverable costs along with their potential significance is also required for decision making. Irrecoverable costs are those costs that, once committed, cannot be recovered if guidance is changed at a later date (Eckermann and Willan, 2007; McKenna et al., 2015; Thijssen and Bregantini, 2017).

Irrecoverable (opportunity) costs are present when initial per-patient losses are compensated by later gains and can be captured by the MCD. The additional costs associated with early PTP are borne immediately at the time of treatment with the benefits of functional recovery accumulating gradually over time. Figure 6.3 illustrates the "investment profile" of how these net functional recoveries accumulate over time for a MCD of 0.5% and 0.1%. From Section 3.2, it takes up to 19 years for the full benefits of functional recovery to be realised. It is assumed that the benefit from the primary outcome accumulates at a constant rate over time while accounting for the discount rate of 3.5%[65]. For an MCD of 0.5%, the initial losses are in excess of the immediate health benefits in the initial period of treatment. These losses are gradually offset with a "breakeven" point of 14 years. For a MCD of 0.1% early PTP breaks even immediately as the initial MCD penalty is modest. The figure also illustrates that for both MCD values, it takes 33 years for the incremental net functional recoveries to reach their long run values; 912 and 1,318 functional recoveries for MCD 0.5% and 0.1% respectively.

[65] The present value of one functional recovery is defined to be 1

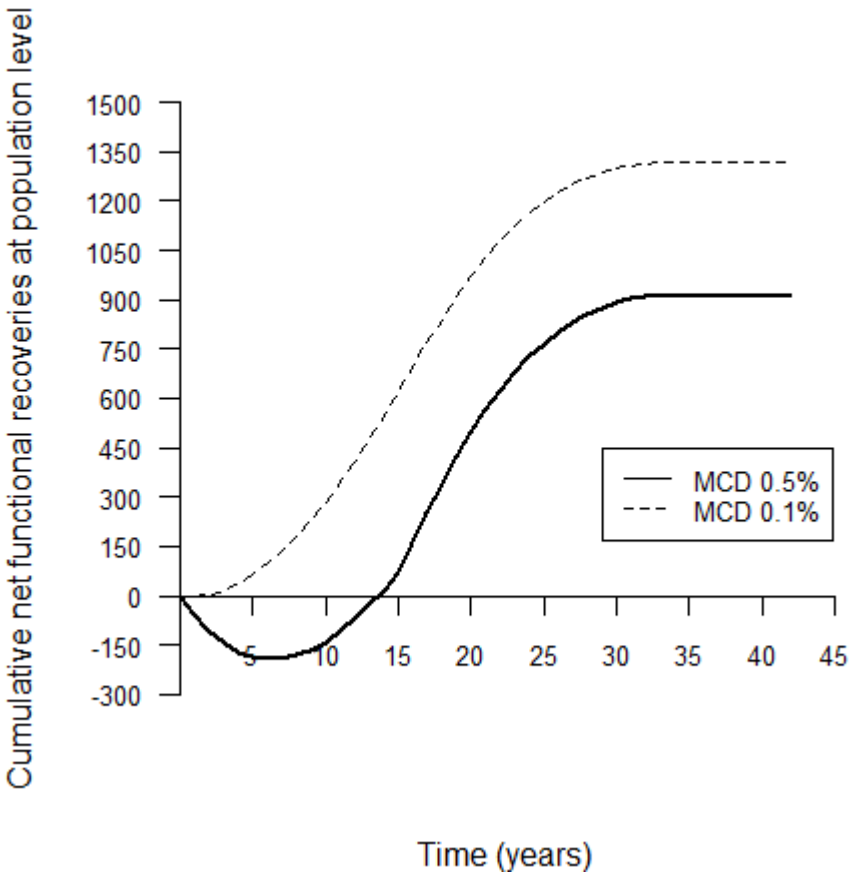


Figure 6.3: Cumulative incremental net functional recoveries of early PTP for the population. The initial costs of early PTP are captured the MCD. For the larger MCD of 0.5% (which would imply larger costs) the initial costs are in excess of the immediate health benefits. These negative NHBs are gradually offset population gains in health after 14 years. For a MCD of 0.1% the initial health benefits immediately compensate for the initial costs incurred. MCD, minimal clinical difference.

The potential significance of the irrecoverable costs illustrated in Figure 6.3, depends on whether there is sufficient flexibility in when a patient's treatment can be initiated (McKenna et al., 2015). If the treatment of a patient can be postponed until uncertainty is resolved, then the initial per-patient losses can potentially be avoided (i.e., they are potentially significant). If the decision to treat cannot be delayed, however, these type of irrecoverable costs cannot be avoided; thus, they will have no influence on the type of guidance (i.e., irrecoverable costs are present but are not potentially significant) (Claxton et al., 2016). Early PTP, is a treatment for acute traumatic head injury in which there is insufficient flexibility to delay the initiation of treatment for presenting patients until the results of research reports become available or other sources of uncertainty resolve. Therefore, any irrecoverable

opportunity costs exhibited by early PTP should not be judged to be potentially significant because they cannot be avoided by delaying the initiation of treatment for particular patients. It should be noted that Figure 6.3 illustrates the investment profile assuming that any costs or side effects are borne at the time of treatment. This is appropriate in the case of P1, however, in other clinical decisions additional costs or side effects may be borne long after initial treatment. In these cases the MCD penalty could be apportioned over the patient time horizon and consequently the investment profile for these treatments would be differ from Figure 6.3.

Irrecoverable costs may also exist at the collective level in the form of one off capital purchases and can be accounted for in terms of the primary outcome. This requires a judgement about the additional number of primary outcomes (ΔP0) which are sufficient to offset the health losses associated with the irrecoverable capital cost. Similar to Equation 6.1, the judgement required can be usefully informed by explicit reference to the capital costs incurred ($\Delta C_{Capital}$), k and INHB.

$$\Delta PO = \frac{\Delta C_{Capital}}{k.INHB}$$

For example if a capital investment of £16,000,000 was required to implement early PTP, in terms of primary endpoints this is equivalent to a required increase of (£16,000,000/(£15,000 x 15.86) =) 67 additional functional recoveries. From point 1, with a MCD of 0.5% early PTP is associated with a net gain of 912 functional recoveries, taking account of the one off capital costs above this would reduce to (912 – 67 =) 845 functional recoveries. Assuming that capital costs were incurred in the initial period, this would change the investment profile in Figure 6.3 by shifting the curves down by 67 functional recoveries meaning that the origin point for both curves would be -67. These capital costs are allocated proportional to the number of individuals that are expected to be treated during the lifetime of the equipment. Treating these upfront capital costs as if they are paid per individual will have no effect on the expected benefit of the treatments as long as guidance is not changed. If the initial approval decision is withdrawn before the end of the lifetime of the equipment (due to research reporting or price changes), the expected future patients will not receive treatment with the technology and so the total cost of the must be allocated to the smaller number of treated individuals which will increase the cost

per patient treated (McKenna et al., 2015). For P1 there does not appear to be any one off capital purchases required for implementation.

6.3.3 POINT 3 - DOES MORE RESEARCH SEEM WORTHWHILE?

Point 3 requires an assessment of whether the potential benefits of conducting additional research are worth the costs of this research. If the public sector does not bear the costs of research then more research will always be worthwhile (see Section 6.3.3.1). However, as trade-offs between the benefits of early access and the benefits of additional research may be required for decision making, some assessment of the benefit of research is required even in this case.

This requires judgments about how uncertain a decision to approve or reject the treatment might be based on expected NHBs. From the assessment in point 1, the new treatment is expected to be worthwhile; however the OR is highly uncertain with a 95% CI from 0.23 to 5.24 meaning that there is a chance that a decision to approve the new treatment will be incorrect. Some assessment of the consequences of uncertainty is required. As illustrated in detail in Chapter 2, this involves drawing samples from the distribution of the OR and combining this with the baseline probability of the outcome to characterise the distribution of the consequences of uncertainty. For each sample, the expected benefits of the new treatment are adjusted to take account of the MCD and calculate the *net* functional recoveries associated with the new treatment. For an MCD of 0.5%, this results in an estimate for the maximum value of research of 590 functional recoveries per year. Extending the yearly consequences of uncertainty over the 15 year time horizon, the maximum value of research is estimated to be 6,791 functional recoveries gained over the full time horizon. The above estimate of the value of research is for an MCD of 0.5%. As discussed under point 1, the appropriate MCD value may be reasonably disputed and/or under control by the manufacturer through price; therefore Figure 6.4 shows how the upper bound on the value of research changes for a range of MCD values.

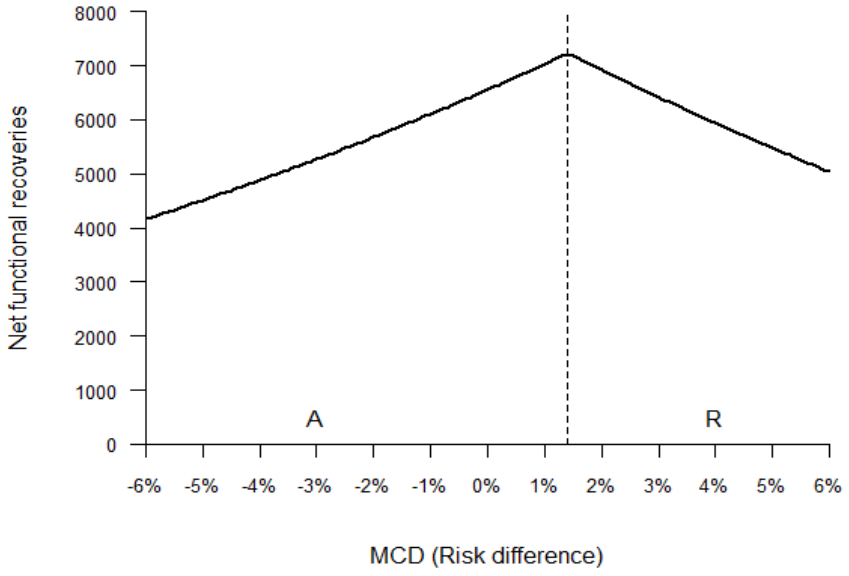


Figure 6.4: Illustration of the upper bound on the value of research, in terms of net functional recoveries for a range of MCD values. The value of research reaches a maximum at a MCD of 1.4% which is the Approve | Reject threshold. MCD, minimum clinical difference. A, approve. R, reject.

As Figure 6.4 illustrates, the value of research reaches as peak of 7,222 functional recoveries at an MCD of 1.4%. This is expected as this is the Approve | Reject threshold and so is the point at which the choice between the two treatments is maximally uncertain (Briggs et al., 2006; Claxton, 1999).

6.3.3.1 *Who pays for research?*

Figure 6.4 does not yet incorporate the costs of research. The decision about whether additional research is worthwhile depends on the opportunity cost of research funding resources. This in turn depends on whether research is funded by the public or from some other source such as the private sector (health care firms) or the charity sector.

Private companies fund research in an attempt to demonstrate the efficacy of their technologies. This incentive for research exists because of the patent system. In the present system, patents are awarded to private companies for developing novel technologies (e.g. drugs or medical devices). These patents provide these companies with monopoly status for the production and sale of novel technologies for a limited time period. As the holder of a patent enjoys monopoly profits during the period of

protection, private companies stand to gain from widespread use of their products. This creates an incentive for the private sector to fund sufficient research to gain access to health systems (Rothery et al., 2017). If manufacturers fund the required research on their own products then these research costs do not fall directly on the health system and so from the perspective of the health system research will always be worthwhile if there is any uncertainty in the decision[66]. This is also the case if charities, acting on behalf of their stakeholders, fund research on promising technologies[67].

If research is publically funded then there will be health opportunity costs associated with funding research. From the proposal, research appropriate to address the question is expected to cost £2,854,000[68]. This allows for an estimate of the upper bound for the value of research by dividing the cost of the research by the expected upper bound for the value of additional research. For a MCD of 1.4% the upper bound for the value of research is (£2,854,000/7,222 =) £395 per functional recovery. The value of research is shown for a range of MCD values.

[66] It is assumed here that treatment price is not renegotiated in response to the results of research. However, as will be discussed in Point 5, price may change in response to generic entry.

[67] As both charities and manufacturers have limited budgets, it may be the case that paying for research on treatment A will result in less funding available for treatment B. This means that carrying out additional research on treatment A can have consequences for the health system even if the health system does not bear the financial costs of research. To reduce the complexity of the analysis, this channel is assumed not to operate in this case. Further research is required to characterise this multisector dynamic.

[68] This approach to calculating the value of additional research is based around the idea of a "definitive trial" i.e. it assumes that the size and therefore the cost of a trial which will adequately address the uncertainty is given by a power calculation or more sophisticated Bayesian methods (O'Hagan et al., 2005; Sutton et al., 2007). This is in contrast to other applications of VOI to decision making which can help inform the size and cost of the required trial (Briggs et al., 2006; Eckermann and Willan, 2007; McKenna and Claxton, 2011). Further research is required to provide a framework to inform the size and design of trials without full economic modelling.

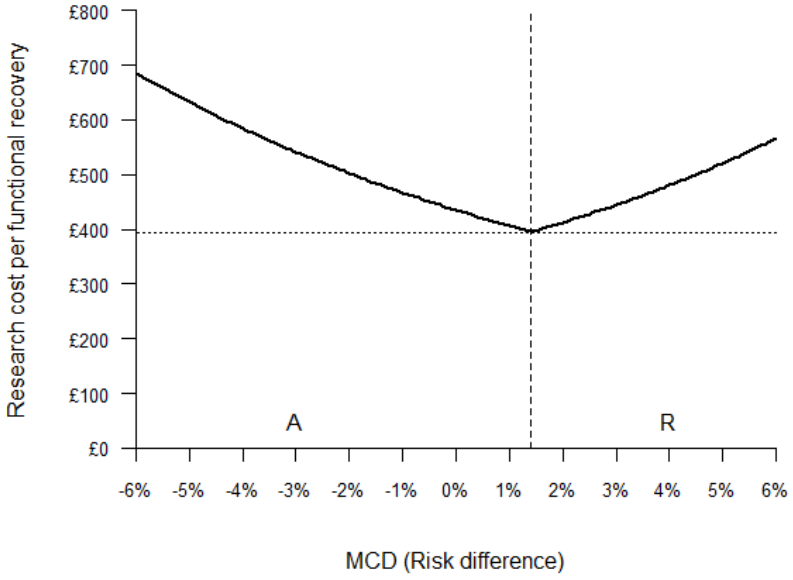


Figure 6.5: Cost per functional recovery gained from research for a range of MCD values for research which is expected to cost £2,854,000. The cost per functional recovery from carrying out research reaches a minimum at a MCD of 1.4% which is the Approve | Reject threshold. MCD, minimal clinical difference.

Figure 6.5 shows that the cost per functional recovery from research declines to a minimum at MCD 1.4% which is the Approve | Reject threshold. At this point, £395 of research expenditure is required to gain an additional functional recovery. As this is the point at which the decision is most uncertain, if the research is not good value at this point then it will not be good value at *any* MCD value. Therefore a necessary condition for research funding is that research is considered worthwhile at this MCD.

Public research funds will either come from a dedicated research budget or general health expenditure. In the case of a dedicated research budget, the value of research must be compared to the value of other research which could be funded with these same resources. For P1, the maximum value of research has been estimated to be £395 per functional recovery. Whether this represents good value for money depends on the value of the alternative research proposals which could be funded with the £2,854,000 required for this research. This judgement involves implicitly comparing the benefits of primary outcomes across different disease areas as discussed in Chapter 2. To aid this process, the value of research can be converted from the primary outcome to generic health units (such as QALYs) using an estimate of the

INHB associated with the primary outcome. From Chapter 3, the health gain from an additional functional recovery is approximately 15.86 QALYs. By arithmetic this means that an upper bound for the value of the proposed research is £395/15.86 = £25 per QALY. This compares very favourably the alternative research proposals discussed in Chapter 3, indicating that more research in this area appears worthwhile.

