Conceptual models in health economic evaluation: a new role

Nicola Jane Michael McMeekin
MSc

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Health Economics and Health Technology Assessment (HEHTA)
Institute of Health and Wellbeing
College of Medical, Veterinary and Life Science
University of Glasgow

March 2021

Author's Declaration

"I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution."

Printed Name: Nicola McMeekin

Signature:

Acknowledgements

First, I would like to express my gratitude to my supervisors Professor Andrew Briggs and Professor Olivia Wu. I would like to thank them for giving me the opportunity to undertake this PhD and for their invaluable encouragement, advice, support, guidance, and most importantly their patience, throughout my PhD journey.

I would like to thank Professor Graham Devereux and Professor Andrew Jahoda for giving their agreement to use the TWICS and BeatIt clinical trials as the case studies in this thesis.

I am grateful to my HEHTA colleagues, past and present, for their support and for their ability to know when to ask, and more importantly when not to ask, me about my PhD.

I am indebted to Craig Runnett, Lesley Hamilton and John Dark, without whom none of this would have been possible, all the gratitude in the world is not and never will be enough.

Many thanks to my husband Peter for his support, willingness to patiently listen to endless procrastination and his unquotable quotes. Thanks are due too to my family for their support and encouragement.

Thank you to Phoebe who ensured that I took regular breaks during the writing up of my thesis, and who has shown me the true meaning of courage, trust and second chances.

And last, but no means least, I would like to dedicate this thesis to my beautiful granddaughter who arrived during my PhD journey and has brightened every day since.

Publications and presentations

Published papers

McMeekin, N., Wu, O., Germeni, E. Briggs A. How methodological frameworks are being developed: evidence from a scoping review. BMC Medical Research Methodology 20, Article number 173 (June 2020). https://doi.org/10.1186/s12874-020-01061-4

Devereux G, Cotton S, Fielding S, McMeekin N, Barnes PJ, Briggs A, et al. Low-dose oral theophylline combined with inhaled corticosteroids for people with chronic obstructive pulmonary disease and high risk of exacerbations: a RCT. *Health Technol Assess*. 2019;23(37).

Jahoda A, Hastings R, Hatton C, Cooper S-A, McMeekin N, Dagnan D, et al. Behavioural activation versus guided self-help for depression in adults with learning disabilities: the Beatlt RCT. Health Technology Assessment. 2018 Sep;22(53):VII-+.

Presentations

Conceptual model guided trial-based health economic analysis: increasing understanding and reducing variance? Health Economists' Study Group (HESG) Winter conference, January 2021, London. Paper presented by another participant followed by questions to author and discussion.

A methodological framework for developing conceptual models in economic evaluation, Poster presentation Health Economists' Study Group Winter conference, January 2019, York

Using conceptual models to unravel negative trial results: Analysis, interpretation and presentation, Oral presentation European Health Economics Association conference, July 2018, Maastricht

Developing a conceptual modelling framework for economic evaluation, Poster

ISPOR 20th Annual European Congress,, November 2017, Glasgow, UK

Conceptual modelling for economic evaluation – can frameworks from other disciplines inform an economic evaluation framework? Poster presentation European Health Economics Association conference, July 2016, Hamburg

Working papers

McMeekin N, Briggs A, Wu O. Conceptual modelling in health economic evaluation: developing a methodological framework.

McMeekin N, Briggs A, Wu O. Conceptual model guided trial-based health economic analysis: increasing understanding and reducing variance?

Abstract

Healthcare budgets are limited, and decisions must be made about which healthcare technologies should be funded from these limited budgets. Decision makers rely on clinical and cost-effectiveness evidence on which to base their decisions, the gold standard vehicle for this evidence is a clinical trial. Guidance exists for conducting economic evaluations alongside clinical trials using standard treatment arm-based comparisons to assess cost-effectiveness, however because of differences in the purposes of the clinical effectiveness and cost-effectiveness analyses this conventional treatment arm-based economic evaluation is often uncertain. Furthermore, economic evaluations focus on the final cost-effectiveness summary measure without consideration of what is driving this measure. There is scope for providing a more detailed understanding of the conventional cost-effectiveness results.

Conceptual models are simplified versions of real-life systems in a visual format, illustrating how key components of the system are linked and interact within it. In the field of economic evaluation conceptual models are recommended for and used as a guide to establish the structure of decision analytic models, in other fields they are also used as a communication tool, to aid understanding of the system and to give direction to research.

The aim of this thesis was to expand the role of conceptual models in economic evaluation. This thesis proposed a new role for conceptual models to provide additional understanding to decision makers, extending the conventional economic evaluation analysis beyond a treatment arm-based analysis. This role was demonstrated using two case studies as illustrative examples, aiming to show how the role could be applied.

This thesis demonstrates the value of an additional conceptual model driven analysis to supplement the conventional treatment arm-based analysis, adding further insight into the trial mechanism and what is driving the economic evaluation results. The contribution of this thesis to the field of economic evaluation is twofold; a new role for conceptual models in economic evaluation and a methodological framework for developing conceptual models in this new role.

Table of Contents

List of Tables

List of Figures

Definitions and abbreviations

Conceptual model

A visual representation of the causal relationships linking key components in a real-life system using simplifying assumptions. A conceptual model provides a framework for hanging information on about the interactions and connections between these key components.

Methodological framework

A structured practical tool for guiding the user through a stepwise process, enabling/facilitating a standardised approach to the given task.

Mathematical model

A simulation-based model commonly described as a decision analytic model in the field of economic evaluation. It is an umbrella term for all models based on mathematical links.

Treatment/intervention and comparator

Treatment is any therapy, pharmaceutical or other care, given to the participants in the clinical trial and compared with other therapies, pharmaceuticals or other care to answer the research question of the clinical trial. An intervention is the new treatment which is being compared to an existing treatment. The comparator is the existing current treatment that the new treatment (intervention) is being compared to.

BNF	British National Formulary
BODE	Body mass index, airflow Obstruction, Dyspnea, and Exercise
C	Cost (in equations)
CBT	Cognitive behavioural therapy
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
CI	Confidence interval
COPD	Chronic obstructive pulmonary disease
CSRI	Client service receipt inventory
E	Effects (in equations)
EQ-5D	Euroqol-5 Dimension
EQ-5D-Y	Euroqol-5 Dimension Youth
EQ-5D-3L	Euroqol-5 Dimension 3-Levels
EVPI	Expected value of perfect information
EVPPI	Expected value of perfect parameter information
GDS-LD	Glasgow Depression Scale for Learning Disabilities
GOLD	Global initiative for chronic Obstructive Lung Disease
HIS	Health Improvement Scotland
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
ICS	Inhaled corticosteroids
ISD	Information Services Division
ISPOR	The Professional Society for Health Economics and Outcomes Research
IRR	Incidence rate ratio
LABA	Long-acting $ß_2$ agonists
mg	Milligram
NMB	Net monetary benefit
NHS	National Health Service
NICE	National Institute for Health and Care Excellence]
NIHR	National Institute for Health Research
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSSRU	Personal Social Services Research Unit
QALY	Quality adjusted life-year
RCT	Randomised controlled trial
SD	Standard deviation
SHTG	Scottish Health Technologies Group
SMC	Scottish Medicines Consortium
SSM	Soft systems methodology
TWICS	Theophylline With Corticosteroids
UK	United Kingdom
US	United States

Chapter 1 Introduction

1.1 Chapter overview

The aim of this chapter is to provide an introduction and orientation to the thesis. In Section 1.2 the rationale of the thesis is explained, in Section 1.3 the research question, aim and objectives are presented, and in Section 1.4 an overview of the thesis structure is provided.

1.2 Rationale

Health care resources in most countries are scarce and limited, therefore not all technologies and treatments can be funded and made available for patients, and decisions need to be made about which technologies and treatments to fund. These decisions are based on clinical efficacy, clinical effectiveness and availability, and cost-effectiveness evidence. To establish the clinical benefits of a technology for informing decision makers, a clinical trial is needed and is considered the gold standard vehicle for this evidence, the cost-effectiveness of the technology is frequently evaluated alongside the clinical trial (1).

The cost-effectiveness of a new technology is assessed using economic evaluation techniques, the purpose of which are to identify the healthcare technologies which deliver the maximum additional health benefits per additional unit of resource, this information is used to inform decision makers and help them to make the best use of the limited healthcare resources (2). An economic evaluation compares both the costs and the health benefits for one or more healthcare technologies to assess which of those technologies is worth funding compared to other technologies which could be funded with the same resources (1, 3). Economic evaluations alongside trials are specifically designed to answer cost-effectiveness questions and can influence the trial design, with input from a health economist. The clinical effectiveness of an intervention is typically the primary outcome of a clinical trial, with the cost-effectiveness typically a secondary outcome.

Guidance for conducting an economic evaluation alongside a clinical trial is well established and provides health economists with best practice (4-6), this

guidance focusses on a treatment arm-based approach where the costs and outcomes of all relevant technologies are compared to assess cost-effectiveness. However, the objectives and audience for a clinical analysis and a cost-effectiveness analysis differ, these conflicting objectives can result in uncertain economic evaluation results (where there is no difference between cost or health benefits). Recent research found that conflicting objectives can lead to conflicting conclusions being drawn in clinical trials; almost a third of trials identified in this research reported at least one 'doubly null' result (where there was no statistical difference between arms in the primary outcome or the cost per participant), but reported a favourable cost-effectiveness result (7). The extend of uncertain cost-effectiveness results is reported by Hollingworth et al. in research which found conclusions on primary outcomes were more likely than conclusions on economic outcomes (42.1% v 15.8%)(8). This research also found that only 1/38 trials considered a sample size based on economic outcomes; powering a trial on economic outcomes requires a larger sample size and longer follow-up to detect statistically significant differences and avoid uncertain results (9). Furthermore, economic evaluations have been described as producing a 'black box' analysis which focusses entirely on the final summary measure, with little or no consideration of what is driving this measure (10). These drawbacks of economic evaluation provide scope to develop an approach which gives additional interpretation to uncertain results and goes beyond the conventional outcomes focussed economic evaluation to provide an understanding of the trial mechanism driving the results. In this thesis I consider capturing the causal relationships inherent in conceptual models to further investigate valuable clinical trial data, providing additional evidence to decision makers for allocating scarce healthcare resources. Proposing and demonstrating this supplementary conceptual model driven analysis is the first objective in this thesis.

Although an early description of a conceptual model from 1976 exists; 'a simplification of reality' (11), today there is no widely accepted definition of a conceptual model, and any definitions are unclear and interpreted in different ways (12). However, despite there being no widely accepted definition there are commonalities in existing definitions of a conceptual model: a top level visual representation of a real-world system in a simplified form (13); it illustrates

causal relationships (14) using simplifying assumptions (15); linking the key constituents in a process or system, and explaining how these interconnect and interact (16); an abstraction of the real world (13, 17, 18), simplifying the decision problem (19); it provides a rack or framework on which to hang information from the system (20); a compass (13), which formalises understanding of the dynamics, and links in the process or system (20).

In disciplines other than economic evaluation, the primary purpose of a conceptual model is an effective communication tool due to its visual and simplified nature, making understanding of the causal relationships straightforward (14). They can also help gain an understanding of the system, objectives, rationale and assumptions used (14), and give shape and direction to the research (16). They can define the study context, research objectives and assumptions (13). While a conceptual model should always precede a mathematical model to inform the development of the mathematical model, not all conceptual models will lead to a mathematical model, it can be an end in itself, helping the user to understand a problem and allowing decisions to be made (15). A conceptual model can also enable the understanding of the relationships between key components (principal characteristics) of a system (15). When conceptual models are part of a mathematical model process they are crucial for a successful mathematical model (21), often developing the conceptual model is more important than developing the mathematical one (19, 20), and should be independent of the software used for a mathematical model (17, 18).

Definitions of conceptual models in the field of economic evaluation include the abstraction, simplification, and depiction of components of reality that are related to the decision problem, allowing understanding of the decision problem to be shared and agreed between interested parties and a mathematical model to be implemented (22, 23).