If the costs of funding further research fall on the budget for general health care expenditure (e.g. the NHS in the UK) then the health benefits of research must be compared to the health benefits of general health system activities. Empirical research estimates that general NHS expenditure requires approximately £15,000 to produce one QALY (Claxton et al., 2015b). Therefore the cost per QALY from funding the planned research (£25 per QALY) compares very favourably to the alternative and the research appears worthwhile.

6.3.4 POINT 4 - IS THE RESEARCH POSSIBLE WITH APPROVAL?

This point entails an assessment of the type of evidence which is required and a judgment about whether the research required can be conducted while the technology is approved for widespread use. An important consideration is whether randomisation is required. If more precise estimates of relative treatment effect are required then a RCT may be necessary to avoid selection bias. However, randomised research may not be considered ethical once a technology is approved for general use (McKenna et al., 2015) which would rule out AWR.

Understanding which type of research is required requires an assessment of the importance of different sources of clinical uncertainty. For binary primary outcomes, there are two sources of clinical uncertainty (i.e. parameters) which can be addressed by further research; uncertainty in the relative effect estimate (the OR in this case) and uncertainty in the baseline probability of the primary outcome. Assessing the importance of these parameters entails judgments about i) how important these parameters are to the decision; ii) the values these parameters would have to take to change the decision; iii) how likely it is that these parameters would take these values; and iv) what the health consequences would be if they take these values. This analysis can provide an estimate of the health gain if the uncertainty could be instantly resolved (McKenna et al., 2015).

With current evidence (point 1), it appears that early PTP provides greater benefit than late PTP. However, if the relative effect and/or baseline probability parameters take certain values then the optimal treatment would switch to late PTP. The simulated probabilities of functional recovery for each treatment described in point 1 can be analysed to determine the contribution of each parameter to overall uncertainty as measured by the expected value of perfect parameter information (EVPPI). Uncertainty in relative treatment effect contributes most to the probability of making an incorrect decision. This is because the baseline probability determines the absolute health effect of the treatments whereas the relative effect can take values which can change the relative benefits of the two treatments.

An understanding of the expected health consequences of this uncertainty is also required. The simulated probabilities of functional recovery for each treatment described in point 1 can again inform this assessment through combining the

uncertainty in the potential values of the parameters with their importance in changing decisions. The health consequences of resolving uncertainty in the relative effect parameter are estimated to be 6,936 functional recoveries and are 146 functional recoveries for the baseline probability[69]. As better estimates of relative treatment effect are most important, randomised research is required to address this question. Therefore a judgement is required as to whether randomised research can be carried out alongside the approval of early PTP. If it is considered unethical to enrol patients into a trial when one comparator is already considered superior and approved for use, randomised research may not be possible. In this case, OIR is the remaining research option. Under points 6 and 7 we will explore the judgements required when AWR is and is not possible.

[69] Note the separate contribution of each parameter (6,936 + 146 = 7,082) does not equal the overall consequences of uncertainty; 7,222 functional recoveries. This is because the value of resolving uncertainty in both parameters simultaneously is greater than the value of resolving the uncertainty in each of them individually.

6.3.5 POINT 5 - WILL OTHER SOURCES OF UNCERTAINTY RESOLVE OVER TIME?

Point 5 requires a judgement about the likelihood of future changes that will influence the relative benefits of the alternative technologies and the expected benefits of research. These uncertain future changes include: i) price changes of the technologies, ii) the appearance of new technologies which make existing technologies obsolete or change their relative benefits, and iii) other relevant trials reporting (McKenna et al., 2015). As this is described in depth elsewhere (Claxton et al., 2016, 2012; McKenna et al., 2015), here we focus on illustrating where MCD can be used to address these judgements.

As discussed in point 1, changes in relative price can be reflected in MCD with a higher price requiring a larger MCD. Therefore, future changes in relative prices (say due to generic entry) can also be captured by future changes in MCD. As shown, changes in MCD influence not only whether the treatment is expected to be worthwhile but also uncertainty and the potential benefits of research to future patients. For patented products, a significant price reduction is expected at patent expiry due to entry of generics (Claxton et al., 2012). This translates to a reduction in the MCD for the technology in the future.

If generic entry occurs before the results of the planned research report, the expected benefits of research will not be realized as the decision to approve the technology will be less uncertain. Naturally, this reduction in uncertainty reduces the value of the planned research. If a technology is expected to be worthwhile, a future price reduction (say, due to patent expiry) will reduce the value of additional research as the technology will become more worthwhile at a lower price and so the uncertainty in the decision will fall.

For example early PTP is expected to be just worthwhile with an MCD of 1.4%. At this MCD the upper bound for the value of research is 7,222 functional recoveries over 15 years (from Figure 6.4). If a price reduction occurred which reduced the MCD to 0% the upper bound for the value of research would be 6,559 functional recoveries over 15 years. The value of research has fallen as the decision is less uncertain at a MCD of 0% (reduced price makes the new treatment more attractive). This can affect the appropriate guidance. For example, OIR may be changed to Approve if the benefits of research fall such that the benefits of early approval

become greater than the value of additional research. For these reasons, information about large changes in price should be considered in decision making.

The entry of a new technology will also influence the expected benefits of treatments and the value of further research. There are two extreme scenarios which may arise. First, a new technology may enter in the future which makes the current treatments obsolete. In this case the value of implementing early PTP will fall to zero at the point at which the new treatment enters, also the value of the research will also tend to zero at this point in time[70]. For research to be of any value, it must report before this point, and it will only affect patients in the window after the research reports and before the new product enters. The second scenario is when a technology enters which is similar to early PTP. In this case the value of research investigating early PTP vs late PTP will increase as the information gained can be used to help inform the decision about the new treatment too.

Trials that are ongoing, funded, or likely to be funded are also relevant because they may have an influence on recruitment rates. It is also possible that the results will change the estimate of relative treatment effectiveness when reported (McKenna et al., 2015).

To avoid complicating the illustration of methods, it is assumed that for P1 there is not any expected change in price, new competitors or ongoing research.

[70] As the entire network of evidence should be considered when considering the value of research, additional evidence on obsolete treatment alternatives will not have zero value if it contributes indirect evidence to inform relevant treatment alternatives.

6.3.6 POINT 6 - ARE THE BENEFITS OF RESEARCH GREATER THAN THE COSTS?

Point 6 on the checklist requires a reassessment of the potential benefits of research and a judgment of whether the benefits of research are likely to exceed the costs of research.

The benefits of research will depend on i) the probability that the planned research will be carried out, ii) how long it will take for the results to report, iii) how much of the uncertainty is expected to be resolved , and iv) the impact of the other sources of uncertainty outlined in point 5. In this section, we show how the expected benefits of research (from point 3) can be adjusted to take account of the time to research reporting among other factors[71].

As discussed under point 3, the costs of research may or may not be funded by the health system. If the costs of research are fully borne by the private or charity sector then the benefits of the research will always be greater than the costs from a health system perspective if there is any decision uncertainty. This means that (from a public sector perspective) the benefits of research are always greater than the costs. If the new treatment is expected to be superior to the current treatment (point 1) and there are no significant irrecoverable costs (point 2) then, as the benefits of research are expected to be greater than the costs the guidance depends on whether or not the required research is possible with approval (point 4). If research is possible with approval then AWR may be appropriate (if the treatment is generally considered safe). Otherwise the benefits of immediate access must be compared to the benefits of delaying approval and carrying out additional research. This is illustrated in point 7 and as will be shown requires an assessment of the benefits of research.

For publically funded research, an assessment of the health benefits of research is always required. The health benefits of funding a particular research proposal must be compared to the health forgone from not funding an alternative research proposal and/or the alternative uses these resources could be put to in the general health system (as discussed in point 3).

[71] The benefits of research will also depend on the sample size with larger sample sizes reducing uncertainty to a greater extent than smaller samples. In this analysis, it is assumed that research resolves all uncertainty. Further methodological work is required to relax this assumption.

Whether research is funded and carried out by the private or public sector there can be cases in which the research does not report. This may be because of problems in carrying out the research such as failure of randomisation or problems with recruitment. A reduced probability of research reporting will reduce the expected value of research. Under point 3, the necessary condition for research funding (with total research cost borne by the health system of £2,854,000) is that research is worthwhile at a cost of (£2,854,000/7,222=) £395 per functional recovery. This assumes that there is a 100% probability of the research reporting. If there was a 80% chance of research reporting then the upper bound for the value of research would be (£2,854,000/(7,222 x 0.8) =) £494 per functional recovery. Table 6.1 illustrates how the necessary condition changes as the likelihood of reporting is varied from 100% to 0%.

Table 6.1: Expected upper bound for health benefits of research for different likelihoods of research reporting when MCD is 1.4%.

Likelihood of research reporting	Maximum benefits of research in functional recoveries	Minimum cost per functional recovery from research
0%	0	NA
10%	722	£3,952
20%	1,444	£1,976
30%	2,167	£1,317
40%	2,889	£988
50%	3,611	£790
60%	4,333	£659
70%	5,055	£565
80%	5,778	£494
90%	6,500	£439
100%	7,222	£395

As can be seen above the value of research is highest when it is certain to report and falls with decreasing probability of reporting. If there is an 80% chance of research reporting, the maximum value of research is £494 for an additional functional recovery.

In addition to the likelihood of research reporting, the potential value of research will depend on the time it takes for research to report. The value of the research falls as the research takes longer to report as the population who can benefit from the

information gets smaller. Table 6.2 illustrates how the expected value of research changes as the time to research reporting is varied from 0 to 15 years.

Table 6.2: Expected upper bound for health benefits of research for different research reporting times when MCD is 1.4% and the likelihood of reporting is 100%.

Time to research reporting (years)	Maximum benefits of research in functional recoveries	Minimum cost per functional recovery from research
0	7,222	£395
1	6,616	£431
2	6,031	£473
3	5,465	£522
4	4,919	£580
5	4,391	£650
6	3,881	£735
7	3,388	£842
8	2,912	£980
9	2,452	£1,164
10	2,007	£1,422
11	1,578	£1,809
12	1,163	£2,454
13	762	£3,745
14	374	£7,631
15	0	-

Table 6.2 illustrates that the value of research falls as the research takes longer to report. As the information is assumed to be relevant for 15 years, the information gained from the research has no value after 15 years and so the final row in Table 6.2 has zero benefits of research.

The above analysis assumes that research resolves all uncertainty for both the relative treatment effect and the baseline probability of functional recovery. Due to limited sample sizes, research will not fully reduce all of this uncertainty; therefore, this represents an upper bound on the value of research. As described in point 5, research cannot completely resolve all uncertainty in a decision and there are other sources of uncertainty, such as patent expiry, which can have an influence on the expected value of research.

Likelihood of reporting, time to reporting, future changes and MCD should be taken account of when making approval and research decisions. The appropriate MCD value may be reasonably disputed and/or under control by the manufacturer through price. Therefore Figure 6.6 shows how the appropriate guidance changes for different MCD values when all research costs are borne by the public sector and AWR is possible. For Figure 6.6, the research is assumed to take 5 years as reported in the proposal and we assume there is a 100% likelihood of reporting.

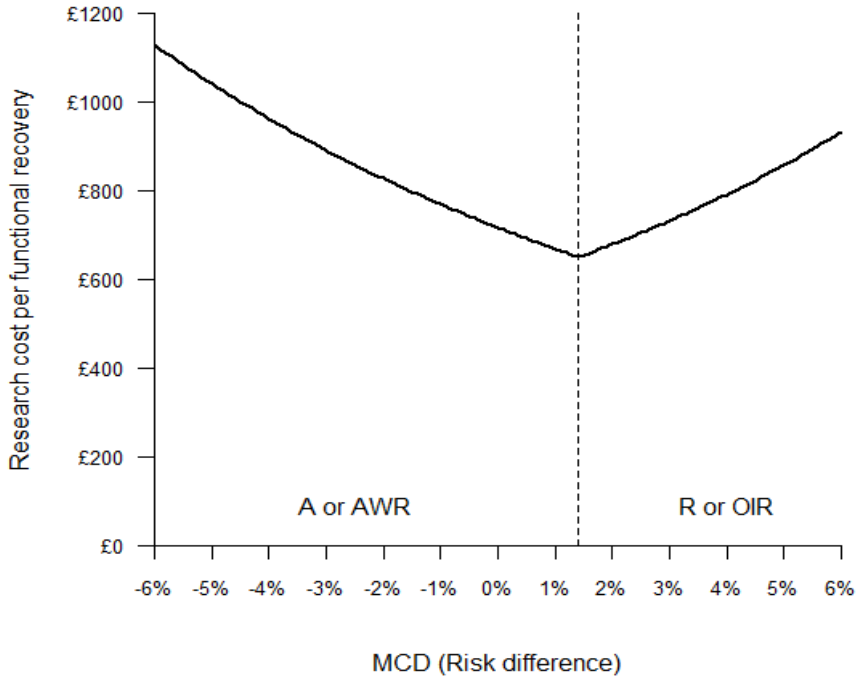


Figure 6.6: Cost per functional recovery gained from research for a range of MCD values for research which is expected to take 5 years to report and costs of £2,854,000. The cost per functional recovery from carrying out research reaches a minimum at a MCD of 1.4% which is the Approve | Reject threshold. MCD, minimal clinical difference. A, approve. R, reject. OIR, only in research.

Figure 6.6 above is the same as Figure 6.5 but adjusted for the time for research reporting. Note that the value of research has declined for every value of MCD in Figure 6.6 compared to Figure 6.5. This is because the estimate of value now takes into account the 5 years it takes to report.

In the scenario illustrated above, the public sector pays for research and AWR is assumed to be possible, therefore research and approval decisions are independent of

one another[72]. This means there is no need to trade off the benefits of early access with the benefits of additional research. If the MCD is below 1.4% (left of the dividing line) then the appropriate guidance is grant immediate access[73]. Whether or not research should be carried out in addition (issue AWR guidance as opposed to Approve) depends on whether additional research is worthwhile. As shown in Figure 6.6, this depends on the MCD. If there is little difference in the secondary outcomes (including relative price), then an appropriate MCD is approximately 0. As this is below 1.4%, immediate access is the optimal policy. From Figure 6.6, additional research should also be carried out (AWR) if £699 per functional recovery is considered value for money. The primary outcome (functional recovery) can be translated to QALYs to aid decision making. As each additional functional recovery is associated with an additional 15.86 QALYs, the upper bound for the value of research is approximately £699/15.86 = £44 per QALY[74]. This compares favourably with the marginal productivity of the general health system; £15,000/QALY (Claxton et al., 2015b), and the other research proposals considered in Chapter 3. Therefore AWR appears to be appropriate guidance. It should be noted that these arrangements have been implicitly assumed in Chapters 2, 3 and 4 with £699 per functional recovery estimated as the value of research in Section 2.4.5.

If research is not possible with approval (discussed in point 4) then AWR is not a viable policy option and so the benefits of early access (Approve) must be traded against the benefits of further research (OIR). This judgement is illustrated in point 7 of the checklist.

[72] It is possible that public sector bodies funding research on patented technologies could have an effect on incentives for the private sector to invest in research before launch. This would mean that research and approval decisions would no longer be independent. Further research is required in this area to understand these incentives and design appropriate policies.

[73] For private companies there is an incentive to reduce price to keep the appropriate MCD below the Approve | Reject threshold.

[74] Note that £44/QALY is the same value as that reported in Table 3.4. This is because calculating the cost per primary outcome using MCD then dividing this by the INHB (as shown here) is equivalent to explicitly accounting for the costs and benefits of the treatment in terms of QALYs (as shown in Chapter 3).

6.3.7 POINT 7 - ARE THE BENEFITS OF APPROVAL GREATER THAN THE COSTS?

If randomised research is required then research may not be possible if the technology is already permitted for widespread use and/or widely used in practice. In this case AWR is not possible and so the benefits of early approval of the new treatment must be compared to the potential health benefits of research through OIR. OIR guidance involves withholding the technology until the results of additional research are clear and then reconsidering the decision. This illustrates the trade-off that must be made between the benefits of immediate approval to current patients and the benefits research which may be forgone as a result of approval (McKenna et al., 2015). Research, which can occur only through OIR can be funded by either the public or some other sector (private or charity). The specific trade-offs and judgements required to make decisions differ depending on the funding source and so are discussed separately below.

6.3.7.1 *Research costs not funded by the public sector*

From a public sector perspective, if the costs of research are borne by manufacturers or charities then they do not result in health opportunity costs[75]. However, trade-offs must be made. Choosing additional research (through OIR) means that access to the new technology is delayed. This delay results in foregone health when the treatment is expected to be effective. If research duration and likelihood of reporting are known, then the judgement about whether the benefits of OIR are greater than the costs of delay can be informed using Figure 6.7. This graph extends Figure 6.2 to include the value of an OIR decision in addition to the value of reject (R) and approve (A).

[75] From a charity perspective allocating limited funds to research as opposed to direct provision of services will come with health opportunity costs. This is not included in the current analysis as a health system perspective is taken.

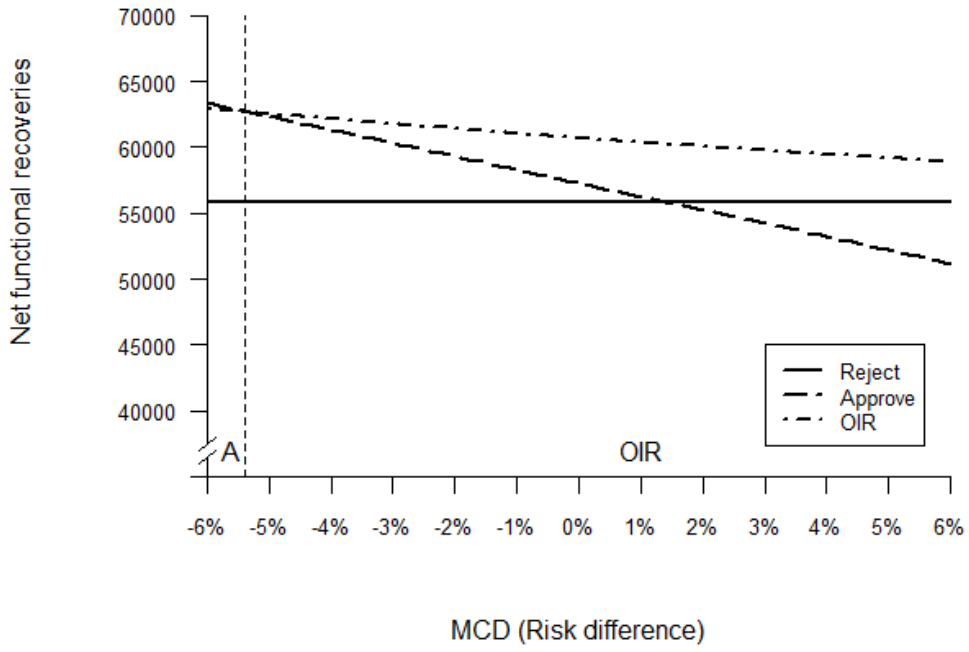


Figure 6.7: Net functional recoveries for different decision options for traumatic brain injury when research takes 5 years to report. Net functional recoveries are expressed at a population level for current and future patients whose treatment choice is informed by the decision. Approve provides more functional recoveries if the MCD is below -5.4%. For all other values of MCD OIR is superior. MCD, minimal clinical difference.

Figure 6.7 illustrates the absolute benefits of approve, reject and OIR guidance for a range of MCD values. The payoff from OIR is calculated as the number of functional recoveries expected from reject in addition to the number of functional recoveries from research. As the benefits of research depend on MCD, the benefits of OIR will depend on the MCD. For example, if the MCD is 0% the value of reject is 55,846 functional recoveries (from point 1) and the upper bound for the value of research is 4,861 resulting in a value of OIR of (55,846 + 4,861=) 60,707 functional recoveries. This compares to 57,265 functional recoveries from approval of the technology. This accounts for the gap between OIR and Approve at MCD 0% in Figure 6.7.

The optimal policy is given by the outer envelope in Figure 6.7[76]. There are two potential policies across the MCD space; Approve (A) and OIR. From the figure, the

[76] This figure is analogous to Figure 4 from (Rothery et al., 2017) except in Rothery et al., the x-axis directly represents changes in price relative to a comparator technology whereas here the x-axis

MCD must be very negative (less than -5.4%) before immediate approval is appropriate. This would require the new treatment to have a very low relative price and/or evidence of substantially reduced side effects. If this is the case then rapid access to the technology provides greater benefits than carrying out research through OIR (which requires a delay in providing access to the new treatment). If the MCD is above -5.4%, then OIR provides a greater number of functional recoveries. For example, at a MCD of 0% OIR provides a greater total number of functional recoveries (3,442) than Approve. This overall difference results from a trade-off between the benefits of early access and the benefits of delaying for additional research. During the 5 years of research Approve provides 556 more functional recoveries than OIR as the new (more effective) treatment is implemented faster. After research reports, the information from OIR provides an additional 3,998 functional recoveries compared to Approve. Overall this results in OIR providing an additional (3,998 – 556 =) 3,442 functional recoveries. This trade-off highlights that research which reports faster will be more valuable as the delay to implement the new treatment will be shorter.

An implication of this approach is that in some cases only the sign of the MCD is required for decision making. In the present case study MCD must be negative for early approval to be appropriate. If it is expected that for a given treatment any reasonable MCD will be positive, due to high relative prices and/or uncertain side effects, then OIR will be appropriate. Precise knowledge about the magnitude of the MCD is not required for decision making only its sign. This is likely to be a common scenario faced by the cancer drugs fund (Grieve et al., 2016) in the United Kingdom and the accelerated approval program in the United States (Gyawali and Kesselheim, 2018).

6.3.7.2 *Research costs funded by the public sector*

If research costs are borne in part or in full by the health system, then decision making must take account of the health opportunity costs of research funding costs. In point 6 we showed how to inform decision making when research is publically funded and AWR is an option. In this case approval and research decisions were independent. Here we illustrate the decision problem when AWR is not an option.

represents changes in MCD relative to a comparator technology. As discussed in point 1 the MCD concept includes relative price in addition to other factors such as side effects.

As research can only be carried out through OIR, it becomes necessary to compare the expected gains from research (OIR) to the gains from immediate access. The resulting decision options and upper bound for the value of additional research for a range of MCD values is illustrated in Figure 6.8.

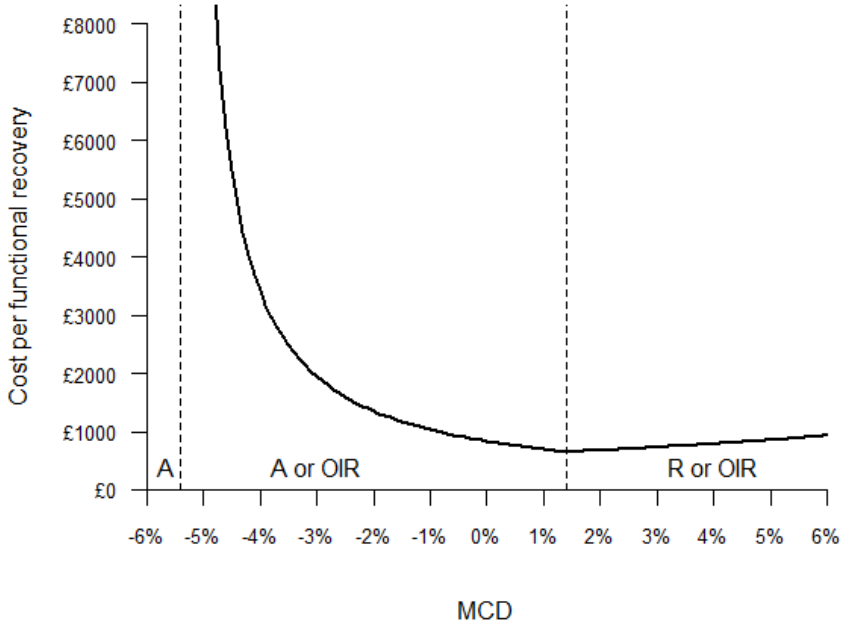


Figure 6.8: Decision options and upper bound for the value of additional research on traumatic brain injury when research takes 5 years to report. Approve is optimal below a MCD of -5.4%. Between -5.4% and 1.4% either approve or OIR is optimal. OIR should be chosen over approve if the cost per functional recovery from research is considered worthwhile at a plausible MCD. Similarly, above 1.4% either reject or OIR is optimal. OIR should be chosen over reject if the cost per functional recovery from research is considered worthwhile. MCD, minimal clinical difference. A, approve. R, reject. OIR, only in research.

In Figure 6.8 the underlying absolute benefits of approve, reject and OIR are the same as those in Figure 6.7. The difference is that the figure above shows the value of choosing research at a given cost to the health system (£2,854,000). The value of research (through OIR) depends on the MCD and is expressed as cost per functional recovery gained. The figure is divided into three sections (from low to high MCD) i) Approve, ii) Approve or OIR, iii) Reject or OIR. These are discussed in turn.

Approve: below MCD -5.4%

If the MCD is expected to be below -5.4% then immediate approval provides more functional recoveries than research through OIR (see Figure 6.7). This is because at

this MCD (low price) there is very little decision uncertainty and so delaying access in order to carry out more research is not worthwhile. Therefore below a MCD of -5.4%, regardless of the cost of research, OIR offers no additional benefit and so research is not worthwhile at any cost per functional recovery. As shown in Figure 6.8 the optimal guidance in this case is to Approve without further research.

Approve or OIR: between MCD -5.4% and +1.4%

For all MCD values above -5.4% OIR is expected to provide a greater number of functional recoveries than the other decision options (Approve or Reject) as can be seen in Figure 6.7. The value of research is calculated by dividing the cost of research by the number of additional functional recoveries through OIR compared to the next best alternative. For example, at an MCD of 0% the benefits of OIR are 60,707 functional recoveries, the benefits of Approve are 57,265 and the benefits of Reject 55,846. This implies that at this MCD (price) research results in (60,707 - 57,265 =) additional 3,442 functional recoveries and so has a value of (£2,854,000/3,442 =) £829 per additional functional recovery[77]. Whether OIR is considered good value depends on how much the health system is willing to pay for an additional functional recovery. Using the heuristic to translate functional recoveries to QALYs, the value of research is approximately £829/15.86 = £52 per QALY. As this compares favourably with the marginal productivity of the general health system; £15,000/QALY (Claxton et al., 2015b) and the other research proposals considered in Chapter 2, OIR appears to be appropriate guidance if AWR is not possible. Reducing price (and therefore MCD) will reduce the additional benefits of OIR and so will make Approve more likely. This can be seen in Figure 6.8 as the cost per additional functional recovery from OIR increases as MCD decreases.

An alternative to starting with a given MCD (price) and calculating whether research is worthwhile at this value, it can be useful to start with the maximum the health system is willing to pay for an additional functional recovery and use this to

[77] If costs were shared with the private sector these costs could simply be subtracted from the costs of research, for example if £1,000,000 was provided by manufacturers the value of research would be (£1,854,000/3,442 =) £539 per additional functional recovery.

determine the required MCD for immediate approval. For example if the health system was willing to pay a maximum of £2,000 per functional recovery, this implies that the new treatment requires a MCD below -3% for immediate approval.

Reject or OIR: above MCD +1.4%

Above an MCD of 1.4% OIR is expected to provide a greater number of functional recoveries than the second best option, which in this case is Reject. As above the decision about whether to carry out research or reject the treatment depends on how much the health system is willing to pay for an additional functional recovery.

6.4 DISCUSSION

This chapter contributes to the literature in two ways. First, we show that the assessments required for research and early approval decisions can be informed without a full economic model. Second we show how early access and research decisions depend on whether or not research costs are borne by the public sector.

As discussed in the introduction, full economic models are expensive and time consuming to construct and this can limit their use by capacity constrained decision making bodies. The method described here does not require a large amount of specialist expertise or time to carry out. This provides three opportunities for supporting evidence based decision making.

First, the method can be routinely carried out as part of reporting the results of a systematic review. This can allow bodies such as Cochrane to provide decision makers with useful metrics to understand the trade-offs involved when deciding on the need for further research.

Second, this framework utilising MCD can help decision makers to inform research and reimbursement decisions. It provides a coherent and transparent basis to trade-off price, uncertainty and effect size when making early access decisions and so can provide a basis to link evidence to pharmaceutical pricing.

Related to the previous point, this method can be used to evaluate the policies which have been implemented in response to calls for earlier access to new technologies such as the cancer drugs fund and the accelerated access collaborative (Grieve et al., 2016; NICE, 2016) in the United Kingdom and the FDA accelerated approval program in the United States (Gyawali and Kesselheim, 2018; Johnson et al., 2011). The methods here can be used to rapidly analyse the health consequences of the decisions taken by these bodies to help determine whether these policies have been worthwhile and whether they should continue in their current form.

Chapter 7

The overall aim of this thesis is to further the use of VOI in research prioritisation by developing methods to calculate VOI which are feasible within the practical constraints of decision making. Throughout the thesis it has been argued that VOI methods can improve the transparency and accountability of decision making by bringing clarity to the discussion around the health benefits of further research and by making use of the available evidence. Despite the benefits of VOI, Chapter 1 identified four barriers to the widespread of VOI methods in practice, namely i) human resources, ii) time, iii) computing resources and iv) familiarity with methods. The rapid VOI methods described in this thesis address the first three of these barriers as they are quick and simple to implement. The online RANE tool greatly reduces the time and technical barriers to carrying out VOI analysis. The final barrier of familiarity has not been surmounted however. This can only be overcome through practical applications and interaction with policy makers as will be discussed in Section 7.1.

The rapid VOI approach necessarily involves simplifying complex clinical processes. Though there may be objections to this, the place of simplified models in decision making may be illustrated with a thought experiment. Imagine comparing VOI results from rapid models with full economic models. For a given set of research proposals, this exercise could compare: i) estimates of the value of research for each proposal and/or ii) the set of projects that would be funded given a certain research budget. Though interesting and potentially useful, this exercise does not include the policy relevant comparator. The rapid VOI methods described in this thesis are not being suggested as an *alternative* to full economic modelling, rather they are being suggested as a *complement* to current decision making processes. Including VOI metrics into this process through rapid VOI has the potential to increase the transparency and accountability of this process. Furthermore, familiarity with VOI can also potentially aid decision making by providing a consistent framework with which to think about uncertainty, opportunity costs, price and the role of research in the health system.