In the economic evaluation discipline conceptual modelling has emerged relatively recently, an early mention of conceptual modelling was in 2000, this briefly described the usefulness of using conceptual models in determining the structure of mathematical models (commonly known as ‘decision analytic models’ in economic evaluation, and from now on in this thesis described as

decision analytic models), but it did not go into detail on how to operationalise this (24). In 2010 Chilcott et al. conducted research into how modellers develop decision analytic models and a conceptual modelling stage was identified by nine out of 12 modellers interviewed (25). In 2012 two sets of guidance were published for developing conceptual models in economic evaluation (23, 26), a further set of guidance was published in 2016 (22). However, all mentions of conceptual modelling in economic evaluation relate to developing the structure of decision analytic models, and almost none relate to a trial-based evaluation. In economic evaluation the main purpose of a conceptual model is the first step in developing a decision analytic model (22); all decision analytic models should be based on a conceptual model (23). A decision analytic model is used to compare costs and health outcomes between treatment arms to estimate cost-effectiveness, particularly when there is scarce evidence. There are many benefits of using a conceptual model in this way in an economic evaluation, all linked to improving the quality of the decision analytic model: ensuring it answers the correct question and meets the needs of the stakeholders; making sure that there is a common understanding and agreement of the problem; helping stakeholders to understand the impact of the intervention on economic outputs (costs and outcomes), ensuring that the decision analytic model is clinically correct and that all relevant components (events, resources, outcomes) are included; it provides a reference point; it highlights any differences in clinical practice; it enables validity and credibility of the decision analytic model; it verifies any structural uncertainty analysis, identifying areas for future research; enabling transparency, and leading to efficient model development (22, 23).

Current guidance for developing conceptual models for use in economic evaluations focusses solely on the purpose of a conceptual model to determine the structure of a decision analytic model. This existing guidance was mentioned above and is described in more detail below.

In 2010 Chilcott et al. conducted research to establish how modellers develop decision analytic models; modellers were interviewed to determine their methods and strategies for developing these models (25). In synthesising the research Chilcott et al. identified five steps that modellers used to develop a

decision analytic model, the second step was conceptual modelling. Nine out of 12 modellers implicitly or explicitly acknowledged ‘conceptualisation and abstraction’ of the problem prior to developing the decision analytic model. The main purpose of this was to come to an agreement on the problem and proposed decision analytic model, as well as sense-checking, and developing ideas. Overall, the abstraction of the decision problem into a conceptual model served as a communication tool between the research team, decision maker and client. The authors acknowledged that without specific conceptual modelling the overall validity of a decision analytic model is compromised and recommended the publication of definitions for conceptual model validation.

Following Chilcott’s research three sets of guidelines for conceptual modelling in health economic evaluations have been published. These guidelines are summarised below.

The first guidance was published in 2012; Tappenden published generalised guidance on developing conceptual models for health economic model development, a process for determining a decision analytic model structure (23). Tappenden’s definition of a conceptual model is the ‘abstraction and representation’ of the decision problem which is used to discuss and agree stakeholder understanding of the decision system and its representation in the decision analytic model. Two types of conceptual model are considered: problem-oriented and design-oriented. Problem-oriented conceptual models are developed to understand the decision problem and the system relevant to the decision problem, analogous to asking, ‘what is relevant’? This stage is particularly helpful for communication with stakeholders and agreeing on the description of the decision system for a clinical understanding of the disease and treatment pathways. A design-oriented conceptual model focusses on designing, specifying and justifying the decision analytic model and structure, analogous to asking, ‘what is feasible’? It sets a boundary around the scope of the decision analytic model. Tappenden echoes Chilcott’s view that conceptual modelling is directly linked to the credibility and validation of the decision analytic model. Tappenden’s paper sets out a ‘practical framework’ for deciding on a decision analytic model’s structure, with ‘practical approaches’ and includes case studies to illustrate the practical guidance. The guidance consists of a list of suggested

questions the modeller could ask to help develop the conceptual models as well as high-level 'recommendations for practice' for each type of conceptual model and potential sources of evidence to inform the models. Tappenden's guidance was summarised in the National Institute for Health and Care Excellence (NICE) Interim Methods Guide for Developing Service Guidance 2014 (27).

Also published in 2012 was the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 'consensus based best practice' guidance for conceptualising a decision analytic model authored by Roberts et al. (26). This best practice was developed by experts in the modelling field from different types of organisations and countries and defines the decision analytic model as one which will 'inform medical decisions and health-related resource allocation questions'. Similar to Tappenden's guidance it focusses on developing the structure of decision analytic models in two steps: conceptualising the problem and conceptualising the model. The first step focusses on understanding the decision problem, perspective of the study, model population, outcomes and valuing them, comparators, time horizon and uncertainties requiring additional sensitivity analysis. The conceptualising the model section focusses on the most suitable model structure for the problem, with key considerations of whether the model represents individuals or groups, and whether there is any interaction between these individuals or groups. Examples of model structure include decision trees and state transition models. Each 'best practice' is accompanied by recommendations. The Roberts et al. best practice is presented as guidance not as a methodological framework and has been described as giving guidance on 'what to do' but not 'how to do it'(22).

The third set of guidance was published more recently in 2016; Squires et al. published a methodological framework specifically for developing public health conceptual models (22), providing the first public health specific standardised approach to developing a decision analytic model structure using conceptual models. The authors argue that due to the complex nature of public health interventions more thought should be applied to developing a public health conceptual model compared to a more straightforward conceptual model, for example the comparison of drug treatments. Squires et al. define conceptual modelling as an 'abstraction of elements of reality at an appropriate level of

simplification for the problem'. Squires et al. developed a draft conceptual modelling framework using information from a literature review and qualitative work, which was then tested with a case study, and revised to produce a final framework. The framework comprises four sections, steps within the last two sections guide the modeller through the process of developing a conceptual model to inform the decision analytic model structure.

Conceptual models are recommended for, and conventionally used in economic evaluation to develop decision analytic model structures. This thesis considers a new role for conceptual models; using their visual nature and the identification of relationships between components within them, to illustrate and determine the links between key components of an economic evaluation alongside a clinical trial to further explore clinical trial data. As previously discussed, current guidance for developing conceptual models in economic evaluation is restricted to developing conceptual models with the sole purpose of informing the structure design of a decision analytic model, therefore new guidance for developing conceptual models to use in the new role is required; this is the second objective in the thesis.

An appropriate format for this new guidance is a methodological framework. Although there is no formal definition of a methodological framework there is unspoken agreement that a methodological framework provides structured practical guidance, or a tool, to guide the user through a process, using phases, stages or a step-by-step approach (22, 28-35). Specific descriptions of a methodological framework have included: an 'algorithm' (36), 'practical guidance' (37), a 'practice based tool' (38),`a body of methods, rules and postulates employed by a particular procedure or set of procedures' (39), a 'set of structured principles', an approach for 'structuring how a given task is performed' (40), and a 'sequence of methods' (41). There are many benefits of using methodological frameworks: the standardised approach provided by a step-by-step guide can improve the consistency, robustness and reporting of the activity (42), enhancing the quality of the research (22), and maximising the trustworthiness of findings (38).

Methodological frameworks for conceptual modelling provide structure and direction for developing conceptual models (43), recommending methods and

good practices (44). They provide guidance for inexperienced modellers, whilst acting as an 'aide memoire' to more experienced modellers (15), and the documents created during the development of a conceptual model can form a helpful audit trail (15, 21).

Using a methodological framework as a format for the new guidance for developing conceptual models would direct the user through a series of standardised steps, providing structure for the user and improving the quality, consistency, robustness and validity of the conceptual model.

At the outset of this thesis there was no consensus on approaches which could be used for guiding the development of methodological frameworks (39). Due to this lack of guidance an extra work strand was added to the thesis with the aim of compiling practical suggestions for developing a methodological framework. This work strand is part of the second objective of this thesis; to propose a methodological framework for developing conceptual models.

In summary, the new role for conceptual models included in this thesis provides additional information beyond the conventional treatment arm-based economic evaluation for decision makers, particularly when evidence is uncertain, providing an understanding of the mechanism driving the economic evaluation results. In this thesis the new role for conceptual models is described and illustrated, and to support this proposed new role for conceptual models new guidance is proposed for their development.

1.3 Research question and aims

The research question posed and answered in this thesis is:

'How can conceptual modelling enhance health economic evaluation?'

The aim of the thesis is to expand the role of conceptual modelling in health economic evaluation.

To answer the research question two objectives were set, these were introduced in the rationale in Section 1.2:

- Propose and demonstrate (with illustrative case studies) a new role for conceptual models in health economic evaluations.

- Propose and demonstrate a methodological framework for developing conceptual models in this new role.

1.4 Thesis layout

An overview of layout of the thesis is provided below which is made up of seven further chapters.

Chapter 2

This chapter introduces the role of economic evaluation in decision making and describes the existing guidance for conducting economic evaluations alongside clinical trials. The practice of conducting economic evaluations alongside clinical trials is critiqued, and drawbacks are highlighted, in particular the uncertainty inherent in economic evaluation results which can lead to scant evidence for decision makers. An explanation of the 'black box' nature of economic evaluations is made. Finally, the concept of the new role for conceptual models in economic evaluation is explained, addressing both uncertainty in results and the lack of consideration to the mechanism driving the economic evaluation results.

Chapter 3

This chapter introduces the two cases studies used for illustration purposes throughout this thesis, both case studies are from clinical trials reporting uncertain results in the economic evaluation. The background and rationale of each clinical trial and clinical results are summarised in turn, then the methods used in the economic evaluation are explained, particularly in relation to the existing guidance described in Chapter 2, and finally the results of the economic evaluations are described.

Chapter 4

To comprehensively present the new role for conceptual models in economic evaluation advice is needed to create the guidance for developing conceptual models in this new role. At the outset of this thesis there was no available guidance for developing a methodological framework, and this chapter takes a step towards filling this gap in the form of practical suggestions. A scoping review identified reported approaches used in developing methodological frameworks; these approaches were then amalgamated and grouped into suggestions for developing methodological frameworks. This chapter partly fulfils the second objective of the thesis and has been published as a peer reviewed article (45).

Chapter 5

The suggestions from Chapter 4 for developing methodological frameworks are applied in this chapter to guide the development of the methodological framework for conceptual models. After identifying evidence from a scoping review a draft methodological framework was developed, which was evaluated by comparing it to the existing guidance for creating conceptual models in economic evaluation, then revised to produce the final methodological framework. This final methodological framework comprises seven discrete stages which are divided into three phases: I) Context, II) Development and III) Finalising. The methodological framework includes a diagram and a written document explaining the seven stages. This chapter fulfils the second objective of the thesis and has been prepared as a manuscript which will be submitted for publication.

Chapter 6

Chapter 6 demonstrates the application of the methodological framework presented in Chapter 5 using the case studies introduced in Chapter 3. These case studies are illustrative in nature with the aim of demonstrating how a conceptual model might be developed. Each case study shows the progression of the conceptual model development, concluding with the final conceptual model that is used in each conceptual model driven analysis in Chapter 7.

Chapter 7

This chapter uses the two conceptual models developed in Chapter 6, from the two case studies, to demonstrate how the new role for conceptual models could be implemented. A recap of the original treatment arm-based results is presented for each case study, followed by an illustrative application of the novel approach proposed in the new role in Chapter 2. This chapter fulfils the first objective of the thesis and a draft manuscript was presented and discussed at the Health Economists' Study Group in January 2021, with a view to submitting for publication.

Chapter 8

The final chapter of the thesis provides a discussion of the work and gives a conclusion to the thesis. First, each chapter is summarised in an overview, then the strengths and limitations of the work in the thesis are discussed, the policy and practice implications are considered and recommendations made, suggestions for future work are described, and finally an overall conclusion is presented.

In summary, the work in this thesis has led to three outputs, which are presented in the thesis in the following order:

1. Practical suggestions for developing a methodological framework

2. A methodological framework for developing conceptual models for health economic evaluations, with illustrative case studies

3. A new approach for the use of conceptual models in health economic evaluations, with illustrative case studies

Chapter 2 The role and practices of economic evaluation alongside clinical trials

2.1 Chapter overview

The aim of this chapter is to give the reader an overview of the role of economic evaluations in the decision-making process, present approaches recommended to conduct them alongside clinical trials and to discuss the strengths and limitations of using clinical trial data to inform economic evaluations, providing an overview of the current position of economic evaluation. The chapter aims to show why there is scope for a supplementary analysis in economic evaluation and presents a proposed new approach for this supplementary analysis.