This thesis has developed methods for the analysis of binary, continuous and survival primary outcomes. The analysis for each type of primary outcome is illustrated using a case study from the original six NETSCC proposals. For each type of primary outcome more complex and potentially more realistic models could be used to link changes in the primary outcome to costs and QALYs. As emphasised throughout this thesis, the downside to using more complex models is the demand on time and analyst resources which makes building complex models impractical in many cases. In this spirit, the proper place of this thesis is to contribute to the development of a "toolkit" of practical methods. As discussed in Chapter 1 and Chapter 5 there exist methods for calculating VOI other than those discussed in this thesis which are less resource intensive than building a full economic model. These are the "minimal modelling" and "hybrid minimal modelling" methods described by Meltzer et al., (2011), Basu et al., (2018a), Bennette et al., (2016) and Carlson et al., (2018). Each approach, including the methods developed in this thesis, has different requirements and limitations, therefore a pragmatic approach is to see each method (including full modelling) as an option in the analyst's toolbox. It then becomes a matter of judgement to decide which method is most appropriate in a given context.

An additional benefit of the rapid VOI methods described in this thesis is that they provide a clear interpretation of MCD for a range of outcomes. MCD is defined here as the improvement in the primary outcome that would need to be detected for the new treatment to be considered worthwhile, including any additional costs or side effects. It is worth noting that there are many definitions of MCD in the medical statistics literature and in policy circles more generally (Beaton et al., 2002; Guyatt et al., 2008; Jaeschke et al., 1989). Importantly the definition of MCD used in this thesis includes both the direct health effect on the recipient of the treatment *and* the indirect health effects which result from the opportunity costs of health expenditure. This extension of the MCD concept is an important contribution as it can be used to introduce resource constraints into deliberations about "required effect sizes" which are used in designing non-inferiority trials and in traditional power calculations (Jones et al., 2003).

7.1 IMPLEMENTING RAPID VOI METHODS IN PRACTICE

In practical policy making, research proposals are sketched out initially and are developed into more detailed proposals as they proceed through prioritisation stages. The earlier in the process that quantitative methods are used, the smaller is the chance of inappropriately discarding a potentially high impact research proposal. To generate VOI metrics to inform decision making, the required inputs could be requested as part of the research application process and VOI analysis carried out internally by the research funding body (e.g. NETSCC). Alternatively, VOI analysis could be incorporated into the application process and carried out by research funding applicants.

The RANE tool (https://shiny.york.ac.uk/rane/) means that VOI can be more easily incorporated into the research funding process. However the time and resource constraints of research applicants and decision makers must still be considered. Therefore there are three important questions which must be answered when applying these methods in practice: (i) which type of analysis is required (cost per primary outcome or cost per QALY)? (ii) who should carry out the analysis (research funders or research applicants)? and (iii) how early in the process should VOI methods be used?

Cost per primary outcome

As discussed in Chapter 2, the set of inputs required to estimate the cost per clinical outcome (e.g. £652 per additional functional recovery) for a given research proposal is the minimum amount of information required to understand the health impact of research. Therefore it may be appropriate that this information should be required at the earliest stage possible. Research applicants are best placed to source these inputs as they have access to topic experts and should have an understanding of the literature and natural history of the disease. There may be concerns from research funders that quantitative research assessments create opportunities to selectively thesis or "game" the analysis if carried out by applicants. It should be noted, however, that these opportunities exist within the current narrative approach to the prioritisation of research proposals. Indeed, the requirement for a standardised set of

explicit inputs, which must be supported by reference to evidence, has the potential to decrease selective reporting and increase transparency and accountability.

Research applicants can also benefit from VOI analysis by using it early in the research development process to determine whether a research idea is worth developing into a full grant application. A large amount of time is spent writing research grant applications with a success rate of approximately 20% for NIHR HTA funding in the UK (NIHR, 2018b). This is an important issue as the time spent writing failed applications could have been used in directly productive research activities. The RANE tool can be used early in application development to estimate the health impact of a given research proposal. If the proposal appears likely to be good value for money, then the VOI analysis can be included into the research proposal to increase the chances of receiving funding. If the planned research proposal appears to offer poor value to the health system then the proposal can either be modified or the organisation can save time and resources by switching to an alternative research proposal.

For decision makers, presenting results in terms of cost per additional primary outcome may help to compare value across proposals by making the trade-offs between the different outcomes clearer, therefore it represents a clear improvement relative to implicit forms of decision making. However, significant implicit scientific judgements are still required to make decisions. Therefore, transparency and accountability to evidence can be improved by linking primary outcomes to QALYs.

Cost per QALY

Estimating the health impact of research in terms of cost per QALY (e.g. £44 per additional QALY) provides important advantages over quantifying the value of research in terms of cost per primary outcome (e.g. £652 per functional recovery) as in Chapter 2. Ideally, cost per QALY analysis would be required for all proposals competing for limited research funding since this make the identification of "best buys" more explicit. However, the additional inputs required for cost per QALY analysis require additional analyst time and expertise. If analyst resources are scarce then projects which are (i) more expensive and/or (ii) of uncertain value may be higher priority for cost per QALY analysis. Analyst time could be allocated preferentially to expensive research proposals as these are the most consequential for

the research budget. Intuitively, if an expensive low value project is funded this will consume a large portion of the available research budget which cannot be used on other, potentially more worthwhile, projects. Analyst time could also be focused on projects whose value is not clear based on the cost per natural outcome analysis. A project which appears to be clearly of superior value from the cost per natural outcome analysis may not require a full cost per QALY analysis. However, caution is required as it may be difficult to compare across research proposals by only considering their cost per natural outcome. For example, it may be difficult to know that P1 is a clear research priority by only looking at Table 2.3. If staged or pilot implementation of VOI methods for research prioritisation is considered, then large expensive projects may be prioritised as these are the most consequential for the research budget and population health.

7.1.1 DELIBERATION, DECISION MAKING AND STRUCTURAL UNCERTAINTY

Whether the analysis carried out reports results in terms of cost per primary outcome or cost per QALY, deliberation and judgement will always be required for reasonable social decision making (Daniels, 2000; Rawlins, 2005). Deliberation is required because both the scientific and social value judgements embedded in an analysis can always be reasonably disputed (Claxton et al., 2013).

The need for social value judgments in decision making has been discussed in the context of rare diseases highlighted by the P3 case study. Though there may be little impact in terms of population health in funding further research in rare diseases there may exist social value concerns that are not exhausted by impact on total population health (Hughes et al., 2005; McCabe et al., 2006). No analysis can fully capture these concerns, all that can be aimed for is that the analysis produced can help inform reasoned consideration of the issues.

Scientific judgements about appropriate model structure and use of evidence are also required in any analysis. In some cases sensitivity analysis such as that presented in Section 3.3.1 can be carried out in which the value of particular inputs is varied to illustrate how results change under different assumptions. In other cases it is the structure of the model used to generate the results that is disputed. As discussed in Chapter 1, this “structural uncertainty” is inherent to all decision models. However as the rapid VOI models discussed in this thesis are designed to be simple and easy

to implement, there are likely important situations in which they will fail to capture aspects of disease natural history. In these cases more complex (and time consuming) modelling approach may be appropriate.

Deciding on whether more complex modelling is required for a given proposal requires judgement and deliberation. This judgement will depend on a host of contextual factors. These factors include: the available resources for carrying out more complex modelling; the consequences of delaying decisions while modelling is taking place; the size of the population who can potentially benefit from the research; the budget impact of the research proposal; and the degree of complexity required to characterise disease natural history. Though further methodological research may contribute to this discussion, balancing these factors will always remain a matter of judgement and practical policy making.

7.2 LIMITATIONS AND AREAS FOR FURTHER RESEARCH

7.2.1 PILOTING AND GENERALISABILITY

As discussed in Chapter 5 there are a number of possible extensions to the current iteration of the RANE tool. In the spirit of letting policy needs guide methodological development, extensions and modifications of RANE may be best informed by piloting the tool with decision makers. Ideally this piloting of the tool should be one component of a wider effort to apply rapid VOI methods in practice. To be effective this work would require buy-in from a range of stakeholders involved in research prioritisation. It should focus on identifying barriers and implementing solutions to integrate VOI methods into routine policy making. Further methodological development may be part of this process but it is likely that efforts to understand and adjust institutional structures will be far more important than methodology.

Relatedly, Section **Error! Reference source not found.** assesses the generalisability o f the rapid VOI methods (and the RANE tool) by determining if it could be applied to a new set of NETSCC proposals. It was found that though there are is room for further research the tool is sufficient to provide useful analysis for each of the proposals in the new supporting its claim to generalisability. However this is based on only three proposals, more intensive piloting of the tool is required to understand how generalizable the tool is and to guide further development where necessary.

7.2.2 INFORMING A JUDGEMENT ABOUT RELATIVE EFFECTIVENESS IN THE ABSENCE OF EXISTING EVIDENCE

Chapter 2 highlighted the importance of methods to explicitly quantify uncertainty in situations in which suitable pre-existing studies do not exist. This is fundamental issue in the use of VOI to aid research prioritisation. Despite its importance this issue has received very little attention in the VOI literature. To address this, Section 2.4.4 outlined three available options for decision makers; expert elicitation, statistical modelling and the use of meta-epidemiological studies. Section 2.4.4 also demonstrated the use of a meta-epidemiological approach related that used by Bennette and Carlson. Though providing a useful starting point for deliberation, this method could be greatly improved with further research.

The primary limitation of the approach used is that applying a generic meta-epidemiological estimate to inform uncertainty across all research proposals does not take account of contextual differences between proposals. There may be good scientific reasons to treat proposals differently. This contextual information could be incorporated by (i) integrating expert elicitation and meta-epidemiological methods, (ii) utilising more sophisticated statistical methods to reflect the fact that different disease areas and types of outcome are associated with different distributions of effect sizes (iii) combining approaches i and ii. Developing these methods with the aim of supporting research prioritisation is potentially an important and fruitful area of further research.

Rapid methods of expert elicitation may be useful to develop and integrate into the research prioritisation process. As judgements about uncertainty must be made either implicitly or explicitly, expert elicitation may be especially useful in cases in which previous randomised trials have not been carried out and the use of meta-epidemiological evidence is deemed inappropriate. The importance of expert elicitation in decision making has been increasingly recognised in recent years and number of user friendly tools have been developed to facilitate its integration into decision making (Mason et al., 2017; O'Hagan and Oakley, 2018).

7.2.3 IRRECOVERABLE COSTS AND ONLY IN RESEARCH (OIR) DECISIONS

As in most applied VOI analysis (e.g. Bennette et al., (2016) and Carlson et al., (2018)), the methods illustrated in this thesis make two simplifying assumptions

about decision making. First, it is assumed that there are no irrecoverable costs associated with the treatments considered. Irrecoverable costs are those costs that, once committed, cannot be recovered if guidance is changed at a later date (Eckermann and Willan, 2007; McKenna et al., 2015; Thijssen and Bregantini, 2017). For example, if large capital investments are required for a treatment to be delivered in a health system then this is associated with irrecoverable costs if the health system cannot easily sell the capital when guidance changes. The second assumption is that decisions to recommend an intervention for widespread use and decisions to carry out additional research on that same intervention are independent. This can be an issues as there are cases in which research cannot be carried out at the same time as the treatment is available for widespread use. In this case a conditional coverage option such as "Only in Research" (OIR) may be appropriate. This means that the use of the new treatment is only allowed in a research setting. Because an OIR decision restricts access to a new intervention this is associated with health opportunity costs of delaying access and so the value of research is affected. A framework exists to relax these assumptions and integrating this perspective with rapid VOI methods is an important area for further research (Claxton et al., 2012; McKenna et al., 2015; Rothery et al., 2017).

7.2.4 MARGINAL PRODUCTIVITY OF RESEARCH BUDGET

By linking primary outcomes to QALYs Chapter 3 constructs a "bookshelf" of research proposals which are ranked from highest to lowest health impact. This framework provides a consistent approach to estimate the marginal productivity of the research budget. This could be estimated by calculating the value of research of proposals which just missed out on funding due to resource constraints (Culyer, 2016). These are the marginal projects which would have been funded if the research budget was larger and so give an indication of the gains from expanding the research budget. In principal, the budget for direct service provision (e.g. the NHS) and the health research budget (e.g. NETSCC), compete for funding. Therefore estimates of their respective marginal productivities can be used to decide whether more resources should be devoted to research rather than to direct patient care (or vice versa). This line of inquiry fits within a wider approach to decision making which aims to reveal the benefits and opportunity costs of public sector expenditure across

multiple sectors of the economy (Claxton et al., 2007; Remme et al., 2017; Sanders et al., 2016).

7.2.5 RESEARCH DESIGN AND INDIVIDUALISED CARE

The methods described in this thesis provide an expected upper bound on the value of research when all uncertainty is resolved. As research designs have a finite sample size, they will only ever partially resolve the total uncertainty. Methods to adjust research value for sample size and other aspects of research design are well developed and further work is required to incorporate this into the rapid VOI methods.

Chapter 4 developed and applied a novel method for estimating the VOI provided by a feasibility/pilot studies. This extension allows the expected health impact of proposals for feasibility/pilot studies to be compared directly to proposals for RCTs or other comparative effectiveness research which is essential if funding for feasibility/pilot studies and full trials come from the same research budget.

In addition to feasibility/pilot studies there are a number of other non-standard research designs. These include sequential RCTs (Wang et al., 2012), trials within cohorts (Relton et al., 2010) and stepped wedge designs (Hemming et al., 2015) among others. The P4 case study requested £3,310,883 to begin a complex adaptive trial. This is a trial in which modifications can be made to a trial's design as it is ongoing. Outcomes are monitored over time and depending on how well or poorly patients perform on treatment arms, patients are preferentially enrolled to particular arms, or arms may be dropped, and/or entirely new treatments added (Chow and Chang, 2008). In principle, it is possible to adapt VOI methods to address this trial design. However, due to the complexity of this design, it is challenging to rapidly quantify the benefits of the adaptive design and so P4 was treated as a standard RCT design in Section 4.2.2. Other innovative research designs are also likely to emerge from the increased attention paid to estimating the effects of interventions on individual patients rather than patient populations i.e. individualised care (Basu et al., 2016; Espinoza et al., 2014; Love-Koh et al., 2018). These non-standard research designs must be funded from the same research budget as feasibility studies and RCTs. VOI methods provide a framework to compare the additional value added from these innovative designs to any additional costs, though additional research is

required to develop methods to quantify the benefits of research in this rapidly developing area.

Appendix

INPUTS FOR PROPOSALS 1-6

A1. PROPOSAL 1

Inputs used to estimate the value of research in terms of natural outcomes for P1

Input	Value	Reference / Justification
Type of outcome	Binary	Functional recovery.
Outcome benefit or harm?	Benefit	
Baseline probability	55% 162 events out of 294 at risk.	Not reported in the proposal. Using the placebo group from Nichol et al., (2015), which reported 162 out of 294 functional recoveries with standard care.
Relative effect of new treatment	Odds ratio with 95% CI from 0.23 to 5.24	Not reported in the proposal. Assumption based on the meta-epidemiological study of 743 publically funded RCTs reported in Djulbegovic et al., (2012), see Section 2.4.4.
Incidence	8,800 per year	Not reported in the proposal. A study by Sauerland and Maegele (2004) estimates that approximately 8,800 patients per year suffer TBI in the UK and it is assumed that all will receive either early or late PTP.
Time information is expected to be valuable	15 years	Not reported in the proposal. Changes to standard practice in this area appear to move relatively slowly. Therefore, it is anticipated that the new information will be valuable for a long time span of 15 years.
Duration of research	5 years	From proposal.
Cost of research to NETSCC	£2,854,000	From proposal.
Discount rate	3.5%	Not reported in the proposal. Guidance from the UK Treasury suggests the use of a discount rate of 3.5% per annum (*HMT Green Book*, 2013).
Current level of utilisation of the	100% receive	Not reported in the proposal. Late PTP is established practice in the UK. Therefore, it is assumed that 100%

interventions	late PTP	of patients currently receive late PTP.

Inputs used to estimate the value of research in terms of costs and QALYs for P1

Input	Value	Reference / Justification
Additional cost of Intervention 1: early PTP	£14.10 per person	3 days of additional doses required due to earlier treatment initiation (Medicines Complete).
Probability of being in GOSE state if functional recovery occurs	42%, 24%, 20% and 14% for GOSE states 5-8	Not reported in the proposal. From Nichol et al., (2015).
Life expectancy if functional recovery occurs	16.73, 16.73, 19.23 and 19.23 years for GOSE states 5-8	Not reported in the proposal. From Shavelle et al., (2006).
Health utility score for functional recovery	0.7, 0.81, 0.96 and 1 for GOSE states 5-8	Not reported in the proposal. From Fuller et al., (2017).
Disease related costs for functional recovery	£27,047, £27,047, £19,575 and £19,575 for GOSE states 5-8	Not reported in the proposal. From Nyein et al., (1999) and Wood et al., (1999).
Probability of being in GOSE state if functional recovery does not occur	29%, 7%, 41% and 23% for GOSE states 1-4	Not reported in the proposal. From Nichol et al., (2015).
Life expectancy if functional does not occur	0, 7.11, 12.52 and 12.52 years for GOSE states 1-4	Not reported in the proposal. (Shavelle et al., 2006)
Health utility score if functional recovery does not occur	0, 0.11, 0.41 and 0.58 for GOSE states 1-4	Not reported in the proposal. (Fuller et al., 2017)
Disease related costs if functional recovery does not occur	£0, £45,450, £154,324, £154,324 for GOSE states 1-4	Not reported in the proposal. (Nyein et al., 1999; Wood et al., 1999)
NHS support and treatment costs	£490,000	From proposal.
Opportunity cost of	£15,000 per	Not reported in the proposal. Endorsed by the UK

health care expenditure	QALY	Department of Health for use in health impact assessments (NHS England, 2015).

A2. PROPOSAL 2

Inputs used to estimate the value of research in terms of natural outcomes for P2

Input	Value	Reference / Justification
Type of outcome	Survival	Time spent in pre-progression state i.e. progression free survival (PFS).
Outcome benefit or harm?	Benefit	Pre-progression is a better state than either post-progression or death.
Baseline survival distribution	Exponential	Not reported in the proposal. Assume exponential distribution.
Rate of transition per month (Lambda)	0.028	Not reported in the proposal. See Section 4.2.3.2
Relative effect of new treatments	Treat for 12 months: HR from 1.05 to 4. Treat for 6 months: HR from 1.1 to 4.39	Not reported in the proposal. The analysis of Djulbegovic et al., (2012) does incorporate survival outcomes and so can be used as a starting point in this analysis see Section 2.4.4 . If the primary outcome is a beneficial survival outcome and so this implies a hazard ratio between 0.19 and 4.39. However, the meta-epidemiological analysis currently does not address plausible results for dose reductions studies. The new treatments are very likely to be less effective than current treatment and so will be associated with higher hazards of transitioning out of pre-progression state to either progression or death; this implies a hazard ratio > 1. As it is very unlikely that treatment for 6 months will be superior to standard care we assume a range for the hazard ratio from 1.1 to 4.39. As treatment for 12 months is likely to be more effective than treatment for 6 months we assume a range from 1.05 to 4. Further meta-epidemiological research and/or expert elicitation is required to better characterize uncertainty in this case.
Incidence	1,137	Not reported in the proposal. The NICE budget impact statements for the appraisal of nivolumab and pembrolizumab *(National Institute for Health and Care Excellence*, 2015) estimates that approximately 1,137 individuals per year will meet the criteria required for use of these drugs.
Time information is expected to be valuable	10 years	Not reported in the proposal. Standard practice in this area appears to move at a moderate pace. Therefore, it is anticipated that the new information will be valuable for a medium time span of 10 years.

Duration of study	6 years	From proposal. The primary analysis of 2 year PFS is planned to report after 6 years. There is also an additional long term follow up study which is planned to report after 10.3 years.
Cost of study to NETSCC	£2,522,710	From proposal.
Discount rate	3.5%	Not reported in the proposal. Guidance from the UK Treasury suggests the use of a discount rate of 3.5% per annum (*HMT Green Book*, 2013).
Current level of utilisation of the interventions	100% receive intensive treatment	Not reported in the proposal. Intensive treatment is the NICE mandated treatment in the UK *(National Institute for Health and Care Excellence*, 2015). Therefore, we assume that 100% of patients currently receive intensive treatment.

Inputs used to estimate the value of research in terms of costs and QALYs for P2

Input	Value	Reference / Justification
Monthly cost for each treatment.	£6,042	From proposal.
Utility associated with pre-progression health state?	0.79	Not reported in the proposal. The pre progression utility score of 0.79 is from the NICE guidance *(National Institute for Health and Care Excellence*, 2015) and come from the CHECKMATE-006 trial (Robert et al., 2015).
Expected disease related costs associated with the pre-transition health state:	£100 per month	The costs of the pre progressed state are the same as those used in the NICE guidance: £100 per month (Johnston et al., 2012).
Duration of treatment with baseline:	Treat to progression	From proposal.
Duration of treatment with intervention 1:	12 months	From proposal.
Duration of treatment with intervention 2:	6 months	From proposal.
NHS support and treatment costs of full study	£-62,410,967	From proposal.
Opportunity cost of health care expenditure	£15,000 per QALY	Not reported in the proposal. Endorsed by the UK Department of Health for use in health impact assessments (NHS England, 2015).

A3. PROPOSAL 3

Inputs used to estimate the value of research in terms of natural outcomes for P3

Input	Value	Reference / Justification
Type of outcome	Binary	The primary outcome for the trial is the number of withdrawal attributable serious adverse event (SAE) prevented during a 2 year follow up.
Outcome benefit or harm?	Harm	
Baseline probability	5.5%, 95% CI from 5% to 6%	In the rationale for research section of the proposal 5% is cited, in the scientific abstract 6% is cited. We assume uncertainty between 5% and 6%
Relative effect of new treatment	Odds ratio with 95% CI from 4.39 to 0.19	Assumption based on the meta-epidemiological study of 743 publically funded RCTs reported in Djulbegovic et al., (2012), see Section 2.4.4. Primary outcome is a harmful binary outcome.
Incidence	26.3 individuals per year	From proposal, the number of individuals facing the uncertain treatment choice in the UK is estimated to be 26.3 per year.
Time information is expected to be valuable	10 years	There are a number of trials currently underway in this area. Therefore, it is anticipated that the new information will be valuable for a medium time span of 10 years. Alternative scenarios may be run to assess the impact of shorter durations.
Duration of research	4 years	From proposal.
Cost of research to NETSCC	£855,403	From proposal.
Current utilization	100% using current practice	Not reported in the proposal. Assumption.

Inputs used to estimate the value of research in terms of costs and QALYs for P3

Input	Value	Reference / Justification
Cost of continuous treatment if SAE occurs	£4,877,749	Not reported in the proposal. See Section 4.2.1
Cost of continuous treatment if SAE does not occur	£7,316,623	Not reported in the proposal. See Section 4.2.1
Cost of withdrawal if SAE occurs	2,643,621	Not reported in the proposal. See Section 4.2.1
Cost of withdrawal if SAE does not occur	£3,965,432	Not reported in the proposal. See Section 4.2.1
Probability of being in state if SAE occurs	33.3%, 33.3% and 33.3% for death, infection and organ injury	Not reported in the proposal. Assumption.
Life expectancy if SAE occurs	0, 35.47 and 35.47 years for death, infection and organ injury	Not reported in the proposal. If individuals survive we assume that they will live for their full additional life expectancy according to NICE report *(National Institute for Health and Care Excellence*, 2015), Eculizumab for treating atypical haemolytic uraemic syndrome.
Health utility score if SAE occurs	0, 0.2 and 0.59 for death, infection and organ injury	Not reported in the proposal. Death is associated with utility zero. If irreversible organ injury occurs we assume that utility will be 0.59 which is the average UK utility in diseases of kidney and ureters (Sullivan et al., 2011). If meningococcal infection occurs we assume that individuals will have the lower utility of 0.2 (Christensen et al., 2013).
Disease related costs if SAE occurs	£0, £3,216 and £115,503 for death, infection and organ injury	Not reported in the proposal. Death is associated with zero additional disease related costs. If meningococcal infection occurs patients incur disease related costs from a spell in hospital (£2936.2) and a follow up appointment (£279.98) (Christensen et al., 2013). If irreversible organ injury occurs we assume this will require kidney transplant costing £17,000 in the first year and £5,000 in subsequent years (*NHS England*, 2013). Individuals are assumed to survive for their remaining 35.47 years of life, resulting in a total cost of £115,503 after discounting.
Life expectancy if SAE does not occur	35.47 years	Not reported in the proposal. Full additional life expectancy according to NICE (*National Institute*

		for Health and Care Excellence, 2015), Eculizumab for treating atypical haemolytic uraemic syndrome.
Health utility score if SAE does not occur	0.59	Not reported in the proposal. 0.59 is the average UK utility in diseases of kidney and ureters (Sullivan et al., 2011).
Disease related costs if SAE does not occur	£0	Not reported in the proposal. Assume zero additional disease related costs (excluding treatment costs).
NHS support and treatment costs	£10,608,500	From proposal.
Opportunity cost of health care expenditure	£15,000 per QALY	Not reported in the proposal. Endorsed by the UK Department of Health for use in health impact assessments (NHS England, 2015).

A4. PROPOSAL 4

Inputs used to estimate the value of research in terms of natural outcomes for P4

Input	Value	Reference / Justification
Type of outcome	Continuous	The primary outcome for the trial is a continuous measure of Alzheimer's severity; the mini mental state examination (MMSE).
Outcome benefit or harm?	Benefit	Higher MMSE score indicates improvement.
Relative effect of new treatments	Mean of 0 and SE of 0.54	Not reported in proposal. See Section 2.4.4.
MCD	1.4 MMSE points	The proposal specified a required change in MMSE of 1.4 points for all of the active treatments.
Incidence	100,000	From proposal.
Time information is expected to be valuable	20 years	Not reported in the proposal. Standard practice in this area appears to move slowly. Therefore, it is anticipated that the new information will be valuable for a long time span of 20 years.
Duration of study	6 years	From proposal.
Cost of study to NETSCC	£3,310,883	From proposal.
Discount rate	3.5%	Not reported in the proposal. Guidance from the UK Treasury suggests the use of a discount rate of 3.5% per annum (*HMT Green Book*, 2013).
Current level of utilisation of the interventions	100% receive no treatment.	Not reported in the proposal. There is no treatment affecting disease progression, therefore, we assume that 100% of patients currently receive no disease modifying treatment (placebo).

Inputs used to estimate the value of research in terms of costs and QALYs for P4

Input	Value	Reference / Justification
Monthly cost of baseline: no treatment	£0	No treatment does not incur any additional costs.
Monthly cost of treatment 1: Exenatide	£73.36	Not reported in the proposal. From (Medicines Complete, 2018) , Exenatide costs £18.34 per week x 4 weeks = 73.36 per month.
Monthly cost of treatment 2: Telmisartan	£14.83	Not reported in the proposal. From (Medicines Complete, 2018), Telmisartan costs £13.61 per 28 tablet pack. £13.61 /28 = £0.49 per day x 30.5 days in a month = £14.83 per month.
Monthly cost of treatment 3: Combination	£88.19	Not reported in the proposal. Combination of Exenatide and Telmisartan (above) costs £73.36 + £14.83 = £88.19 per month.
How is the primary outcome expected to be related to health utility?	0.01242 increase in utility for 1 unit increase in MMSE	Not reported in the proposal. From Technical Appraisal of Alzheimer's drugs (Bond et al., 2012) see Section 4.2.2.6.
How is the primary outcome expected to be related to disease related costs?	£14 decrease in costs for 1 unit increase in MMSE	Not reported in the proposal. Wolstenholme et al., (2002) estimated the relationship between disease progression and cost of care in dementia. Each one-point decline in the MMSE score is associated with a £56 increase in the four-monthly costs. £56/4 = £14.
How long is the treatment effect expected to last	12 months	Not reported in proposal. Assumption, should be informed by expert opinion.
NHS support and treatment costs of research	£1,297,789	From proposal.
Opportunity cost of health care expenditure	£15,000 per QALY	Not reported in the proposal. Endorsed by the UK Department of Health for use in health impact assessments (NHS England, 2015).

A5. PROPOSAL 5

Inputs used to estimate the value of research in terms of natural outcomes for P5

Input	Value	Reference / Justification
Type of outcome	Binary	The proposal discusses the Positive and Negative Syndrome Scale (PANSS) score as an outcome in the full trial data could not be found to link to costs and QALYs. As relapse is a major concern in first episode psychosis we assume that the definitive trial will use relapses prevented as the primary outcome (Craig et al., 2004).
Outcome benefit or harm?	Harm	
Baseline probability	47.5% 29 events out of 61 at risk.	Not reported in the proposal. Using the treatment as usual arm of a UK (London) based, RCT by Craig et al., (2004) we estimate that there is a (29/61 =) 47.5% risk of relapse with treatment as usual for young people (aged 16-40) with first episode psychosis.
Relative effect of new treatments	Odds ratio with 95% CI from 0.19 to 4.39	Not reported in the proposal. Assumption based on the meta-epidemiological study of 743 publically funded RCTs reported in Djulbegovic et al., (2012), see Section 2.4.4.
Incidence	1,563 per year	Not reported in the proposal. From Kirkbride et al., (2013), 5,939 cases for individuals between 16-35 years old. Assuming constant rate of events over all age ranges: (5939/19 =) 312.58 cases for each year group. Number of individuals between the ages of 14-18 (312.58 x 5 =) 1,563 each year.
Time information is expected to be valuable	15 years	Not reported in the proposal. Standard practice in this area appears to move relatively slowly. Therefore, it is anticipated that the new information will be valuable for a long time span of 15 years.
Duration of feasibility study	2 years	From proposal.
Duration of full trial (including time to apply for funding)	6 years	Not reported in the proposal. From Morgan et al., (2018) full trials funded by the National Institute for Health Research (NIHR) Research for patient benefit (RfPB) programme which arise out of feasibility studies take 42 months on average to report (range 26 to 55). Application for a full trial takes an average of 10 months (range 7 – 29) after feasibility trial reporting and trials start an average of 18 months (range 13 to 28) after application. Therefore the

		average time from feasibility study report to full trial report is approximately (42 +10+18)/12 = 6 years.
Cost of feasibility study to NETSCC	£601,481	From proposal
Discount rate	3.5%	Not reported in the proposal. Guidance from the UK Treasury suggests the use of a discount rate of 3.5% per annum (*HMT Green Book*, 2013).
Cost of full trial to NETSCC	£1,000,000	Not reported in the proposal. From Morgan et al., (2018) full trials funded by the NIHR RfPB programme cost £1,163,996 on average (range £321,403 to £2,099,813). If the full trial is feasible, it will take place it is expected to take place 10 months after the feasibility trial reports i.e. (24 + 10)/12 = 2.8 years. Discounting to present value the expected cost of the full trial is £1,163,996/(1.035)^2.8 = £1,057,106.
Current level of utilisation of the interventions	100% receive APs.	Not reported in the proposal. The proposal states that standard treatment is APs, which are chosen based on *(National Institute for Health and Care Excellence*, 2014) NICE guidance. Therefore, we assume that all individuals receive this treatment.
Likelihood of delivering a future full trial:	64%	Not reported in the proposal. From Morgan et al., (2018) (57/89 =) 64% of feasibility studies were considered feasible by self-report of principal investigators. 20 were judged as not feasible and 12 as uncertain.