The current role of health economic evaluations and how clinical trials are used to inform decisions in healthcare are discussed in Section 2.2. Existing guidance on conducting health economic evaluations alongside clinical trials is presented in Sections 2.3 to 2.7, and a critique of conducting economic evaluations alongside clinical trials highlights the strengths (Section 2.8), and limitations (Section 2.9) of this process. The consequences of these limitations underline the scope for an additional new approach to provide further evidence to decision makers and the new approach proposed in this thesis is presented in Section 2.10. The chapter is summarised in Section 2.11.

2.2 The current role of health economic evaluations alongside clinical trials

Healthcare resources such as medication, staff time, blood supply, donated organs or hospital facilities are limited, and as a consequence of these limited resources the demand for health resources outstrips supply (46). In healthcare systems around the world decisions have to be made on how best to allocate scarce resources when there is increasing demand and limits on budget, to meet the needs of populations and to achieve the efficient use of scarce healthcare resources (47). The reasons for this increasing demand include advances in technology, aging populations and higher chronic disease prevalence (48).

To make decisions on how to best allocate these scarce resources, options are ranked to allow the costs and benefits to society of alternative technologies to be assessed(49). This ranking is typically achieved by using either a welfarism or extra-welfarism framework. Both frameworks allow ranking of two or more options but differ in how they measure the benefits received from the changes in healthcare resulting from the technologies. Welfarism assumes that individuals want to maximise their utility (welfare) and are best placed to assess this utility and show this by making choices(50). In standard economics, inferences about the benefits received from these choices is achieved by observing individual choices made, this is termed 'revealed preference'. However, in healthcare, which is often publicly funded, choices cannot be studied, so 'stated preferences' are used; stated preferences assess hypothetical choices made by a sample of the public to obtain willingness-to-pay monetary amounts(49). However, some researchers are not comfortable with valuing health in monetary terms, leading to the extra-welfarism approach. Culyer introduced the concept of extra-welfarism in 1989, this measures benefits from changes in healthcare using a health state utility: the quality adjusted life-year (QALY)(49, 51). The QALY is a generic measure combining length and quality of life, it has the benefit of enabling comparison of different technologies and different disease areas.

However, evidence shows that society is not only concerned about maximising health; capabilities are important too (50). Capability relates to an individual's potential functioning based on choice and opportunity; the more choices and opportunities an individual has the more improved wellbeing, whether or not the individual chooses to make use of those opportunities (50, 52). Furthermore, when a technology is complex, or the system the technology is situated in is complex, restricting the measurement of benefits to health has the potential to underestimate the benefits of the health technology. In these complex interventions a broader approach is needed to assess the wider impacts of the technology, including benefits such as friendships, dignity, self-respect, spillover effects to family and community and multisectoral benefits, for example to the education and criminal sectors (50, 53).

The extra-welfarist approach of measuring benefits as QALYs has been widely adopted in health economics and this is the approach considered in this thesis.

The purpose of the research discipline 'Health Technology Assessment' is to provide evidence to inform decision makers about whether to adopt new technologies. In many countries Health Technology Assessment (HTA) organisations oversee evaluations, specifically for their jurisdiction, on clinical and cost-effectiveness for informing local decisions about adopting new technologies (54). For example, in Scotland the umbrella organisation Health Improvement Scotland (HIS) includes the Scottish Medicines Consortium (SMC) to oversee pharmaceutical decisions and the Scottish Health Technologies Group (SHTG) which oversees non-pharmaceutical decisions, and in England and Wales NICE has topic specific committees to oversee decisions on funding.

Clinical effectiveness evaluations compare the health outcomes of a new technology to the health outcomes of one or more existing technologies, or a placebo, to assess the size of the treatment effect. Cost-effectiveness evaluations investigate whether the health improvements resulting from the new technology represent value for money compared to the existing technology, or a placebo, by calculating the additional benefits per additional resources used (2). The cost-effectiveness component of a Health Technology Assessment is referred to as an 'economic evaluation' (1).

An economic evaluation comprises of two measurements: the difference in costs and the difference in health benefits of a new technology compared to an existing alternative (or placebo). The total costs and total health benefits accumulated by each participant in a clinical trial are estimated and averaged across all participants in the new technology group (t) and the existing technology group (or placebo) (c), resulting in a mean cost (C) and health benefit (E) per participant in each group. A summary measure for the cost-effectiveness of the new technology compared to the existing technology is calculated by applying the following equation (Equation 1), dividing the difference in costs between the technologies by the difference in health benefits (6).

$$\frac{Ct - Cc}{Et - Ec} = \frac{\Delta C}{\Delta E}$$

Equation 1: Incremental cost-effectiveness ratio

The product of this equation is referred to as an 'incremental cost-effectiveness ratio', typically shortened to the acronym 'ICER'. To evaluate whether the resultant ICER is cost-effective it is assessed against the local willingness-to-pay threshold, which is currently £20,000 to £30,000 in the United Kingdom (UK) (55); if the ICER is below the willingness-to-pay threshold it is considered cost-effective.

However, using the ICER to assess cost-effectiveness by simply comparing it to the threshold is not always straightforward; it can cause problems when the ICER is not a simple product of higher costs and improved health benefits, to give the ICER full interpretation it should be plotted onto a cost-effectiveness plane (Figure 1); the quadrant where the ICER sits explains the interpretation of cost-effectiveness. The vertical axis of the cost-effectiveness plane represents incremental costs, above the horizontal axis (incremental costs are positive) the new technology is more costly than the existing technology, below the horizontal axis (incremental costs are negative) the new technology is less costly than the existing technology. The horizontal axis of the cost-effectiveness plane represents incremental health benefits, to the right of the vertical axis (incremental health benefits are positive) the new technology has more health benefits than the existing technology, and to the left of the vertical axis (incremental health benefits are negative) the new technology has less health benefits than the existing technology. If the ICER is in the southeast quadrant the new technology is less costly and has more health benefits than the existing health technology, in this scenario the new technology is described as 'dominant' and should be adopted. If the ICER is in the northwest quadrant the new technology is more costly and has less health benefits that the existing technology, in this scenario the new technology is described as 'dominated' and should be rejected. If the ICER is in the southwest quadrant the decision on whether to accept or reject the new technology will depend on the decision-

making body's willingness to accept; there will be cost savings, but this comes at the price of losing health benefits and the willingness to accept decision is the amount of decrease in health benefits for a cost saving that decision makers are willing to accept. If the ICER is in the northeast quadrant the decision will depend on the willingness to pay threshold; the willingness to pay extra for a defined amount of additional health benefits. As previously discussed, if the ICER is below this threshold in the northeast quadrant the new technology should be adopted. As previously mentioned, relying on only the ICER is problematic; if the ICER is positive it can be in either the northeast or southwest quadrants where the decision is based either on willingness to pay to willingness to accept, if it is negative it could be in either the northwest or southeast quadrants where the decisions is to adopt or reject. If the ICER is plotted on the cost-effectiveness plane the decision is clear.

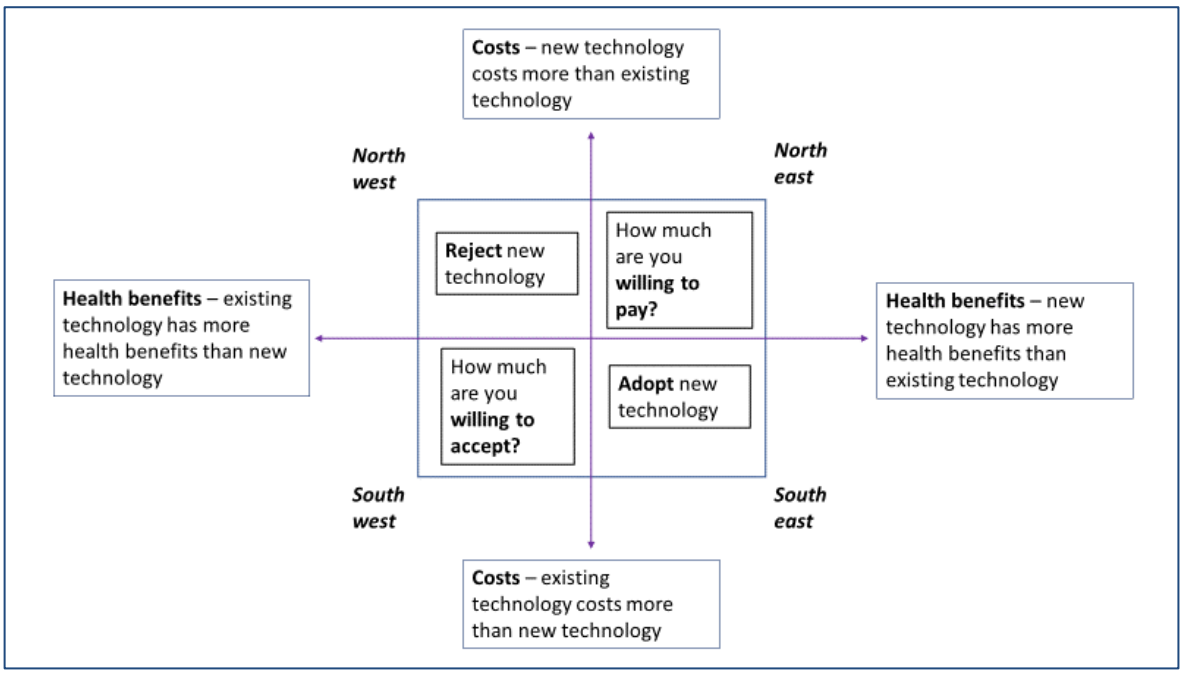


Figure 1: Cost-effectiveness plane interpretation

Evidence used to inform HTA policy decisions typically comes from clinical trials (4, 5), however it can also come from decision analytic models and routinely collected data (56). The focus of this thesis is economic evaluations alongside clinical trials and the evidence discussed in it will be clinical trial data; this data provides evidence on the safety, efficacy and clinical effectiveness of medicine, medical devices and procedures, and a rising number of clinical trials are also

collecting data for economic evaluations. The aforementioned growth in demand for healthcare has resulted in an increase in the number of economic evaluations (5), and over the past 30 years economic evaluation alongside clinical trials have become more popular, with increasing numbers being published (4). In a clinical trial two or more treatments (intervention and comparator) are compared to assess one or more outcomes agreed in advance and participants in each treatment arm are followed for a specified length of time to test the safety, efficacy and effectiveness of the intervention(s) (57). There are typically four phases in clinical trials: Phase 1 studies assesses the safe dose range and potential side effects of the intervention in a small number of patients; Phase 2 studies assess potential adverse events of the intervention in a larger group of people; Phase 3 studies assess the effectiveness of the intervention compared to a similar treatment or placebo in a large group of people, across different sites and countries, this Phase is typically the last step in assessing the intervention before it is either approved or not approved by different jurisdictions (although recently some decisions are being based on Phase 2 evidence (58)), and Phase 4 trials only take place after a new technology is approved and if required for further testing. For the clinical effectiveness and cost-effectiveness discussed in this thesis the clinical trial referred to is a Phase 3 study.

Clinical trials are considered the 'gold standard' for providing evidence on the efficacy and clinical effectiveness of a new health technology, assessing the causal relationship and association between the intervention and outcome of interest. The ability to examine causal relationships is possible because of the practice of randomisation of participants in the trials, this removes much of the bias seen in other study designs. Randomisation ensures that there is a balance in participant characteristics between the arms, this allows any difference in outcomes to be accredited to the effects of the intervention only, these effects are known as 'treatment effects'. Clinical trials are closely monitored to ensure that they are managed correctly, the management of clinical trials includes: participants recruited according to strict inclusion and exclusion criteria; the concealment of randomisation to trial arms; randomisation and blinding of allocation to trial arm where either the researcher, participant or both are not aware of the treatment they are randomised to; calculating an appropriate sample size to ensure the trial has enough power to determine reliable

treatment effect results; an intention-to-treat analysis is conducted, and all information regarding the clinical trial management should be pre-specified in a protocol. Whilst clinical trials have disadvantages such as expense, generalisability and dropouts, they provide the best available evidence on causality without the inherent bias of other study designs, for example observational studies (59).

Although clinical trials are considered the 'gold standard' vehicle for providing evidence on the efficacy and clinical effectiveness of a new health technology, and despite the increase in their use for economic evaluations, there have been doubts raised about their sole use in assessing the cost-effectiveness of a new health technology, resulting in a restricted analysis (60). The specific limitations of using a clinical trial as evidence for an economic evaluation are discussed in detail in Section 2.9.

The increase in the number of economic evaluations alongside clinical trials has led to a need for guidance for conducting robust economic evaluations, this need has been met by published recommendations for economic evaluation methods (4-6). This guidance is summarised and described in Section 2.3.

2.3 Current guidance for conducting economic evaluations alongside clinical trials

The guidance described in this section has been mainly taken from three sources: two journal papers – the ISPOR 2015 Good Research Practices Report (Ramsey et al.) (5) and Petrou et al. guidance (6), and the Glick et al. book on the subject (4). An earlier version of the ISPOR good research practices report was published in 2005 in response to a growing number of clinical trials which included health economic endpoints (61), this was updated in 2015.

The economic evaluation guidance is presented in four sections: Section 2.4 describes the initial step of designing economic evaluations alongside clinical trials; Section ▫2.5 describes methods of data collection, Section 2.6 describes analysis of the data once it has been collected and Section 2.7 describes the recommended layout for reporting an economic evaluation.

Chapter 2

2.4 Current guidance - design of trial based economic evaluations

The intrinsic features of a clinical trial design will feed directly into the quality of the economic data collected in the trial, so it is important that the health economist is involved in the design of the study and collaborates closely with the clinical trialist to ensure that the best quality economic data is collected (5, 6). Details of this collaboration should be recorded in the standard operating procedures of the clinical trials unit (6).

The health economist should assess whether the clinical trial design is a suitable vehicle for conducting an economic evaluation, it may be the case that the clinical trial design is an exploratory non-comparative Phase 1 or 2 trial, or otherwise inappropriate (5). To assess the suitability of the clinical trial data the underlying principles of economic evaluations should also be considered, these include:

- The sample size should be large enough to reliably detect differences in economic outcomes (62).

- All relevant evidence should be included in the economic evaluation, this includes clinical effectiveness of the intervention, relevant resource use and health benefits (3).

- Comparators should reflect all available treatments routinely used in the National Health Service (NHS) and those considered best practice (3, 63).

- The economic resource use should be representative of usual clinical practice (5, 6).

- The preferred outcome measure to assess health benefit is one that is estimated using a preference-based measure (63), if this is not available or not considered sensitive enough for the disease in question then the outcome should reflect an episode of care which then triggers a cost (for example a stroke) (4). However, in clinical trials surrogate endpoints are often used, these allow a reduced sample size and/or trial duration,

reducing costs and speeding up the decision-making progress (64, 65). The surrogate endpoint should be associated with the final endpoint such that all treatment effect on the final endpoint is captured by the surrogate endpoint (65). NICE has made decisions based on economic evaluations using surrogate endpoints, however it has been shown that the final endpoint may show no clinical or cost-effectiveness once more evidence is collected (64). To avoid subsequent evidence on final endpoints reversing treatment effect sizes and direction based on surrogate endpoints, robust validation of the surrogate endpoint should be conducted. This comprises assessing the level of evidence and strength of association between the surrogate and final endpoints, and quantifying the relationship between them, preferably in terms of QALYs (64).

- A sufficient follow-up period is required to capture all costs and health benefits relating to the comparators to estimate cost-effectiveness, preferably over a lifetime (3, 5, 63). If there are strong links between an intermediate end point in the clinical trial and a long-term disease episode of care, then more reliance can be given to these intermediate end points. The length of time between data collection time points is also important, care should be taken not to leave too long between data collection time points; this enables the participant to remember exactly what resources were used in the previous period.

- Finally, uncertainty in the results should be characterised, in particular decision uncertainty which explores the probability that the decision based on the available evidence is the correct one (3).

2.5 Current guidance - data collection

As discussed previously in this chapter, the two main components of an economic evaluation are costs and health benefits. Costs are a combination of the resources used by participants during and prior to the clinical trial (for example nurse visits or inpatient hospital stays) and the relevant unit cost for each resource used. For health benefits this should be a preference-based outcome, but may be restricted to disease specific clinical outcomes, or a

mixture of both. Careful consideration should go into the collection of economic data (5).

2.5.1 Resource use and costs - identification

Typically calculating costs requires two steps; the first ascertains relevant resource use categories and collects these identified resources used by trial participants, and the second applies unit costs to these resources to estimate total costs. It is possible but less common to have access to routine resource data collection that has costs already included in the economic data. The typical categories of resource use collected in a clinical trial include: the treatments; healthcare use related to the disease or treatment; treatment of side effects related to the intervention, and other resource use deemed important (6). The economic evaluation guidance recommends several techniques which can be used to identify relevant resource use, these are discussed below.

The perspective of the economic evaluation will help to determine the resource use categories that should be included; a healthcare perspective (also known as payer perspective) will include use of health and social services as already described, this is the basic perspective required in UK by NICE (63). A more broad societal perspective also includes resource use incurred outside of health and social care services and includes personal costs, informal care by friends and family and productivity losses from being unable to work (6). Another method for ascertaining relevant resource use is to conduct an assessment of the typical resource use categories in a care pathway to identify key resource categories (4, 5), in addition to administrative data and patient logs.

Generally, the economic evaluation guidance recommends the collection of data use categories that should affect the results of the economic evaluation (5). In particular, recommending the identification of two types of resource uses: those most likely to differ between arms (cost drivers), but not necessarily linked to the disease or intervention, and high value resources. The first type (resource use that may differ between arms) estimates the intervention's effect on costs, and the second estimates overall variability on costs (4, 5). Other resource use

categories recommended for collection are those which are used by many participants (62).

Suggestions for reducing the burden of resource use collection include restricting the resource use categories to those related to the disease or intervention, those provided by the study site(s), and reducing the number of participants who provide resource use data, however a balance should be struck between minimising the burden and undermining the results from the economic evaluation (4). The general view is that including a restricted range of resource use categories should be avoided as there may be unexpected consequences of the intervention (4, 5).

Ideally the unit costs attached to the resource use should represent the opportunity cost of that resource; the value of the next best alternative. However for practical reasons, site- or nation-specific unit costs are more frequently used, the most appropriate unit cost to use will depend on the research question (4, 6). Tariffs are one alternative to using opportunity costs, these are set by governments and either relate to a diagnosis related group or health resource group payments, the benefits of using these costs are that they represent the true cost of what is spent by governments (4, 6), however the costs for different procedures or stays are often put in wide categories that do not accurately reflect the actual costs. Another alternative to opportunity costs are site-specific unit costs, micro-costing techniques are needed to calculate these, and the benefit of using site-specific unit costs is that an accurate cost for each site can be calculated and applied to resource use (4). Micro-costing also has the benefit of reflecting the actual resources used and therefore displaced similar to the concept of opportunity cost, and could arguably more accurately measure differences between treatment arms (66). Whatever type of unit cost is used to value resource use, it should always be for the same price year, inflated using healthcare price indices if necessary. In multinational trials costs should be converted to the same currency using purchasing power parity techniques (6).

2.5.2 Resource use and costs - data collection options

Conventional collection of participant data during clinical trials is through the case report form which is completed by all trial participants at their follow-up visits (67), it is common for economic resource use data to be collected using the same case report form that is used to collect other relevant (non-economic evaluation) data during the clinical trial. However, other instruments which have been designed to collect data directly from participants (or via a carer or proxy) during a clinical trial may be used, these including patient questionnaires and diaries. It is also possible to include resource use which is not collected directly from participants during the clinical trial, this is known as routinely collected data, for example patient medical records (ie GP or hospital) (5, 6). Case report forms are often adapted specifically for each trial which can lead to variability in the data collected, therefore best practice is to use validated resource collection instruments (5, 68, 69) and instruments for calculating productivity costs (70-72). If it is not possible to use validated resource collection instruments then any new versions of forms or instruments used to collect economic data should be tested and evaluated for suitability of use (6).

Recall bias can affect the completeness and accuracy of resource use collected directly from participants, this occurs when the participant is unable to accurately remember the healthcare resources that they have used since the last follow-up (6). It is often the case that, for practicality, the collection of economic data coincides with the clinical trial follow-up visits, and if the timing is unlikely to result in recall bias this is deemed suitable for the economic evaluation. However, if there is likely to be a lot of resources used by trial participants during the follow-up period, then using a case report form which is completed at set trial follow-up points may not be suitable, in this case a diary may be more practical in helping the participant to record resource use more frequently, and results in resource use data collection not subject to recall bias (5). Another way of minimising recall bias is to verify the resource use directly reported by participants using secondary sources (5).

Using electronic databases for collecting clinical trial data is becoming increasingly popular, this enables a simplified method of clinical and economic data collection for sites; data is input into an electronic case report form which

is uploaded to the trial database. The advantages of these databases are that data can be checked for accuracy and completeness throughout the time period of the trial and any data issues spotted by trial staff monitoring the database can be highlighted before the end of the trial and queries can sent to trial site staff (5). More recent additions to the electronic trial database are the ability to use electronic health care records data to partially populate the trial database, and using the internet, smartphones and mobile health applications to upload data (5). The main benefit of these electronic methods is an improvement in efficiency, however further research is needed to understand the quality and completeness of the resulting trial databases (5). Whatever process is chosen to collect economic data it should be piloted prior to use (6).

2.5.3 Outcomes measures

When conducting an economic evaluation it is best practice for the health benefit measure included to allow for comparisons across diseases and interventions, and not be limited to the disease and intervention of the specific clinical trial (4). As discussed previously, the favoured outcome measures for assessing health benefits in an economic evaluation are preference-based quality of life measures, however the health benefit may also be measured in disease specific clinical outcomes such as blood pressure, cholesterol or strokes avoided, if possible alongside the preference-based measure or, if it is not possible to include a preference-based measure, as the sole measure of health benefit (4).

Preference-based quality of life measures combine a health utility score (quality of life) with a measure of length of life to calculate a QALY. The QALY is the preferred health benefit outcome in many jurisdictions as it can be used to compare cost-effectiveness across different disease areas, as recommended in the guidance. The value of the health utility is on a scale from 0 to 1, where 0 represents death and 1 represents full health. For example, in combining health utility and length of life, a QALY of 0.6 signifies that the health state the participant is in is worth 0.6 years at full health or one year at less than full health (4). However using the QALY may not pick up differences in health related quality of life in an economic evaluation due to its restrictive and insensitive nature, it has been reported that QALYs are not responsive to changes in disease states (4, 6).

There are two main preference-based techniques for calculating QALYs, the first uses a 'pre-scored' questionnaire asking participants about their health state. Each participant completes a questionnaire (for example EuroQol-5 Dimension (EQ-5D), Health Utilities Index or SF-6D) (73-76) at pre-determined time points during the clinical trial follow-up. The responses to these questionnaires are valued using existing population-based value sets using population preferences, the value sets can be country specific. There are quality of life questionnaires specifically for children (for example EQ-5D-Y and HUI2) (75, 77, 78), and for seriously ill participants or those with cognitive impairment proxies can be used (79). The second technique for assessing quality of life is using preferences obtained directly from participants using techniques such as time trade-off or standard gamble (79, 80), these ask each participant about their health status and then to value it, however these latter techniques can be time consuming and expensive and therefore not suitable for use in many trial based economic evaluations which are constrained by time and costs.

Where preference based quality of life measures are not collected from participants in the trial, there is the option to map non-preference based responses from the trial to preference-based measures (81), however this technique has limitations due to few mapping algorithms being validated and issues with the reliability of results (5).

To calculate QALYs over the trial period the utility scores at each follow-up time point are combined with the length of time between follow-ups using area-under-the-curve techniques (4).

The frequency of quality of life data collection will depend on disease severity, disease progression and the perceived burden on participants, if relevant and for convenience the trial follow-up points used for collecting clinical outcomes and resource use data, can also be used to collect the quality of life data (6).

Chapter 2

2.6 Current guidance - analysis

2.6.1 Analysis plan

Each economic evaluation should be planned in advance and a written analysis plan should be produced and agreed before the trial data is unblinded (5, 82). Key approaches common to all economic evaluations include: 1) the analysis should use an intention-to-treat population where the analysis is based on the treatment arms that participants were randomised to regardless of which treatment they received; 2) the within trial analysis should use the same time horizon for costs and outcomes; 3) uncertainty around each result should be included; 4) if a trial follow-up is longer than a year then a discount rate should be applied to costs and outcomes for all years after the first year, this adjusts for time preference, (for England and Wales the recommended discount rate is 3.5%), and 5) if there is missing data, techniques should be employed to deal with this (4-6, 63).