Inputs used to estimate the value of research in terms of costs and QALYs for P5

Input	Value	Reference / Justification
Patient time horizon considered	1 year	Not reported in the proposal. Here we only consider differences in costs and outcomes over a 12 month period since there is no information reported on long term effects of treatment. This is an important limitation of the analysis (Briggs et al., 2006) since long term effects on costs and outcomes are not captured.
Cost of baseline treatment: Antipsychotics	£687	Not reported in the proposal. NICE guidance *(National Institute for Health and Care Excellence*, 2014) recommends that the choice of an AP for a particular individual should be based on its side effect profile for that individual. Since there is no single AP used as established standard of care in the UK for CYP, we use the average yearly cost over the range of APs listed in the NICE guidance of £687 per individual. Monthly costs from *National Institute for Health and Care Excellence* (2014) guidance: £57.23 (Amisulpride), £14.35 (Haloperidol), £85.13 (Olanzapine), £108.89 (Aripiprazole) £156.34 (Paliperidone), £67.52 (Risperidone), £63.03 (Zotepine), £6.7 (Flupentixoldecanoate). Average cost = £57.23 per month. £57.23 x 12 months = £687 per year.
Cost of treatment 1: Psychological interventions	£3,492	Not reported in the proposal. The psychological intervention (PI) is described in the proposal as up to 30 sessions of cognitive behavioral therapy (CBT) over 6 months, with an extra 6 sessions of family intervention with parents/carers. Based on a cost of £97 per CBT session (Curtis and Burns, 2016) this implies a total cost of (£97 x 36 =) £3,492 per individual for PI.
Cost of treatment 2: Combination	£4,179	Not reported in the proposal. The combination therapy (AP + PI) is assumed to be simply the cost of APs plus the cost of PI, (£3,492 + £687 =) £4,179 per individual.
Health state utility if relapse does not occur	0.94	Not reported in the proposal. If relapse does not occur we assume a similar health utility as general population below 30 years old (Ara and Brazier, 2011).
Disease related costs if relapse does not occur	£5,401	Not reported in the proposal. Costs from NICE *(National Institute for Health and Care Excellence*, 2014) guidance on adult schizophrenia, using only outpatient, primary and community care costs = £5,401 per year.
Health state utility if relapse occurs	0.805	Not reported in the proposal. Utility during relapse estimated from adults with schizophrenia (0.67) (Lenert et al., 2004). From NICE *(National Institute for Health and Care Excellence*, 2014) guidance on adult schizophrenia relapse is associated with a 6 month decrement in utility. Average utility over 12 month period = 0.67 x 0.5 + 0.94 x 0.5 = 0.805

Disease related costs if relapse occurs	£19,210	Not reported in the proposal. From NICE *(National Institute for Health and Care Excellence*, 2014) guidance on adult schizophrenia disease related costs are £33,018 per year for relapse. As above assume 6 months of increased costs associated with relapse and 6 months of non-relapsed costs = £33,018 x 0.5 + £5,401 x 0.5 = £19,210
NHS support and treatment costs for feasibility study	£150,000	From proposal.
NHS support and treatment costs of full study	£490,000	Not reported in the proposal. Because of similar size and duration of planed research, additional NHS support and treatment costs are assumed to be close to those of P1, which is £490,000.
Opportunity cost of health care expenditure	£15,000 per QALY	Not reported in the proposal. Endorsed by the UK Department of Health for use in health impact assessments (NHS England, 2015).

A6. PROPOSAL 6

Inputs used to estimate the value of research in terms of natural outcomes for P6

Input	Value	Reference / Justification
Type of outcome	Binary	Death in preferred place.
Outcome benefit or harm?	Benefit	Dying in a preferred place is considered better to dying in a place not of the patients choosing.
Baseline probability	30%, 95% CI from 20% to 40%	From proposal, of those who wish to die at home approximately 30% achieve their wish (*Office for National Statistics*, 2015). Assume +/- 10% uncertainty.
Relative effect of new treatment	Odds ratio with 95% CI from 1.1 to 5.24	Not reported in the proposal. The analysis of Djulbegovic et al., (2012) can be used as a starting point in this analysis see Section 2.4.4. The primary outcome is a beneficial binary outcome and so this implies a hazard ratio between 0.23 and 5.24. However, the new treatment is very likely to be more effective than current treatment and so will be associated with odds achieving preference for place of death, this implies odds ratio > 1. Therefore we assume a range for the odds ratio from 1.1 to 5.24. Further meta-epidemiological research and/or expert elicitation is required to better characterize uncertainty in this case.
Incidence	259,150 per year	The Palliative Care Funding Review (Tom Hughes-Hallett et al., 2011) estimates that 355,000 patients need palliative care each year and 73% of these individuals would prefer to die at home (*Macmillan*, 2010). Therefore the population of interest is approximately (355,000 x 0.73 =) 259,150 individuals per annum.
Time information is expected to be valuable	15 years	Not reported in the proposal. Changes to standard practice in this area appear to move relatively slowly. Therefore, it is anticipated that the new information will be valuable for a long time span of 15 years.
Duration of research	3 years	From proposal.
Cost of research to NETSCC	£882,177	From proposal.
Discount rate	3.5%	Not reported in the proposal. Guidance from the UK Treasury suggests the use of a discount rate of 3.5% per annum (*HMT Green Book*, 2013).
Current level of	100%	Not reported in the proposal. We assume that 0% of

utilisation of the interventions	receive no treatment	patients currently receive the booklet.

Inputs used to estimate the value of research in terms of costs and QALYs for P6

Input	Value	Reference / Justification
Additional cost of Intervention 1: booklet	£9.06 per person	Assuming the estimated NHS support and treatment costs of £4,104 are to cover the cost of the booklet and advice in the treatment arm (n=453), this will cost the NHS an additional (£4,104/453 =) £9.06 per person.
Life expectancy if death at home does not occur (death in hospital)	0.3 years	Those included in the trial have a prognosis of 8-24 weeks. The midpoint in this range is 16 weeks which is equivalent to 16/52 = 0.3 years.
Health utility score if death at home does not occur (death in hospital)	0.76	Not reported in the proposal. Estimated utility of palliative care patients a review of utilities at the end of life compiled by Dixon et al., (2009). The average utility in this sample (extracted from column 7 of Table 2 of report) was 0.76. It is assumed that the health state utility is not affected by whether the patient dies at home or at hospital.
Disease related costs if death at home does not occur (death in hospital)	£3,000	Not reported in the proposal. A review carried out by the End of Life Care Programme (2012) reported the average cost of a hospital stay ending in death to be approximately £3,000.
Life expectancy if death at home does occur	0.3 years	Those included in the trial have a prognosis of 8-24 weeks. The midpoint in this range is 16 weeks which is equivalent to 16/52 = 0.3 years.
Health utility score if death at home does occur	0.76	Not reported in the proposal. Estimated utility of palliative care patients a review of utilities at the end of life compiled by (Dixon et al., 2009). The average utility in this sample (extracted from column 7 of Table 2 of report) was 0.76. It is assumed that the health state utility is not affected by whether the patient dies at home or at hospital.
Disease related costs if death at home does occur	£2,108	Not reported in the proposal. A review carried out by the End of Life Care Programme (*Reviewing end of life care costing information to inform the QIPP End of Life Care Workstream*, 2012) report provided a range of £1,415 to £2,800 for the cost of dying in the community. Assume the midpoint of this range £2,108.
NHS support and treatment costs	£4,014	From proposal.
Opportunity cost of health care expenditure	£15,000 per QALY	Not reported in the proposal. Endorsed by the UK Department of Health for use in health impact assessments (NHS England, 2015).

A7. R CODE TO CONVERT BETWEEN ODDS RATIO, RISK DIFFERENCE AND RISK RATIO

Odds ratio (OR) to risk difference (RD)

```
# Events_t0 = events in control arm
# AtRisk_t0 = number at risk in control arm
# OR_t1 = odds ratio for treatment effect
OR_to_RD <- function(Events_t0, AtRisk_t0, OR_t1){
  set.seed(5)
  # RD = E(P_t1) - E(P_t0)
  # E(P_t0) = 162/294 = 0.5510204
  E_P_t0 <- Events_t0/AtRisk_t0
  # therefore
  # RD = E(P_t1) - 0.5510204

  # E(P_t1) = E(Odds_t1/(1 + Odds_t1))
  # Odds_t1 = Odds_t0*OR_t1
  # Odds_t0 <- Prob_t0/(1 - Prob_t0)
  Prob_t0 <- rbeta(100000, Events_t0, (AtRisk_t0 - Events_t0))
  Odds_t0 <- Prob_t0/(1 - Prob_t0)
  Odds_t1 <- Odds_t0*OR_t1
  E_P_t1 <- mean(Odds_t1/(1 + Odds_t1))

  RD_t1 <- E_P_t1 - E_P_t0
  return(RD_t1)
}
```

Risk difference (RD) to odds ratio (OR)

```
# Note: require: OR_to_RD() function
# Events_t0 = events in control arm
# AtRisk_t0 = number at risk in control arm
# RD_t1 = risk difference for treatment effect
# search_range = range of odds ratios tried in maximisation algorithm
RD_to_OR <- function(Events_t0, AtRisk_t0, RD_t1, search_range = c(0, 6)){
  set.seed(5)
  # minimse squared distance between OR_variable and the target RD
  # x = OR_variable
  f2 <- function(x) (OR_to_RD(Events_t0, AtRisk_t0, x) - RD_t1)^2
  # optimise is for single variable minimisation:
  # c(search_range[1], search_range[2])is the area for searching
  r2 <- optimise(f2, c(search_range[1], search_range[2]))    # single variable minimisation: c(0, 2) is the area for searching
  return(r2$minimum)
  # note: estimation correct to 13 decimal places.
}
```

References

Ara, R., Brazier, J.E., 2011. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. Value Health 14, 539–545.

Arrow, K.J., 1963. Uncertainty and the Welfare Economics of Medical Care. Am. Econ. Rev. 53, 941–973.

Bae, J.-M., 2014. Meta-epidemiology. Epidemiol. Health 36.

Barczyk, H., 2018. Opinion | Easier Drug Approval Isn't Cutting Drug Prices. N. Y. Times.

Basu, A., Carlson, J.J., Veenstra, D.L., 2016. A framework for prioritizing research investments in precision medicine. Med. Decis. Mak. Int. J. Soc. Med. Decis. Mak. 36, 567–580. https://doi.org/10.1177/0272989X15610780

Basu, A., Sullivan, S.D., 2017. Toward a Hedonic Value Framework in Health Care. Value Health 20, 261–265.

Basu, A., Veenstra, D.L., Carlson, J.J., Wang, W.-J., Branch, K., Probstfield, J., 2018a. TECHNICAL APPENDIX for the VICTOR Platform.

Basu, A., Veenstra, D.L., Carlson, J.J., Wang, W.-J., Branch, K., Probstfield, J., 2018b. Value of Information for Cardiovascular Trials and Other Comparative Research (VICTOR). Washington.

Beaton, D.E., Boers, M., Wells, G.A., 2002. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. Curr. Opin. Rheumatol. 14, 109.

Bennette, C.S., Veenstra, D.L., Basu, A., Baker, L.H., Ramsey, S.D., Carlson, J.J., 2016a. Development and Evaluation of an Approach to Using Value of Information Analyses for Real-Time Prioritization Decisions Within SWOG, a Large Cancer Clinical Trials Cooperative Group. Med. Decis. Mak. Int. J. Soc. Med. Decis. Mak. 36, 641–651. https://doi.org/10.1177/0272989X16636847

Bill and Melinda Gates foundation, 2018. URL https://www.gatesfoundation.org/ (accessed 10.25.18).

Bindels, J., Ramaekers, B., Ramos, I.C., Mohseninejad, L., Knies, S., Grutters, J., Postma, M., Al, M., Feenstra, T., Joore, M., 2016. Use of value of information in healthcare decision making: exploring multiple perspectives. Pharmacoeconomics 34, 315–322.

BNF, 2018. BNF: British National Formulary - NICE [WWW Document]. URL https://bnf.nice.org.uk/medicinal-forms/dalteparin-sodium.html (accessed 7.26.18).

Bojke, L., Claxton, K., Sculpher, M., Palmer, S., 2009. Characterizing Structural Uncertainty in Decision Analytic Models: A Review and Application of Methods. Value Health 12, 739–749. https://doi.org/10.1111/j.1524-4733.2008.00502.x

Bond, M., Rogers, G., Peters, J., Anderson, R., Hoyle, M., Miners, A., Moxham, T., Davis, S., Thokala, P., Wailoo, A., 2012. The effectiveness and cost-effectiveness of donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease (review of Technology Appraisal No. 111): a systematic review and economic model.

Brennan, A., Chick, S.E., Davies, R., 2006. A taxonomy of model structures for economic evaluation of health technologies. Health Econ. 15, 1295–1310.

Brennan, A., Kharroubi, S.A., 2007. Efficient computation of partial expected value of sample information using Bayesian approximation. J. Health Econ. 26, 122–148. https://doi.org/10.1016/j.jhealeco.2006.06.002

Briggs, A., Claxton, K., Sculpher, M.J., 2006a. Decision Modelling for Health Economic Evaluation. Oxford University Press.

Carlson, J.J., Kim, D.D., Guzauskas, G.F., Bennette, C.S., Veenstra, D.L., Basu, A., Hendrix, N., Hershman, D.L., Baker, L., Ramsey, S.D., 2018. Integrating value of research into NCI Clinical Trials Cooperative Group research review and prioritization: A pilot study. Cancer Med.

Carlson, J.J., Thariani, R., Roth, J., Gralow, J., Henry, N.L., Esmail, L., Deverka, P., Ramsey, S.D., Baker, L., Veenstra, D.L., 2013. Value-of-information analysis within a stakeholder-driven research prioritization process in a US setting: an application in cancer genomics. Med. Decis. Making 33, 463–471.

CG 178 Psychosis and Schizophrenia in adults | Clinical Guidance | National Institute for Health and Care Excellence, 2014.

Chalmers, I., Bracken, M.B., Djulbegovic, B., Garattini, S., Grant, J., Gülmezoglu, A.M., Howells, D.W., Ioannidis, J.P., Oliver, S., 2014. How to increase value and reduce waste when research priorities are set. The Lancet 383, 156–165.

Chalmers, I., Nylenna, M., 2014. A new network to promote evidence-based research. The Lancet 384, 1903–1904. https://doi.org/10.1016/S0140-6736(14)62252-2

Chang, W., Cheng, J., Allaire, J., Xie, Y., McPherson, J., 2017. shiny: Web Application Framework for R. R package version 1.0.5.

Chelladurai, Y., Stevens, K.A., Haut, E.R., Brotman, D.J., Sharma, R., Shermock, K.M., Kebede, S., Singh, S., Segal, J.B., 2013. Venous thromboembolism prophylaxis in patients with traumatic brain injury: a systematic review. F1000Research 2.

Chow, S.-C., Chang, M., 2008. Adaptive design methods in clinical trials–a review. Orphanet J. Rare Dis. 3, 11.

Christensen, H., Hickman, M., Edmunds, W.J., Trotter, C.L., 2013. Introducing vaccination against serogroup B meningococcal disease: an economic and mathematical modelling study of potential impact. Vaccine 31, 2638–2646.

Claxton, K., 1999. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J. Health Econ. 18, 341–364.

Claxton, K., Briggs, A., Buxton, M.J., Culyer, A.J., McCabe, C., Walker, S., Sculpher, M.J., 2008. Value based pricing for NHS drugs: an opportunity not to be missed? BMJ 336, 251–254. https://doi.org/10.1136/bmj.39434.500185.25

Claxton, K., Griffin, S., Koffijberg, H., McKenna, C., 2015a. How to estimate the health benefits of additional research and changing clinical practice. BMJ 351, h5987. https://doi.org/10.1136/bmj.h5987

Claxton, K., Griffin, S.C., Koffijberg, H., McKenna, C., 2013. Expected health benefits of additional evidence: principles, methods and applications. University of York, Centre for Health Economics.

Claxton, K., Martin, S., Soares, M., Rice, N., Spackman, E., Hinde, S., Devlin, N., Smith, P.C., Sculpher, M., 2015b. Methods for the estimation of the National

Institute for Health and Care Excellence cost-effectiveness threshold. Health Technol. Assess. 19, 1–504. https://doi.org/10.3310/hta19140
Claxton, K., Palmer, S., Longworth, L., Bojke, L., Griffin, S., McKenna, C., Soares, M., Spackman, E., Youn, J., 2012. Informing a decision framework for when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development.
Claxton, K., Palmer, S., Longworth, L., Bojke, L., Griffin, S., Soares, M., Spackman, E., Rothery, C., 2016. A comprehensive algorithm for approval of health technologies with, without, or only in research: the key principles for informing coverage decisions. Value Health 19, 885–891.
Claxton, K.P., Sculpher, M.J., 2006. Using Value of Information Analysis to Prioritise Health Research. PharmacoEconomics 24, 1055–1068. https://doi.org/10.2165/00019053-200624110-00003
Claxton, K.P., Sculpher, M.J., Culyer, A.J., 2007. Mark versus Luke? Appropriate Methods for the Evaluation of Public Health Interventions (CHE Research Paper No. 31). Centre for Health Economics, York.
Collaborators, C.T., 2005. Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury—outcomes at 6 months. The Lancet 365, 1957–1959.
Conti, S., Claxton, K., 2009. Dimensions of Design Space: A Decision-Theoretic Approach to Optimal Research Design. Med. Decis. Making 29, 643–660. https://doi.org/10.1177/0272989X09336142
Cooper, N.J., Jones, D.R., Sutton, A.J., 2005. The use of systematic reviews when designing studies. Clin. Trials 2, 260–264.
Craig, T.K., Garety, P., Power, P., Rahaman, N., Colbert, S., Fornells-Ambrojo, M., Dunn, G., 2004. The Lambeth Early Onset (LEO) Team: randomised controlled trial of the effectiveness of specialised care for early psychosis. Bmj 329, 1067.
Culyer, A.J., 2016. Cost-effectiveness thresholds in health care: a bookshelf guide to their meaning and use. Health Econ. Policy Law 11, 415–432.
Culyer, A.J., 2012. The Humble Economist. University of York and Office of Health Economics.
Curtis, L., Burns, A., 2016. Unit Costs of Health & Social Care | Personal Social Services Research Unit.
Dakin, H., Abel, L., Burns, R., Yang, Y., 2018. Review and critical appraisal of studies mapping from quality of life or clinical measures to EQ-5D: an online database and application of the MAPS statement. Health Qual. Life Outcomes 16, 31.
Daniels, N., 2000. Accountability for reasonableness. BMJ 321, 1300–1301.
Deaths registered in England and Wales: 2014 | Office for National Statistics, 2015.
department for international development, 2018.
Dixon, S., Longworth, L., Wailoo, A., 2009. Assessing technologies at the end of life: a review of empirical evidence. Decis. Support Unit 41.
Djulbegovic, B., Kumar, A., Glasziou, P.P., Perera, R., Reljic, T., Dent, L., Raftery, J., Johansen, M., Di Tanna, G.L., Miladinovic, B., 2012a. New treatments compared to established treatments in randomized trials. Cochrane Database Syst. Rev. Online 10, MR000024.
Dolan, P., 1997. Modeling Valuations for Euroqol Health States. Med. Care 35, 1095–1108.

Drummond, M.F., Sculpher, M.J., Claxton, K., Stoddart, G.L., Torrance, G.W., 2015. Methods for the Economic Evaluation of Health Care Programmes. Oxford University Press.

Duarte, A., Mebrahtu, T., Goncalves, P.S., Harden, M., Murphy, R., Palmer, S., Woolacott, N., Rodgers, M., Rothery, C., 2017. Adalimumab, etanercept and ustekinumab for treating plaque psoriasis in children and young people: systematic review and economic evaluation. Health Technol. Assess. 1–244.

Eckermann, S., Willan, A.R., 2009. Globally optimal trial design for local decision making. Health Econ. 18, 203–216. https://doi.org/10.1002/hec.1353

Eckermann, S., Willan, A.R., 2008a. The Option Value of Delay in Health Technology Assessment. Med. Decis. Making 28, 300–305. https://doi.org/10.1177/0272989X07312477

Eckermann, S., Willan, A.R., 2008b. Time and Expected Value of Sample Information Wait for No Patient. Value Health 11, 522–526. https://doi.org/10.1111/j.1524-4733.2007.00296.x

Eckermann, S., Willan, A.R., 2007. Expected value of information and decision making in HTA. Health Econ. 16, 195–209. https://doi.org/10.1002/hec.1161

Eculizumab for treating atypical haemolytic uraemic syndrome | Highly specialised technologies guidance | National Institute for Health and Care Excellence, 2015.

Edney, L.C., Haji Ali Afzali, H., Cheng, T.C., Karnon, J., 2018. Estimating the Reference Incremental Cost-Effectiveness Ratio for the Australian Health System. PharmacoEconomics 36, 239–252. https://doi.org/10.1007/s40273-017-0585-2

Edwards, E., 2018. Ireland urgently needs access to new drugs, says pharma body [WWW Document]. Ir. Times. URL https://www.irishtimes.com/news/health/ireland-urgently-needs-access-to-new-drugs-says-pharma-body-1.3540257 (accessed 9.19.18).

Espinoza, M.A., Manca, A., Claxton, K., Sculpher, M.J., 2014. The Value of Heterogeneity for Cost-Effectiveness Subgroup Analysis: Conceptual Framework and Application. Med. Decis. Making 34, 951–964. https://doi.org/10.1177/0272989X14538705

European and developing countries clinical trials partnership, 2018. URL http://www.edctp.org/ (accessed 10.25.18).

Fenwick, E., Claxton, K., Sculpher, M., 2008. The Value of Implementation and the Value of Information: Combined and Uneven Development. Med. Decis. Making 28, 21–32. https://doi.org/10.1177/0272989X07308751

Fleurence, R.L., Meltzer, D.O., 2013. Toward a Science of Research Prioritization? The Use of Value of Information by Multidisciplinary Stakeholder Groups. Med. Decis. Making 33, 460–462. https://doi.org/10.1177/0272989X13486979

Fleurence, R.L., Torgerson, D.J., 2004. Setting priorities for research. Health Policy 69, 1–10. https://doi.org/10.1016/j.healthpol.2003.11.002

Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 12, 189–198.

Fremeaux-Bacchi, V., Fakhouri, F., Garnier, A., Bienaimé, F., Dragon-Durey, M.-A., Ngo, S., Moulin, B., Servais, A., Provot, F., Rostaing, L., 2013. Genetics and outcome of atypical hemolytic uremic syndrome: a nationwide French

series comparing children and adults. Clin. J. Am. Soc. Nephrol. CJN–04760512.
Fuller, G.W., Hernandez, M., Pallot, D., Lecky, F., Stevenson, M., Gabbe, B., 2017b. Health state preference weights for the Glasgow Outcome Scale following traumatic brain injury: a systematic review and mapping study. Value Health 20, 141–151.
Grieve, R., Abrams, K., Claxton, K., Goldacre, B., James, N., Nicholl, J., Parmar, M., Parker, C., Sekhon, J.S., Smeeth, L., 2016. Cancer Drugs Fund requires further reform. Bmj 354.
Griffin, S.C., Claxton, K.P., Palmer, S.J., Sculpher, M.J., 2011. Dangerous omissions: the consequences of ignoring decision uncertainty. Health Econ. 20, 212–224.
Grimm, S.E., Dixon, S., Stevens, J.W., 2017. Assessing the Expected Value of Research Studies in Reducing Uncertainty and Improving Implementation Dynamics. Med. Decis. Making 37, 523–533. https://doi.org/10.1177/0272989X16686766
Guyatt, G.H., Oxman, A.D., Vist, G.E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., Schünemann, H.J., 2008. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 336, 924–926. https://doi.org/10.1136/bmj.39489.470347.AD
Gyawali, B., Kesselheim, A.S., 2018. Reinforcing the social compromise of accelerated approval. Nat. Rev. Clin. Oncol. 1.
Heath, A., Manolopoulou, I., Baio, G., 2018. Efficient Monte Carlo Estimation of the Expected Value of Sample Information Using Moment Matching. Med. Decis. Making 38, 163–173. https://doi.org/10.1177/0272989X17738515
Heath, A., Manolopoulou, I., Baio, G., 2017. A review of methods for analysis of the expected value of information. Med. Decis. Making 37, 747–758.
Heller, J., 1961. Catch-22. Simon and Schuster.
Hemming, K., Haines, T.P., Chilton, P.J., Girling, A.J., Lilford, R.J., 2015. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. Bmj 350, h391.
Higgins, J.P.T., Green, S., 2011. Cochrane Handbook for Systematic Reviews of Interventions [WWW Document]. URL http://handbook-5-1.cochrane.org/ (accessed 7.4.19).
Higginson, I.J., Sen-Gupta, G.J.A., 2000. Place of care in advanced cancer: a qualitative systematic literature review of patient preferences. J. Palliat. Med. 3, 287–300.
HMT Green Book: supplementary guidance, 2013.
Howard, R.A., 1966. Information value theory. IEEE Trans. Syst. Sci. Cybern. 2, 22–26.
Hughes, D.A., Tunnage, B., Yeo, S.T., 2005. Drugs for exceptionally rare diseases: do they deserve special status for funding? QJM Int. J. Med. 98, 829–836. https://doi.org/10.1093/qjmed/hci128
Hughes-Hallett, Tom, Craft, A., Davies, C., 2011. Palliative Care Funding Review.
Hughes-Hallett, T., Craft, A., Davies, C., Mackay, I., Nielsson, T., 2011. Creating a fair and transparent funding system; the final report of the palliative care funding review. Lond. Palliat. Care Funding Rev.
Ibrahim, J.G., Chu, H., Chen, L.M., 2010. Basic Concepts and Methods for Joint Models of Longitudinal and Survival Data. J. Clin. Oncol. 28, 2796–2801. https://doi.org/10.1200/JCO.2009.25.0654

Ioannidis, J.P.A., Greenland, S., Hlatky, M.A., Khoury, M.J., Macleod, M.R., Moher, D., Schulz, K.F., Tibshirani, R., 2014. Increasing value and reducing waste in research design, conduct, and analysis. The Lancet 383, 166–175. https://doi.org/10.1016/S0140-6736(13)62227-8

Jaeschke, R., Singer, J., Guyatt, G.H., 1989. Measurement of health status: Ascertaining the minimal clinically important difference. Control. Clin. Trials 10, 407–415. https://doi.org/10.1016/0197-2456(89)90005-6

Jalal, H., Alarid-Escudero, F., 2018. A Gaussian Approximation Approach for Value of Information Analysis. Med. Decis. Making 38, 174–188. https://doi.org/10.1177/0272989X17715627

Jamjoom, A.A., Jamjoom, A.B., 2013. Safety and efficacy of early pharmacological thromboprophylaxis in traumatic brain injury: systematic review and meta-analysis. J. Neurotrauma 30, 503–511.

Johnson, J.R., Ning, Y.-M., Farrell, A., Justice, R., Keegan, P., Pazdur, R., 2011. Accelerated Approval of Oncology Products: The Food and Drug Administration Experience. JNCI J. Natl. Cancer Inst. 103, 636–644. https://doi.org/10.1093/jnci/djr062

Johnston, K., Levy, A.R., Lorigan, P., Maio, M., Lebbe, C., Middleton, M., Testori, A., Bédane, C., Konto, C., Dueymes, A., 2012. Economic impact of healthcare resource utilisation patterns among patients diagnosed with advanced melanoma in the United Kingdom, Italy, and France: results from a retrospective, longitudinal survey (MELODY study). Eur. J. Cancer 48, 2175–2182.

Jones, S.R., Carley, S., Harrison, M., 2003. An introduction to power and sample size estimation. Emerg. Med. J. 20, 453–458. https://doi.org/10.1136/emj.20.5.453

Jönsson, L., Andreasen, N., Kilander, L., Soininen, H., Waldemar, G., Nygaard, H., Winblad, B., Jönhagen, M.E., Hallikainen, M., Wimo, A., 2006. Patient-and proxy-reported utility in Alzheimer disease using the EuroQoL. Alzheimer Dis. Assoc. Disord. 20, 49–55.

Kemp, R., Prasad, V., 2017. Surrogate endpoints in oncology: when are they acceptable for regulatory and clinical decisions, and are they currently overused? BMC Med. 15, 134.

Kieny, M.P., Viergever, R.F., Adam, T., Boerma, T., Røttingen, J.-A., 2016. Global platform to inform investments for health R&D. The Lancet 387, 1157.

Kirkbride, J.B., Jackson, D., Perez, J., Fowler, D., Winton, F., Coid, J.W., Murray, R.M., Jones, P.B., 2013. A population-level prediction tool for the incidence of first-episode psychosis: translational epidemiology based on cross-sectional data. BMJ Open 3, e001998.

Laplace, P.-S., 1829. Essai philosophique sur les probabilités. H. Remy.

Lenert, L.A., Sturley, A.P., Rapaport, M.H., Chavez, S., Mohr, P.E., Rupnow, M., 2004. Public preferences for health states with schizophrenia and a mapping function to estimate utilities from positive and negative symptom scale scores. Schizophr. Res. 71, 155–165.

Lomas, J., Claxton, K., Martin, S., Soares, M., 2018. Resolving the "Cost-Effective but Unaffordable" Paradox: Estimating the Health Opportunity Costs of Nonmarginal Budget Impacts. Value Health 21, 266–275.

Love-Koh, J., Peel, A., Rejon-Parrilla, J.C., Ennis, K., Lovett, R., Manca, A., Chalkidou, A., Wood, H., Taylor, M., 2018. The Future of Precision Medicine: Potential Impacts for Health Technology Assessment.