2.6.2 Costs

The appropriate cost measure for economic evaluations is the arithmetic mean – an average cost per person, this allows decision makers to estimate total population costs for the intervention. Costs are often right skewed; most participants incur low or no costs while a small number will incur high costs. To estimate the difference in mean costs between arms, if there is a large sample it can be assumed that the costs are normally distributed and parametric techniques can be used, however generally best practice is to use a generalised linear model to analyse skewed data, using a gamma family and log link (4-6).

2.6.3 Outcomes

For health benefits best practice is to replicate the primary clinical analysis methods for any clinical outcomes included in the economic evaluation, however the primary clinical trial outcome may be presented as a time to event while the economic evaluation will include all clinical outcome events. Recommended methods for analysing QALYs include adjusting for baseline characteristics of participants and baseline EQ-5D (83), and, as with the cost analysis, a

generalised linear model is a suitable method to estimate the differences in QALY, using a Gauss family and log link (5, 6, 83).

2.6.4 Missing data

There is an inherent problem with missing data in economic evaluations alongside clinical trials (6), there are two main types of missing data and the techniques for dealing with these types of missing data are explained next.

The first type of missing data occurs when participants do not report or record all the data that they are asked for. The trial team can help reduce the amount of missing data by auditing the trial database, and health economists can help by reminding trial researchers the importance of minimising missing data. Naive methods of dealing with missing data are presenting complete case and full sample (available case) analyses, although these methods are not generally recommended. However if the amount of missing data is less than 5% of the observations, and the quantity and pattern of missing data do not differ between treatment arms, a complete case analysis is permissible (5). A more suitable approach for dealing with missing data is the use of multiple imputation techniques, the exact multiple imputation technique used will depend on the type of missing data, if the data is missing at random the multiple imputation using chained equations is recommended (84).

The second type of missing data is censored data, which is when a participant drops out of the trial and the data on them is restricted to their time in the clinical trial, suitable methods for dealing with censored data in an economic evaluation are available for use (5, 6). The health economists should work with the trial team to minimise the amount of censored data, and similar to missing data, if around 5-10% of the total data is censored it is not believed that it will have an impact on the results so not further methods are required. Again complete case or full sample analyses are naïve methods often applied, however more suitable methods are available, these include: Lin 1997 method, Lin 2000 methods, Bang and Tsiatis method, Carides methods and multiple imputation (4).

2.6.5 Summary measures

The difference in costs and health benefits between arms should be summarised in a measure, this summary measure provides cost-effectiveness evidence to the decision maker. There are three measures commonly used to demonstrate the value of the interventions (5):

1. The ICER described in Section 2.2 where the cost of gaining or losing one QALY, or other health benefit, is presented, and is often the main summary measure reported in an economic evaluation. However, the nature of ratios can prove challenging; the difference in health benefit (either QALY or alternative clinical outcome) may be near to zero which results in a large ICER, and negative ICERs points (represented in the northwest and/or southeast quadrants of the cost-effectiveness plane) can be the result of higher costs and outcomes lower outcomes, or lower costs and higher outcomes in the new technology arm compared to the existing technology arm. Furthermore, negative ICERs make confidence intervals difficult to calculate and interpret (6), the net benefit measure overcomes these issues and is discussed next.

2. The net monetary benefit (NMB) measure incorporates the healthcare payer's willingness-to-pay threshold (λ) with incremental costs (ΔC) and incremental QALYs (ΔE) (Equation 2):

$$Net\ monetary\ benefit = \lambda.\Delta E - \Delta C > 0$$

Equation 2: Net monetary benefit

If the net monetary benefit measure is positive it should be interpreted as the new technology being cost-effective and it should be adopted, if the net monetary benefit it is negative then the new technology is not cost-effective and should be rejected; the costs outweigh the value assigned to the health benefits (6).

3. The third summary measure is a measure of probability, the method for estimating this summary measure is to present net monetary benefit

results on a cost-effectiveness acceptability curve using different levels of willingness-to-pay thresholds (2, 6). This shows the probability of the new technology being cost-effective at different levels of willingness-to-pay thresholds (5).

2.6.6 Uncertainty

When there is uncertainty around the economic evaluation results this should be characterised and presented to decision makers as evidence for basing their funding decision on. There are four key types of uncertainty that can occur in the results: sampling, parameter, imputation and heterogeneity uncertainty.

Sampling uncertainty occurs when a sample of people is extracted from the population of interest and an analysis is based on this limited pool of people who may not be representative of the entire population (5). Guidance includes four recommendations for characterising this uncertainty: 1) Measures of variability should be reported for mean costs and mean health benefits for each arm of the trial, as well as for the differences between arms for mean costs and health benefits, and summary measures such as the ICER and net benefit (5). The measure of variability most useful when estimating differences is the 95% confidence interval. 2) Results can also be shown on a cost-effectiveness plane using non-parametric bootstrapped samples, this has the benefit of avoiding complications from using a ratio, and uncertainty can be represented on the plane by including confidence ellipses (4). 3) Sampling uncertainty can also be presented on a cost-effectiveness acceptability curve using the bootstrapped samples and varying the levels of the willingness-to-pay threshold, this was discussed above in relation to summary measures (2, 5). 4) Using the value of information measure; the expected value of perfect information (EVPI) can be calculated to inform decision makers on the value of removing uncertainty from the data ; the EVPI represents difference between the value of a decision based on current information and the value of a decision based on perfect information (85). The cost of acquiring perfect information is based on the probability of making the wrong decision when deciding whether to adopt or reject the intervention, and the cost of making the wrong decision (4, 5). If the population EVPI is greater than the cost of further investigation in a future trial or other research, this this is a necessary, but not sufficient, condition to conduct further

research; it is potentially worthwhile. To establish whether future research is worthwhile the expected value of sample information (EVSI) should be calculated (85). The EVSI measures how much the uncertainty, and related consequences of this uncertainty, are diminished by evidence from additional research. This is a result of the design of the future research, for example sample size, follow-up and endpoints, all these factors affect the cost of the research (86). The difference between the population EVSI and the expected cost of the research is termed the expected net benefit of sampling (ENBS) and should be calculated for a range of study designs. If the ENBS is positive this suggests that further research is worthwhile, this is a sufficient condition, if the ENBS is negative further research is not worthwhile (86). EBNS is useful for determining the study design of future research; the design with the highest ENBS should be chosen (87).

Parameter uncertainty is the uncertainty in the parameter estimates due to uncertainties in the economic data (5). There are two recommendations for dealing with parameter uncertainty in economic evaluations: 1) Sensitivity analysis, where uncertain parameters are assessed for their impact on the results by varying these parameters within probable ranges. 2) Another technique is to use a value of information approach similar to EVPI to calculate the expected value of perfect parameter information (EVPPI); this approach focusses on calculating the value of gaining perfect information on specific parameters.

Imputation uncertainty occurs when imputing missing values artificially shrinks sampling uncertainty estimates, a suggestion for tackling this type of uncertainty is to bootstrap the results, the results of this technique can be used to create a cost-effectiveness plane and a cost-effectiveness acceptability curve (5).

Heterogeneity uncertainty is driven by differences (heterogeneity) between the trial participants, this can be investigated using regression techniques applied to the results, using characteristics of participants as covariates, and by sub-group analyses (6).

2.6.7 Extending the analysis beyond the trial time horizon

The economic evaluation should capture all costs and health benefits related to the intervention, this is often not possible within the timescale of a clinical trial, in these circumstances an extrapolation of trial results is needed. Life expectancy can be estimated using survival analysis techniques such as Cox proportional hazards or Weibull models to extrapolate survival data (6), and if the trial period is long enough to capture sufficient data for modelling then direct modelling can be used, if this is not feasible the trial data can be combined with longer-term observational data for conducting decision analytic modelling (5).

2.7 Current guidance - reporting

This section briefly describes the recommendations for reporting economic evaluation results, this is included for completeness only as the main focus of this chapter is the analysis of the economic evaluation.

Guidance is available for reporting economic evaluations alongside clinical trials (88-91). The audience for economic evaluation is often varied so detailed reporting of the analysis is required (5).

2.7.1 Trial information

The economic evaluation report should begin with clinical trial related information, such as a description of the clinical trial, the setting and location, the inclusion and exclusion criteria for participants, a description of the treatments, details of the protocol driven treatments, the clinical trial time horizon, participants' baseline characteristics and a link to the registry of the trial. Following this overview, a summary of the clinical results should be included (5).

2.7.2 Economic data

An overall description of the economic data used in the economic evaluation should be reported, including the economic data collected (costs and outcomes),

a timetable of what data was collected when, unit cost sources and how much missing and censored data there is (5).

2.7.3 Methods

The methods section should describe the component parts of the costs and outcomes and how they were calculated, if the outcome is the clinical endpoint and differs to the clinical analysis results an explanation of how this difference occurred should be included. Planned methods for dealing with missing and censored data should be included as well as planned statistical methods for the comparison of resource use, costs and outcomes. If the time horizon is to be extended beyond the follow-up period of the clinical trial the methods and assumptions planned to do this should be reported. Finally, any deviations from the analysis plan should be reported and justified (5).

2.7.4 Results

The point estimates and corresponding measures of uncertainty should be reported for resource use, costs and outcomes. Results for the within trial time horizon and any longer time horizons (if appropriate) should be presented, and any results not suited to table presentation should be presented as graphs and other suitable formats (5).

2.8 Strengths of conducting economic evaluations alongside clinical trials

As discussed in Section 2.2 clinical trials are considered to be the gold standard vehicle for assessing safety, efficacy and effectiveness of an intervention, and necessary to gain a licence for a new pharmaceutical product, or an existing pharmaceutical product used for a new indication (1). One of the main strengths of a clinical trial is that any confounding issues experienced in an observational study are removed by applying the system of randomisation, this gives an unbiased estimate of the size of benefit or risk from using the new technology compared to the existing technology, thus clinical trials are deemed to be the best vehicle for estimating the relative treatment effect between the intervention and comparator (55, 92).

Many funders are calling for economic evaluations to be included in the design of clinical trials to add cost-effectiveness evidence as well clinical evidence for decision makers (6). Economic evaluations conducted alongside clinical trials benefit from the robust clinical trial design, unbiased estimates of effectiveness and from the trial systems for collecting data, giving a unique chance to collected individual resource use data (62, 93). Having this individual level data gives the opportunity for robust statistical analysis (1, 62) and provides valuable information on resource use, health related quality of life and the disease being researched (92). They also provide additional information on the cost-effectiveness of an intervention with high internal validity, and with external validity if the clinical trial is designed and conducted correctly (1, 5, 62).

There are also potential financial benefits from conducting an economic evaluation alongside a clinical trial, as it gives a practical opportunity to conduct a cost-effectiveness analysis, which provides reliable evidence with little extra cost on top the existing large, fixed cost of the trial. Conducting clinical trials is expensive and using the trial data to conduct an economic evaluation adds small marginal costs for arguably a large analytical gain (1, 6, 62).

2.9 Limitations of conducting economic evaluations alongside clinical trials

Despite these strengths of using a clinical trial as a vehicle to conduct an economic evaluation there are inherent weaknesses too, which despite the robust guidance set out in 2.3 to 2.7 can undermine the results of the economic evaluation. Whilst it is agreed that a clinical trial is the gold standard vehicle for providing clinical effectiveness evidence, it is also agreed that clinical trials are not gold standard for providing cost-effectiveness evidence as discussed by Sculpher at al. in 2006 (60).

The objectives of a clinical trial (an unbiased protocol driven assessment of treatment safety, efficacy and effectiveness) are inconsistent with the objectives of an economic evaluation (estimation of the costs and health benefits of a intervention provided in real clinical practice to all, not carefully selected patients, compared to current best practice, over a suitable follow up period) (62, 63, 94). The audience for these two analyses is also different; the

clinical evaluation from a clinical trial informs regulatory licensing and clinical decision makers, whereas economic evaluations inform healthcare policy makers and payers (94).

These contrasting objectives and audiences lead to inherent shortcomings when conducting economic evaluations alongside clinical trials, which challenge the underlying principles of conducting an economic evaluation (set out in Section 2.4) as a result of the necessary features of a clinical trial (5, 6). The main conflicts between the economic principles and features of a clinical trial are described in detail in the following sub-sections.

2.9.1 Protocol driven care

The care provided to participants in clinical trials is driven by the trial protocol not clinical practice, so is not always representative of resource use in 'real-life'. Blinding of participants and/or researchers exacerbates this additional resource use with participants in the control arm potentially being given tests and other procedures that are only needed in the intervention arm, which can potentially underestimate any true difference in costs. More frequent monitoring can lead to 'case findings'; which is when an undiagnosed condition is discovered during a protocol driven visit or test, and, as it is almost impossible to know whether this condition would have been diagnosed in the absence of the protocol driven visit or test, it is unclear whether resources used in diagnosing or treating should be included in the economic evaluation (5, 62). In protocol-driven care participants are encouraged to attend visits and comply with medication, in a real-life setting compliance is often lower, leading to an artificially high resource use in the economic evaluation (62), this can lead to problems with external validity. There are three recommended approaches for mitigating these issues; 1) the economic evaluation can omit resource use resulting solely from the influence of the trial protocol; 2) the trial team could run a pragmatic trial where the care is based on clinical practice; and 3) the trial team could include a usual care arm, either as part of the trial or run parallel to the trial (62).

2.9.2 Single comparator arm, which may not be current best practice

The comparator chosen in a clinical trial is often a placebo, therefore the estimated treatment effect from the clinical trial is the maximum possible treatment effect and may overestimate the effectiveness of the new technology. A key principle of health economic evaluations is to compare the new technology to all treatments currently being routinely used, especially those considered to be current best practice, so a clinical trial comparing the new technology to usual care is considered more appropriate than one comparing the new technology to a placebo (62, 92). The recommended approach for mitigating this limitation is to conduct an analysis using a decision analytic model which includes all possible and existing comparators used in current practice (3).

2.9.3 Unrepresentative study sites

Study sites included in the clinical trial may not be representative of the sites that would provide the intervention if it is adopted; study sites are often chosen on their ability to recruit participants only. This misrepresentation can lead to a variation in clinical practice between study sites, and the participants from the study sites may differ to those at other sites too, for example severe morbidity compared to mild. This can result in resource use at study sites not representing overall clinical practice and undermining external validity, this problem can be overcome by including several study sites with different characteristics (62).

2.9.4 Strict inclusion and exclusion criteria

Strict inclusion and exclusion criteria are necessary in exploratory trials to minimise variation between participants and, as accurately as possible estimate the treatment effect, this can lead to the trial participants not being representative of the target population, undermining external validity (5, 6, 62). One recommendation for dealing with this weakness is to use a pragmatic trial design; pragmatic trials enrol participants who are representative of all patients, the new technology is compared to current practice, and follow-up is carried out under typical clinical practice routines, however internal validity for pragmatic trials can be low (62, 94).

2.9.5 Clinical endpoints do not match economic evaluation endpoints

The clinical end points used in a clinical trial are often not suited to economic evaluations, which use an episode of care as an endpoint which then triggers a cost (4). There are typically three types of clinical endpoints used, each of which are discussed in terms of their suitability as an economic endpoint next: 1) composite endpoints are when several endpoints are combined to give greater statistical power to the results, for example all-cause mortality, recurrent myocardial infarction and new congestive heart failure (3). The cost per composite endpoint is not suitable for an economic evaluation, mainly because the endpoints do not have the same importance, so when composite endpoints are used in a clinical trial it is recommended that they are presented separately and are common across different diseases and treatments for comparison (5). 2) Intermediate or surrogate endpoints are used primarily when the expense of a longer follow-up in the clinical trial is prohibitive. An intermediate or surrogate endpoint is not usually an episode of care, an example a percentage decrease in blood pressure. The time horizon of the clinical trial is usually based on assessing the clinical effectiveness of the intervention, which when the endpoint is an intermediate one is not always long enough to capture all important health economic resources and outcomes related to the disease for all comparators, a key principle for conducting economic evaluations (92). In these cases, the intermediate clinical endpoint should be linked to long-term costs and outcomes, and if a link is not available the health economist should push for longer follow-up to allow for the collection of a suitable economic outcome. When deciding on a suitable length of follow-up for economic endpoints this may be longer than the clinical endpoint, for example a clinical endpoint may be a blood pressure reading whereas the related economic endpoint will be an episode of care triggering a cost, for example number of strokes linked to an increase or decrease in blood pressure. 3) The third type of clinical endpoint is aimed at assessing how a participant feels, functions, or survives, these are the preferred endpoints for economic evaluations. Preference-based quality of life scores can be combined with survival data to produce a QALY, the preferred economic evaluation outcome in many jurisdictions (5).

A suggested approach to deal with this limitation is to use a decision analytic model and epidemiologic data to predict final outcomes from intermediate outcomes (1, 62).

2.9.6 Censored data

Censored data reduces the economic evidence, either as a result of the participants being lost to follow up or due to data collection stopping at a pre-determined point. Approaches recommended to mitigate this issue are to use modelling techniques to extrapolate trial results, to predict final economic outcomes using clinical intermediate outcomes, and to analyse censored data using techniques suggested in sub-section 2.6.4. (62).

2.9.7 Insufficient follow-up

A key principle in economic evaluation is using an appropriate length of follow-up to capture all costs and health benefits to estimate cost-effectiveness, preferably over a lifetime, however often this is limited to a much shorter period (5). If there are strong links between an intermediate end point and a long-term disease episode of care more reliance can be given to these intermediate end points. Another approach for dealing with insufficient follow-up is to extrapolate the clinical trial results using modelling techniques.

2.9.8 All relevant evidence not collected

Another key principle is based on the requirement to include all relevant evidence in the economic evaluation, however using a clinical trial as a vehicle for economic evaluation invariably narrows the range of evidence available on health related quality of life and resource use (92). A suggested approach to mitigate this limitation is the use of systematic reviews and evidence to inform a decision analytic model (92).

2.9.9 Insufficient sample size

One of the key issues of using clinical trials for economic evaluations is sample size; a sample size (power) calculation informs the investigators on how many participants are needed to reach a significant conclusion on the size of

treatment effect from the intervention in the clinical trial (62, 95). If too few participants are recruited the clinical trial will not be able to evaluate effectiveness reliably, and if too many participants are recruited the clinical trial becomes too expensive and hard to run. The power calculation ensures the correct number of participants are recruited to the trial to answer the research question (95). To determine the power calculation the null hypothesis, the type I error, and the type II error should be defined. The null hypothesis is 'there is no difference between treatment A and treatment B', a type I error is the probability of rejecting the null hypothesis when it is true and is typically set at two sided 0.05, and a type II error is the probability of not rejecting the null hypothesis when it is false which is typically set high at 0.80, the higher this is set the larger the sample size. The clinically acceptable margin of difference (treatment effect size) should also be estimated, either from a pilot study or extracted from the literature (96). The type of clinical trial design is also needed to calculate the sample size; designs comprise superiority, equivalence, and non-inferiority. Superiority trials assess whether one treatment is more effective than the other either in statistical or clinical terms, equivalence trials assess whether the treatments are equally effective, and a non-inferiority trial assesses whether one treatment is as effective as the other within a previously set margin. Finally, the effectiveness outcome is defined as either dichotomous or continuous (95).

When an economic evaluation is conducted alongside a clinical trial this often results in the underpowering of the economic evaluation (62). The clinical outcome often requires a smaller sample size than the economic outcomes due to a large variability in resource use and costs relating to the economic evaluation, leaving clinical trials underpowered to detect economic differences (5, 6). Even if it was possible (financially and ethically) to recruit additional participants so that the economic evaluation is not underpowered, there is no agreed definition of an economically meaningful difference for an economic outcome sample size calculation, and estimating the joint distribution of the difference in costs and health benefits between treatment arms is a complex issue (6, 62). Economic evaluations are not powered to test a hypothesis and instead focus on estimating the differences in costs and health benefits, and the likelihood that an intervention is cost-effective (6, 97). Several approaches have

been proposed to address the limitations of an inadequate sample size, these relate to the statistical analysis of the economic evaluation: estimating uncertainty around the ICER, presenting a cost-effectiveness plane and using the net monetary benefit summary measure (62).

2.10 An introduction to the new role for conceptual models

This section introduces the new role for conceptual models in economic evaluation. It begins with an overview of the role and explanation of why it is needed, then presents suggestions for how the new role will work, including the contents of the conceptual model, how the expected relationships within the conceptual model can be assessed for accuracy, and how the conceptual model could be used for an additional analysis. The section finishes with a discussion of the new role.

2.10.1 Overview

In conventional health economic evaluations alongside clinical trials, study data are analysed by comparing the costs and outcomes of each treatment arm to assess the cost-effectiveness of a new technology, or for the new use of an existing technology. However, while clinical trials are essential for evaluating the clinical safety, efficacy and effectiveness of an intervention they do not always provide economic data that matches the key principles of economic evaluation. Due to not always achieving these key principles it is common for the results of treatment-arm based analyses to be uncertain (reporting no significant difference in costs or outcomes between arms), particularly as a result of underpowering of the economic evaluation, the suboptimal collection of economic data and protocol-driven costs (98). Because of this inherent uncertainty in economic evaluation results it is recommended that the focus in the results should be estimation and not hypothesis testing (99). The role of economic evaluations is to provide cost-effectiveness evidence to decision makers for allocating healthcare resources, when this evidence is uncertain the decision makers' role is made harder, but a decision must still be made.

Methods have been developed to represent this uncertainty in results (62, 94) and are reported in sub-section 2.6.6, they comprise presenting confidence intervals around the results, producing a cost-effectiveness plane, providing a cost-effectiveness acceptability curve, and presenting value of information estimates. When uncertain results are plotted on a cost-effectiveness plane, the plots will straddle the axis' and crowd around the origin, this is described as 'Scenario 9' (Figure 2) as depicted by Briggs and O'Brien (100), where there are no differences between arms in either costs or health benefits.

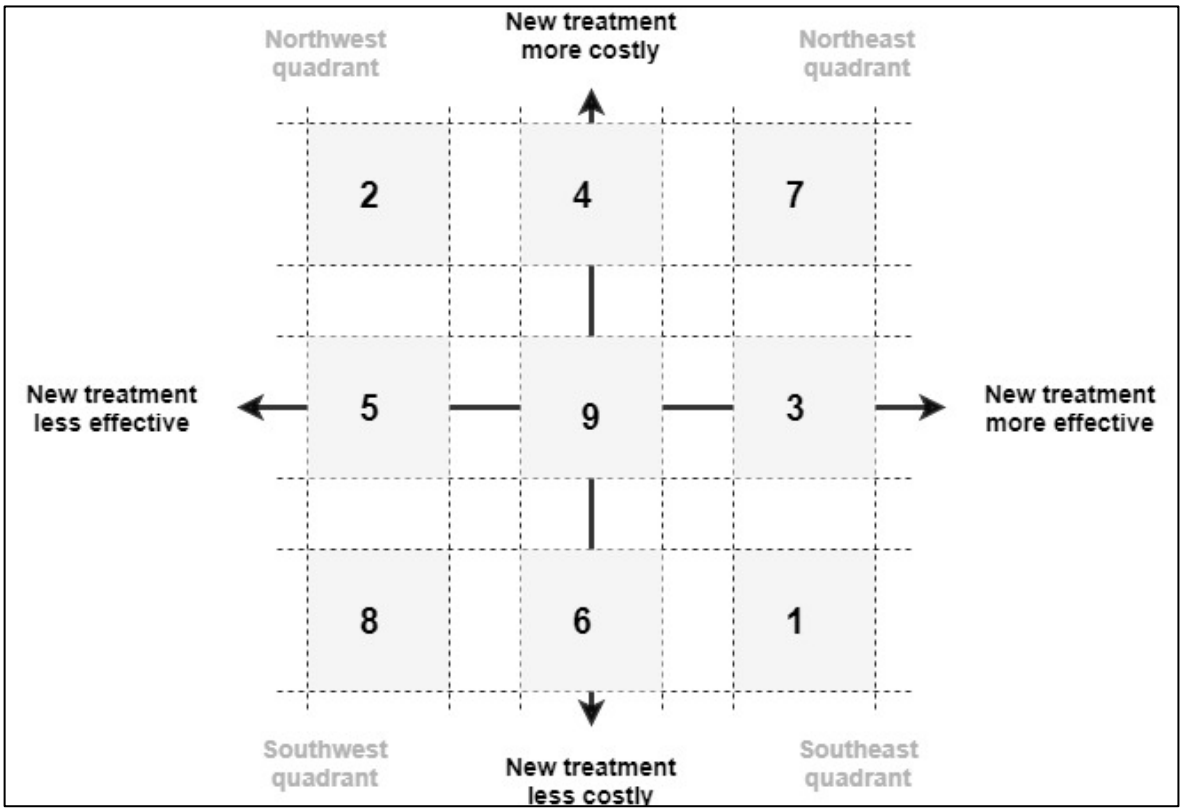


Figure 2: Cost-effectiveness plane from Briggs et al. (100)

Despite this uncertainty in economic evaluation results the treatment costs in a clinical trial are often known and precise; a greater cost in the intervention arm compared to comparator arm, resulting in a highly statistically significant difference between treatment arms. Conversely, non-treatment costs often have a statistically non-significant difference. It is when the uncertain non-treatment incremental costs are combined with known and precise incremental treatment costs the resultant plots on the cost-effectiveness plane fall into Scenario 9. Despite treatment costs often being known and precise and non-treatment costs less precise, it is rare to see a discussion in the literature about

the certainty of treatment costs being overshadowed by the uncertainty of non-treatment costs; of the 72 UK National Institute for Health Research (NIHR) HTA reports published in 2020, 32 included a within trial economic evaluation, 17 of these discussed the costs of different resource categories, only 5 of these mentioned increased treatment costs being offset by cost savings in other categories, and none of the reports confirmed that treatment costs were known and certain (Appendix 1: NIHR HTA reports published in 2020 (Volume 24) (*Chapter 2*).