PharmacoEconomics 36, 1439–1451. https://doi.org/10.1007/s40273-018-0686-6
Macmillan Online survey of 1,019 people living with cancer, 2010.
Mason, A.J., Gomes, M., Grieve, R., Ulug, P., Powell, J.T., Carpenter, J., 2017. Development of a practical approach to expert elicitation for randomised controlled trials with missing health outcomes: Application to the IMPROVE trial. Clin. Trials Lond. Engl. 14, 357–367. https://doi.org/10.1177/1740774517711442
McCabe, C., Paul, A., Fell, G., Paulden, M., 2016. Cancer Drugs Fund 2.0: A Missed Opportunity? PharmacoEconomics 34, 629–633. https://doi.org/10.1007/s40273-016-0403-2
McCabe, C., Tsuchiya, A., Claxton, K., Raftery, J., 2006. Orphan drugs revisited. QJM Int. J. Med. 99, 341–345. https://doi.org/10.1093/qjmed/hcl033
McKenna, C., Claxton, K., 2011. Addressing adoption and research design decisions simultaneously: the role of value of sample information analysis. Med. Decis. Making 31, 853–865.
McKenna, C., Griffin, S., Koffijberg, H., Claxton, K., 2016. Methods to place a value on additional evidence are illustrated using a case study of corticosteroids after traumatic brain injury. J. Clin. Epidemiol. 70, 183–190. https://doi.org/10.1016/j.jclinepi.2015.09.011
McKenna, C., Soares, M., Claxton, K., Bojke, L., Griffin, S., Palmer, S., Spackman, E., 2015. Unifying Research and Reimbursement Decisions: Case Studies Demonstrating the Sequence of Assessment and Judgments Required. Value Health 18, 865–875. https://doi.org/10.1016/j.jval.2015.05.003
Medicines Complete, 2018. URL www.medicinescomplete.com (accessed 1.11.18).
Meltzer, D.O., Hoomans, T., Chung, J.W., Basu, A., 2011. Minimal Modeling Approaches to Value of Information Analysis for Health Research. Med. Decis. Making 31, E1–E22. https://doi.org/10.1177/0272989X11412975
Menzies, N.A., 2016. An Efficient Estimator for the Expected Value of Sample Information. Med. Decis. Making 36, 308–320. https://doi.org/10.1177/0272989X15583495
Mills, E.J., Thorlund, K., Ioannidis, J.P.A., 2012. Calculating additive treatment effects from multiple randomized trials provides useful estimates of combination therapies. J. Clin. Epidemiol. 65, 1282–1288. https://doi.org/10.1016/j.jclinepi.2012.07.012
Montori, V.M., Permanyer-Miralda, G., Ferreira-González, I., Busse, J.W., Pacheco-Huergo, V., Bryant, D., Alonso, J., Akl, E.A., Domingo-Salvany, A., Mills, E., Wu, P., Schünemann, H.J., Jaeschke, R., Guyatt, G.H., 2005. Validity of composite end points in clinical trials. BMJ 330, 594–596. https://doi.org/10.1136/bmj.330.7491.594
Mooney, G., 2012. The Health of Nations, 1st ed. Zed Books, London.
Morgan, B., Hejdenberg, J., Hinrichs-Krapels, S., Armstrong, D., 2018. Do feasibility studies contribute to, or avoid, waste in research? PloS One 13, e0195951.
Myers, E., McBroom, A.J., Shen, L., Posey, R.E., Gray, M.R., Sanders, G.D., 2012. Value-of-information analysis for patient-centered outcomes research prioritization. Rep. Prep. Duke Evid.-Based Pract. Cent. Patient-Centered Outcomes Res. Inst.
NHS England, 2015. Proposal for supplementary prescribing by dietitians.
NHS England NHS standard contract for adult kidney transplant service, 2013.

NICE, 2016. Accelerated Access Collaborative (AAC) [WWW Document]. URL https://www.nice.org.uk/aac (accessed 5.2.19).
NICE, 2014. CG 176 Head Injury: Triage, assessment, investigation and early management of head injury in children, young people and adults.
NICE, 2013. Guide to the methods of technology appraisal 2013.
Nichol, A., French, C., Little, L., Haddad, S., Presneill, J., Arabi, Y., Bailey, M., Cooper, D.J., Duranteau, J., Huet, O., 2015. Erythropoietin in traumatic brain injury (EPO-TBI): a double-blind randomised controlled trial. The Lancet 386, 2499–2506.
NIHR, 2019a. Research for Patient Benefit | Transparency table.
NIHR, 2019b. Efficacy and Mechanism Evaluation | Reviewer Assessment Form.
NIHR, 2018a. NIHR Evaluation, Trials and Studies Coordinating Centre [WWW Document]. URL https://www.nihr.ac.uk/about-us/how-we-are-managed/managing-centres/nihr-evaluation-trials-and-studies-coordinating-centre/ (accessed 9.4.18).
NIHR, 2018b. HTA application success rates [WWW Document]. URL https://www.nihr.ac.uk/funding-and-support/documents/funding-for-research-studies/research-programmes/HTA/hta-application-success-rates-2018.pdf
NIHR, 2017. Supporting information for applicants applying to the hta programme.
NIHR, 2015. Invention for Innovation (i4i) | Guidance for panel members [WWW Document]. URL https://www.nihr.ac.uk/about-us/CCF/i4i/Guidance%20for%20i4i%20Panel%20Members_Challenge.pdf
NIHR HTA, 2018. Health Technology Assessment [WWW Document]. URL https://www.nihr.ac.uk/funding-and-support/funding-for-research-studies/funding-programmes/health-technology-assessment/ (accessed 9.4.18).
Noris, M., Remuzzi, G., 2010. Thrombotic microangiopathy after kidney transplantation. Am. J. Transplant. 10, 1517–1523.
Nyein, K., Turner-Stokes, L., Robinson, I., 1999. The Northwick Park Care Needs Assessment (NPCNA): A measure of community care needs: Sensitivity to change during rehabilitation. Clin. Rehabil. 13, 482–491.
Ochalek, J., Lomas, J., Claxton, K., 2018a. Estimating health opportunity costs in low-income and middle-income countries: a novel approach and evidence from cross-country data. BMJ Glob. Health 3, e000964. https://doi.org/10.1136/bmjgh-2018-000964
Ochalek, J., Revill, P., Manthalu, G., McGuire, F., Nkhoma, D., Rollinger, A., Sculpher, M., Claxton, K., 2018b. Supporting the development of a health benefits package in Malawi. BMJ Glob. Health 3, e000607.
Office for National Statistics, 2014. Mortality statistics: Deaths registered in England and Wales in 2013. London.
O'Hagan, A., Buck, C.E., Daneshkhah, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E., Rakow, T., 2006. Uncertain judgements: eliciting experts' probabilities. John Wiley & Sons.
O'Hagan, A., Stevens, J.W., Campbell, M.J., 2005. Assurance in clinical trial design. Pharm. Stat. 4, 187–201. https://doi.org/10.1002/pst.175
O'Hagan, T., Oakley, J.E., 2018. SHELF version 3.0 [WWW Document]. URL http://www.tonyohagan.co.uk/shelf/

Pembrolizumab for advanced melanoma not previously treated with ipilimumab | Technology appraisal guidance | National Institute for Health and Care Excellence, 2015.
Phelps, C.E., Mushlin, A.I., 1988. Focusing Technology Assessment Using Medical Decision Theory. Med. Decis. Making 8, 279–289. https://doi.org/10.1177/0272989X8800800409
Philips, Z., Claxton, K., Palmer, S., 2008. The Half-Life of Truth: What Are Appropriate Time Horizons for Research Decisions? Med. Decis. Making 28, 287–299. https://doi.org/10.1177/0272989X07312724
Prasad, V., 2017. Do cancer drugs improve survival or quality of life? British Medical Journal Publishing Group.
Programme, N.E. of L.C., 2012. Reviewing End of Life Care Costing Information to Inform the QIPP End of Life Care Workstream. Whole Syst. Partnersh.
R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rawlins, M.D., 2005. Pharmacopolitics and deliberative democracy. Clin. Med. 5, 471–475. https://doi.org/10.7861/clinmedicine.5-5-471
Relton, C., Torgerson, D., O'Cathain, A., Nicholl, J., 2010. Rethinking pragmatic randomised controlled trials: introducing the "cohort multiple randomised controlled trial" design. Bmj 340, c1066.
Remme, M., Martinez-Alvarez, M., Vassall, A., 2017. Cost-effectiveness thresholds in global health: taking a multisectoral perspective. Value Health 20, 699–704.
Reviewing end of life care costing information to inform the QIPP End of Life Care Workstream, 2012.
Rhodes, K.M., Turner, R.M., Higgins, J.P.T., 2015. Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. J. Clin. Epidemiol. 68, 52–60. https://doi.org/10.1016/j.jclinepi.2014.08.012
Riley, R.D., 2009. Multivariate meta-analysis: the effect of ignoring within-study correlation. J. R. Stat. Soc. Ser. A Stat. Soc. 172, 789–811.
Robert, C., Long, G.V., Brady, B., Dutriaux, C., Maio, M., Mortier, L., Hassel, J.C., Rutkowski, P., McNeil, C., Kalinka-Warzocha, E., Savage, K.J., Hernberg, M.M., Lebbé, C., Charles, J., Mihalcioiu, C., Chiarion-Sileni, V., Mauch, C., Cognetti, F., Arance, A., Schmidt, H., Schadendorf, D., Gogas, H., Lundgren-Eriksson, L., Horak, C., Sharkey, B., Waxman, I.M., Atkinson, V., Ascierto, P.A., 2015. Nivolumab in Previously Untreated Melanoma without *BRAF* Mutation. N. Engl. J. Med. 372, 320–330. https://doi.org/10.1056/NEJMoa1412082
Rothery, C., Claxton, K., Palmer, S., Epstein, D., Tarricone, R., Sculpher, M., 2017. Characterising uncertainty in the assessment of medical devices and determining future research needs. Health Econ. 26, 109–123.
Rothery, C., Woods, B., Schmitt, L., Claxton, K., Palmer, S., Sculpher, M., 2018. Framework for value assessment of new antimicrobials 178.
Røttingen, J.-A., Regmi, S., Eide, M., Young, A.J., Viergever, R.F., \AArdal, C., Guzman, J., Edwards, D., Matlin, S.A., Terry, R.F., 2013. Mapping of available health research and development data: what's there, what's missing, and what role is there for a global observatory? The Lancet 382, 1286–1307.
Sanders, G.D., Neumann, P.J., Basu, A., Brock, D.W., Feeny, D., Krahn, M., Kuntz, K.M., Meltzer, D.O., Owens, D.K., Prosser, L.A., Salomon, J.A., Sculpher,

M.J., Trikalinos, T.A., Russell, L.B., Siegel, J.E., Ganiats, T.G., 2016. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. JAMA 316, 1093–1103. https://doi.org/10.1001/jama.2016.12195

Sauerland, Stefan, Maegele, M., 2004. A CRASH landing in severe head injury. The Lancet 364, 1291–1292.

Schlaifer, R., Raiffa, H., 1961. Applied statistical decision theory.

Schneider, L.S., Mangialasche, F., Andreasen, N., Feldman, H., Giacobini, E., Jones, R., Mantua, V., Mecocci, P., Pani, L., Winblad, B., 2014. Clinical trials and late-stage drug development for Alzheimer's disease: an appraisal from 1984 to 2014. J. Intern. Med. 275, 251–283.

Shavelle, R., Strauss, D., Day, S., 2006. Life expectancy, in: Brain Injury Medicine: Principles and Practice. Demos Medical Publishing.

Soares, M.O., Bojke, L., Dumville, J., Iglesias, C., Cullum, N., Claxton, K., 2011. Methods to elicit experts' beliefs over uncertain quantities: application to a cost effectiveness transition model of negative pressure wound therapy for severe pressure ulceration. Stat. Med. 30, 2363–2380. https://doi.org/10.1002/sim.4288

Soares, M.O., Walker, S., Palmer, S.J., Sculpher, M.J., 2018. Establishing the value of diagnostic and prognostic tests in Health Technology Assessment. Med. Decis. Making 38, 495–508.

Strong, M., 2012. Managing Structural Uncertainty in Health Economic Decision Models. University of Sheffield.

Strong, M., Oakley, J.E., Brennan, A., 2014. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. Med. Decis. Making 34, 311–326.

Strong, M., Oakley, J.E., Brennan, A., Breeze, P., 2015b. Estimating the Expected Value of Sample Information Using the Probabilistic Sensitivity Analysis Sample: A Fast, Nonparametric Regression-Based Method. Med. Decis. Making 35, 570–583. https://doi.org/10.1177/0272989X15575286

Sullivan, P.W., Slejko, J.F., Sculpher, M.J., Ghushchyan, V., 2011. Catalogue of EQ-5D scores for the United Kingdom. Med. Decis. Making 31, 800–804.

Sutton, A.J., Cooper, N.J., Jones, D.R., Lambert, P.C., Thompson, J.R., Abrams, K.R., 2007. Evidence-based sample size calculations based upon updated meta-analysis. Stat. Med. 26, 2479–2500.

Svensson, S., Menkes, D.B., Lexchin, J., 2013. Surrogate outcomes in clinical trials: a cautionary tale. JAMA Intern. Med. 173, 611–612.

The James Lind Alliance [WWW Document], 2019. URL http://www.jla.nihr.ac.uk/ (accessed 7.15.19).

Thijssen, J.J.J., Bregantini, D., 2017. Costly sequential experimentation and project valuation with an application to health technology assessment. J. Econ. Dyn. Control 77, 202–229. https://doi.org/10.1016/j.jedc.2017.01.016

Thorlund, K., Golchi, S., Mills, E., 2017. Bayesian adaptive clinical trials of combination treatments. Contemp. Clin. Trials Commun. 8, 227–233.

Treasury, H.M., 2003. The green book: appraisal and evaluation in central government. London TSO.

US national institutes of health, 2018. URL https://www.nih.gov/ (accessed 10.25.18).

Vallejo-Torres, L., García-Lorenzo, B., Serrano-Aguilar, P., 2018. Estimating a cost-effectiveness threshold for the Spanish NHS. Health Econ. 27, 746–761. https://doi.org/10.1002/hec.3633

Walker, S., Sculpher, M., Claxton, K., Palmer, S., 2012. Coverage with Evidence Development, Only in Research, Risk Sharing, or Patient Access Scheme? A Framework for Coverage Decisions. Value Health 15, 570–579. https://doi.org/10.1016/j.jval.2011.12.013

Wang, L., Rotnitzky, A., Lin, X., Millikan, R.E., Thall, P.F., 2012. Evaluation of viable dynamic treatment regimes in a sequentially randomized trial of advanced prostate cancer. J. Am. Stat. Assoc. 107, 493–508.

Willan, A.R., 2008. Optimal sample size determinations from an industry perspective based on the expected value of information. Clin. Trials 5, 587–594. https://doi.org/10.1177/1740774508098413

Willan, A.R., Eckermann, S., 2012. Value of Information and Pricing New Healthcare Interventions. PharmacoEconomics 30, 447–459. https://doi.org/10.2165/11592250-000000000-00000

Willan, A.R., Eckermann, S., 2009. Optimal clinical trial design using value of information methods with imperfect implementation. Health Econ. n/a-n/a. https://doi.org/10.1002/hec.1493

Willan, A.R., Pinto, E.M., 2005. The value of information and optimal clinical trial design. Stat. Med. 24, 1791–1806. https://doi.org/10.1002/sim.2069

Wilson, E.C.F., 2015. A Practical Guide to Value of Information Analysis. PharmacoEconomics 33, 105–121. https://doi.org/10.1007/s40273-014-0219-x

Wilson, E.C.F., Usher-Smith, J.A., Emery, J., Corrie, P.G., Walter, F.M., 2018. Expert Elicitation of Multinomial Probabilities for Decision-Analytic Modeling: An Application to Rates of Disease Progression in Undiagnosed and Untreated Melanoma. Value Health 21, 669–676. https://doi.org/10.1016/j.jval.2017.10.009

Wolstenholme, J.L., Fenn, P., Gray, A.M., Keene, J., Jacoby, R., HOPE, T., 2002. Estimating the relationship between disease progression and cost of care in dementia. Br. J. Psychiatry 181, 36–42.

Wood, R.L., McCrea, J.D., Wood, L.M., Merriman, R.N., 1999. Clinical and cost effectiveness of post-acute neurobehavioural rehabilitation. Brain Inj. 13, 69–88.

Woods, B.S., Rothery, C., Revill, P., Hallett, T., Phillips, A., 2018. Setting research priorities in Global Health: Appraising the value of evidence generation activities to support decision-making in health care.