In the absence of alternative guidance, the established recommendations for conducting economic evaluations alongside clinical trials described in this chapter should be followed. However, when there is uncertainty in results decision makers are often left contemplating a cloud of dots on the cost-effectiveness plane, with the only certainty being that the new intervention cost is greater than the comparator.

As well as the uncertainty in economic evaluation results, economic evaluations have been described as 'intervention-focussed' and 'outcomes -driven' (10); the analysis is driven by a direct comparison of costs and outcomes between treatment arms, so the results are defined only by the treatment arm the participants are randomised to. This is also known as a 'black box' evaluation where there is little interest in how the outcomes occur (10) and no understanding of the causal mechanism linking the intervention to the outcomes (101). This intervention focussed approach results in the analyst concentrating only on the ICER and not considering the mechanism driving the ICER. By focussing on a treatment arm-based analysis no attention is given to the causal mechanism that is believed to drive the economic outcomes of costs and health benefit (10). This misses an opportunity to provide additional insight into the trial data, identifying the mechanism linking invention to outcomes and what is driving the outcomes, and giving further understanding of the treatment arm-based results.

Despite the limitations highlighted in Section 2.9 the patient-level data from clinical trials is the preferred choice of many and conducting economic evaluations alongside clinical trials provides an appropriate vehicle to collect this patient-level data. There will continue to be demand for economic

evaluations alongside clinical trials, and there are approaches recommended to address the limitations (62), which were presented in Section 2.9.

Decision makers consider two questions when allocating scarce healthcare resources; 1) should the new technology be adopted, based on the results of the economic evaluation and measures of uncertainty and 2) is any additional information needed to help make the adoption decision (3)? There are four distinct conditions required of economic evaluation methods; defining the objective and limitations of the 'health care provision', defining the decision problem, portraying uncertainty, and providing a method for interpreting the results (3). The conventional guidance for economic evaluation described earlier meets the first three conditions, however the latter consideration still is not completely filled. Furthermore, when there are uncertain results and limited understanding of how the economic outputs occur and are driven, the answer to question 2 above ('is any additional information needed to help make the adoption decision?') is 'yes'. Clinical trials are expensive and time consuming; in England alone clinical research is worth £2.7 billion annually, supporting over 47,000 jobs, and delivering over £28.6 million savings to the NHS (102). As there will always be demand for economic evaluations alongside clinical trials there is scope for proposing a new approach to analysing valuable clinical trial data, which addresses the limitations (uncertainty and 'black box' analysis) of economic evaluations and provides a supplementary analysis to the conventional treatment arm-based analysis to give additional information to decision makers.

This thesis proposes a conceptual model driven analysis of the clinical trial data to provide additional interpretation and understanding of the economic evaluation results based on the expected causal mechanism in the clinical trial, linking the intervention to outcomes. This proposed new approach would act as a supplement to the conventional treatment arm-based approach of analysing clinical trial data for economic evaluations, giving additional evidence to decision makers.

While the use of conceptual models is recommended and frequently adopted for developing decision analytic models in economic evaluation, they are rarely used in the context of trial-based analyses. The proposed new role for conceptual

models is to further investigate, interpret and understand trial data, extending the analysis beyond the conventional treatment-arm based comparison.

Conceptual models visually illustrate the casual relationships between key components in a real-life system or process (13-16). The aim of the new approach is to use the causal relationships inherent in conceptual models to further analyse clinical trial results; the conceptual model will illustrate a simplification of the relationships (causal or otherwise) in the trial mechanism, linking the key economic evaluation components, from treatment (inputs) to health benefit outcomes (outputs). There are a variety of possible types of trial mechanism, including biological, behavioural or a policy change.

To omit bias in this new role the conceptual model should be developed prior to the analysis, however any unexpected outcomes can be explored with additional conceptual model analyses as sensitivity analyses; these could potentially use either a different trial mechanism or different components. Having a conceptual model depicting the anticipated consequences of the intervention can also help to limit data mining and post hoc analysis, although care should be taken not to ignore genuine unexpected consequences of the intervention.

This thesis suggests three aspects of the conceptual model driven analysis:1) As a communication tool, illustrating the contents of the model, depicting expected associations (causal or otherwise) between key components in the trial mechanism, linking inputs to economic outputs; 2) Assessing the expected associations in the conceptual model to confirm (or otherwise) the validity and accuracy of the conceptual model in terms of the expected mechanisms driving the study results, and 3) conducting additional analysis on the clinical trial data to provide additional and more detailed interpretation of the treatment arm-based results.

2.10.2 Conceptual model

The conceptual model should show how the economic inputs (ie patients, facilities and treatment) are linked through the trial mechanism (and mediators) to the economic outputs (ie costs, QALYs or other health benefits).

The key components included in the conceptual model will relate to the assumed trial mechanism and economic inputs and outputs, these should be identified when the conceptual model is developed (a methodological framework for guiding the development of the conceptual model is presented in Chapter 5). These key components are expected to include the disaggregated components of economic resource use and health benefits, and the clinical components of the trial mechanism.

A suggestion of the typical disaggregated economic resource use collected in a clinical trial was discussed previously in sub-section 2.5.1, as reported by Petrou et al., these include: the treatments; healthcare use related to the disease or intervention; treatment of side effects related to the intervention, and other resource use deemed important (6).

Further guidance to help determine relevant disaggregated economic components can be taken from the equations presented by Weinstein and Stason in 1977, these equations illustrate the individual components making up the cost-effectiveness calculus; one relating to incremental costs and one to incremental QALYs (41). The equations give an overview of the individual components expected in an economic evaluation; however they should be applied with relevance to each situation individually as all economic evaluations differ, and therefore so will conceptual models. It is anticipated that each economic evaluation comprises a different set of costs and outcomes, but the equations presented by Weinstein and Stason provide a helpful guide to typical components of an economic evaluation.

In the equation for total incremental costs, the cost categories are broken down into four components (Equation 3): medical and healthcare costs (ΔCRx); costs resulting from treatment side effects (ΔCSE); cost savings resulting from the intervention alleviating and preventing disease (ΔCMorb), and costs from treating diseases resulting from living longer (ΔCRxΔLE).

$$\Delta C = \Delta CRx + \Delta CSE - \Delta CMorb + \Delta CRx\Delta LE$$

Equation 3: Weinstein and Stason's incremental cost equation

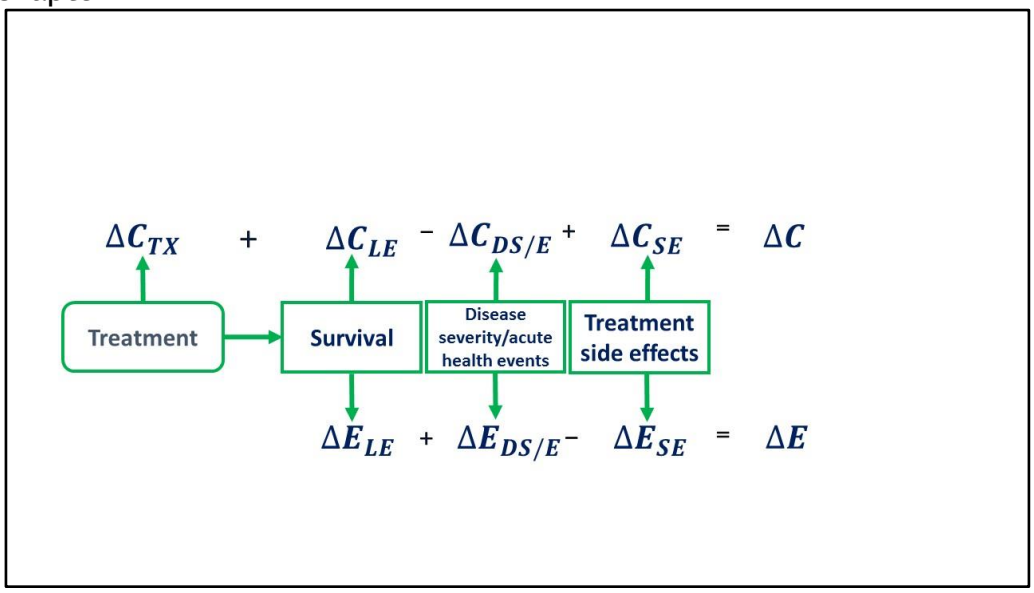
The effectiveness equation representing QALYs is broken down into three components: number of life years (ΔY); improvement in quality of life as a result of the intervention alleviating or preventing disease (ΔYMorb), and a decrease in quality of life from side effects of the treatment (ΔYSE) (Equation 4).

$$\Delta E = \Delta Y + \Delta YMorb - \Delta YSE$$

Equation 4: Weinstein and Stason's incremental effectiveness equation

The number of life years represents the quantity-of-life element of a QALY and the results of the intervention alleviating or preventing disease and detrimental effect of treatment side effects represent the quality of life aspects of the QALY.

From identifying these cost and outcome categories, the analyst specifies the relevant cost components that make up the total net-incremental costs and total net-incremental QALYs. The two components from Weinstein and Stason's disaggregated calculus; costs and health benefits, are combined in a simple template in Figure 3, this template can be used as a guide on which to base the conceptual model. The hypothetical mechanism in this template is based on disease progression, and the intervention is assumed to have an association with survival, acute events linked to the disease, disease severity and side effects from the treatment. The total incremental costs (ΔC) would comprise: incremental treatment costs (ΔC_{TX}); incremental healthcare costs resulting from living longer because of an increase in survival (ΔC_{LE}); incremental cost savings from a decrease in disease severity/number of acute health events ($\Delta C_{DS/E}$), and incremental costs resulting from treatment side effects (ΔC_{SE}). The total incremental effects (ΔE) components comprise: an incremental improvement in the effects measure from an increase in survival (ΔE_{LE}); an incremental improvement in the effects measure as a result of a decrease in disease severity/number of acute health events ($\Delta E_{DS/E}$), and finally an incremental decrement in the effects measure as a result of treatment related side effects (ΔE_{SE}).



TX - treatment, LE - life expectancy, DS/E - disease severity/acute health events, SE - side effects

Figure 3: Adaptation of Weinstein and Stason calculus into a cost-effectiveness template which demonstrates combining a disease process with disaggregated costs and effects

In the adaption of the Weinstein and Stason equations in Figure 3, it is clear that survival, disease severity/acute health events and treatment side effects are mirrored in both the incremental total costs and incremental total effects, however this is not the case with the treatment component of the diagram. The treatment component relates solely to incremental total costs, however it could be argued that process utility is the equivalent to treatment costs in the effects side of the equation. As previously mentioned, in the UK, the concept behind economic evaluation is typically extra-welfarism, where the focus is on the QALY. However, the effects of a patient actually receiving the treatment could also be considered, this is called 'process utility'. Process utility includes elements such as dignity, reassurance and treatment type such as surgery v. drugs or oral v. intravenous drugs (103). Research shows that EQ-5D based QALYs do not capture all factors relating to healthcare, and that process utility outcomes can be successfully captured by 'bolting on' a process utility element to the SF-6D questionnaire (104).

Because treatment costs are often known and precise, if we remove these from the equation, we are left to establish whether there is evidence of cost savings in the remaining cost components. As discussed by Petrou et al. other types of resource use may also be important (6), therefore the template is a guide rather

than a prescriptive measure. This template is referred to as the ‘disaggregated cost-effectiveness template’ throughout the rest of the thesis.

The disaggregation of a mechanism into constituent parts is an approach also used in realist evaluations, where the mechanism behind how an intervention works, focussing on context, mechanisms and outcomes, is broken down into constituent parts (101, 105). It could be argued that the proposed conceptual model approach also focusses on context, mechanisms and outcomes: the context of the clinical trial, the mechanism of the trial and the outcomes driven by the trial mechanism. Realist evaluations are typically conducted in social science, however there is a growing interest in using this approach for complex interventions in economic evaluations (10).

The suggested format of the conceptual model is based on a simple path analysis diagram. Path analysis is a method developed by Sewall Wright in the early twentieth century to test whether a hypothesised model structure is consistent with observed data(106). The path analysis method is based on a set of hypothesised nested causal relationships within a system represented by linear regressions. The structure is derived from the causal path hypothesis and variables (inputs) at the start of the model do not always directly link to the variables at the end (outputs). The links may be mediated through the pathway, some variables are causes, some are effects(107). The path diagram describes the relationships between variables, linking them with directional arrows. This format is similar to directed acyclic graphs (DAGs), the difference being that in path analysis the models predetermine linear causal effects, whereas in DAGs the models may be linear or non-linear (108). It should be noted that path analysis is used in this instance as a basis for the visual aspect of the conceptual model, not for methods of analysis.

The components of the conceptual model should all be represented by a shape (rectangle or square etc.) and the relationships between the components are represented by arrows, the relationships may be direct or mediated through further components. The arrows will either be solid or dashed; solid lines represent assumed relationships and dashed lines represent possible relationships. Components with arrow heads entering them are dependant variables (for example all model outputs). Components with arrows leaving

them are independent (explanatory) variables (for example all model inputs); the component to which their arrow enters (dependent variable) is predicted by them. Mediating entities are both explanatory and dependant; they have arrows entering and leaving. Each dependant variable should be allocated a consecutive number which is used to identify regressions in the conceptual model, these regressions are used to test the assumed and possible associations linked to each numbered entity, this is explained in sub-section 2.10.3. The suggestion of using regressions is only one way to demonstrate how the conceptual analysis could work and is included as an illustrative example to test the associations depicted in the conceptual model in a simple and straightforward way.

The purpose of suggesting a format for the visual representation of the conceptual model is to ensure consistency and clarity of presentation in terms of design and structure.

2.10.3 Testing the accuracy of the conceptual model

The variables from the clinical trial data used in the regressions should be presented and described for clarity and to provide understanding of how each of the conceptual model components are represented in the regressions and which variables they are sourced from.

The numbered regressions identified in the conceptual model should be conducted, the results will confirm, or otherwise, the accuracy and validity of the relationships presented in the conceptual model. Testing the accuracy of the conceptual model helps to identify how the intervention is working; how the key components in the economic evaluation are related to each other and which components are driven by the expected trial mechanism. Any regression model could be used for this purpose, however, as costs and count data are typically skewed the analyst may choose a regression technique that takes account of this and covariates can also be included.

2.10.4 Analysis

Regression results that provide confirmation of the conceptual model structure can be taken forward and utilised to conduct an additional analysis on the clinical trial data to provide additional interpretation of the treatment arm-based results. One suggested illustrative method for this additional analysis is to use bootstrapped samples. Bootstrapped samples can be produced to calculate point estimates for differences in costs and outcomes (ie QALYs), uncertainty around these point estimates, and to calculate an ICER. Bootstrapped results can also be used to produce cost-effectiveness planes, cost-effectiveness acceptability curves (CEAC), net monetary benefit calculations, and value of information analyses.

To operationalise this step the regression results are used to produce bootstrapped estimates, this is achieved by incorporating the regressions that do predict and drive the mechanisms into the conceptual model and omitting those that do not contribute to the mechanism in the conceptual model, as proven by the regression results. Where there are mediating relationships between components in the trial mechanism these should be included as nested equations where the output of one regression feeds into another.

After the outputs of the bootstrapping exercise have been used to calculate point estimates the conceptual model driven analysis bootstrapped samples can be plotted on the cost-effectiveness plane to provide evidence of cost-effectiveness. If the conceptual model driven analysis has replicated the trial mechanism accurately, the spread of the samples should be smaller than the treatment arm-based analysis and may occupy fewer quadrants due to less uncertainty and fewer outliers. Any change in shape, placement or spread of plots between approaches will aid interpretation of the results, particularly in terms of how the trial mechanism has driven the components of the economic analysis, and how these components interact and the dynamics in them. These bootstrapped samples can also be used to produce a CEAC, net monetary benefit measures and value of information estimates. The CEAC will provide estimations of the probability of the intervention being cost-effective at different willingness-to-pay thresholds. The net monetary benefit provides another measure of cost-effectiveness which is not affected by the complexities of

negative ICERs. The value of information measure places a value on eliminating uncertainty; if the results from a conceptual model driven analysis are less uncertain than the treatment arm-based analysis they can provide more certain information on the value of future research, as a whole or based on individual parameters.

Finally, the results of the conceptual model driven analysis can be compared to the results of the treatment arm-based analysis to provide further information and interpretation of both sets of results.

2.10.5 Discussion

The new role for conceptual models has the potential to provide evidence to decision makers, in addition to the conventional treatment arm-based analysis, to help inform the adoption decision. Associations are identified between components in the conceptual model, while regression techniques test the validity of these links and are used to produce bootstrapped samples. These bootstrapped samples are then used to calculate summary measures, measures of uncertainty, and the summary measures from the conventional treatment arm-based analysis can be compared to the conceptual model driven analysis to add further interpretation. This new role for conceptual models provides a structured way of incorporating more information to inform the decision-making process, simultaneously considering the mechanism of the disease or system central to the clinical trial, and the dynamics of the key components of that mechanism relevant to the economic evaluation. Conducting clinical trials and obtaining the associated data is time consuming and expensive, this proposed further analysis of the trial data makes the best use of valuable evidence.

Section 2.9 discussed the limitations inherent when conducting economic evaluations alongside clinical trials, the conceptual model approach just introduced is based solely on the evidence from one clinical trial so has some of these limitations of conducting an economic evaluation alongside a clinical trial. The new approach addresses a sub-set of the limitations in particular uncertain results. The new approach also addresses the criticism of intervention focussed, outcomes-driven economic analysis.

This section would not be complete without a discussion about how this proposed new role for conceptual modelling fits into the ongoing debate about the benefits of a within trial analysis compared to a decision analytic model analysis (24). Many of the suggestions for minimising the limitations described in Section 2.9 are to use a decision analytic model, and it has been argued that to overcome the drawbacks of basing an economic evaluation solely on the data from a clinical trial, additional evidence synthesis and decision-modelling should be included (60). There is a common belief that using clinical trial evidence or decision analytic models for economic evaluation are mutually exclusive (24), this is compounded by the ISPOR task force 'Good research practices for cost-effectiveness analysis alongside clinical trials' (Ramsey et al.) guidance, and the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) reporting guideline treating each approach separately (5, 90). However, treating these approaches as mutually exclusive is a false dichotomy as economic evaluations alongside clinical trials and decision analytic models synthesising evidence are not substitutes for each other, both are important approaches for informing decision makers (24).

The proposed new role for conceptual models could arguably be described as bringing together these two approaches by going beyond the conventional treatment arm-based analysis of clinical trial data and investigating a causal (or associated) pathway based on the key economic components of the trial mechanism using patient level data, introducing a more nuanced understanding of the trial data. Decision analytic models attempt to represent associations in the trial mechanism at an aggregate level using assumed mathematical relationships, considering the biological or clinical processes driving them and basing the model states on the well-known and understood natural history of disease, attaching specific costs and quality of life measures to each state (3).

2.11 Summary

This chapter explained the role of economic evaluation alongside clinical trials and presented the current guidance for conducting economic evaluations. Conventional economic evaluations alongside clinical trials often result in uncertain results, giving limited evidence to decision makers. There is also criticism of the 'black box' approach of economic evaluations, which focusses on treatments and outcomes with no consideration of the mechanism driving the outcomes. However, despite these drawbacks there will continue to be demand for economic evaluations alongside clinical trials, the collection of clinical trial data is expensive and full use should be made of it to provide evidence to decision makers. Therefore, there is scope for introducing a supplementary new approach for analysing the clinical trial data, to provide additional understanding to decision makers and make further use of the valuable clinical trial data.

This chapter proposed such a supplementary new approach, a new role for conceptual models, beyond the conventional role of designing a decision analytic model structure. This new role for conceptual models illustrates the expected trial mechanism resulting from the intervention and identifies the key elements of this mechanism driving the costs and outcomes of the economic evaluation. By simultaneously considering the conceptual model of the disease or system being analysed in the economic evaluation and the disaggregated impact of the components of the economic evaluation, a more nuanced analysis can be achieved; one that remains true to the key principle of estimation not hypothesis testing. This new role for conceptual models provides a structured way of incorporating more evidence to inform the decision-making process. The new role for conceptual models in economic evaluation is the third output and first objective of this thesis.

The next chapter demonstrates the application of the current guidance for treatment arm-based analysis using two case studies.

Chapter 3 Case studies: an illustration of current economic evaluation guidance

3.1 Chapter overview

The previous chapter described the role of economic evaluations in providing decision makers with evidence for funding new healthcare technologies, it summarised the existing guidance for conducting treatment arm-based economic evaluations alongside clinical trials, highlighted the strengths and limitations of conducting economic evaluations alongside clinical trials and introduced the new role for conceptual models.

The aim of this chapter is to illustrate the guidance for conducting standard economic evaluations alongside clinical trials described in Chapter 2, applying it to two case studies. Both case studies are the economic evaluation component of NIHR funded clinical trials, conducted alongside a clinical trial. The first case study is introduced in Section 3.2, this case study is the Theophylline With Corticosteroids (TWICS) study; the second case study, introduced in Section 3.3, is the BeatIt study. The layout of both case studies begin with the background and rationale of the clinical trial with a summary of the clinical outcome results, then the economic evaluation methods using the existing guidance are described, and finally the results of the economic evaluations are reported. The chapter is summarised in Section 3.4.

For both of these case studies my role was the analysis of the data constituting the economic evaluation. The design of the economic evaluation (in terms of resources and outcomes collected for analysis) was completed prior to my involvement with them.

3.2 Case study #1 TWICS

3.2.1 Background

The TWICS study was a multicentred UK clinical trial and the aim of the trial was to determine the clinical and cost-effectiveness of adding a low dose of theophylline to usual care for people with chronic obstructive pulmonary disease

(COPD), compared to usual care plus a placebo. The protocol for the trial and full trial results are reported in detail in published articles (109-111).

COPD is a lung disease and is associated with breathlessness on exertion, it can result in disability, absence from work, morbidity, early retirement, and premature death; the main cause of COPD is smoking, and most cases are diagnosed from the age of 50 onwards. COPD is a progressive disease and patients are likely to deteriorate over time. The burden from COPD in the UK is high; in 2014-2015 there were over 1 million diagnosed cases of COPD, this accounted for almost 2% of the population. In 2005 COPD accounted for 5% of global deaths and it is estimated that in 2030 COPD will be the third leading cause of death globally (112). In the UK COPD was recorded as the cause of approximately 30,000 deaths per year between 2008 and 2018 (113). The financial burden from COPD on the National Health Service (NHS) is about £1.9 billion per year (114).

A key feature of COPD are exacerbations, these are defined as a 'sustained worsening of the patient's condition from the stable state and beyond normal day-to-day variations that is acute in onset and may warrant additional treatment in a patient with underlying COPD' (115). Symptoms of an exacerbation include breathlessness, coughing, and expelling mucus; this sudden worsening of a patient's health is likely to result in a decrease in a patient's quality of life. Many exacerbations require treatment to manage them, this includes treatment with antibiotics and corticosteroids; less severe exacerbations can be managed at home or in the community, however more severe exacerbations require treatment in hospital, with associated higher costs of treatment. Around 15% of COPD patients will experience an exacerbation annually, and of those hospitalised with an exacerbation 12% will die within a year of hospitalisation (116). Exacerbations are linked to an increased decline in lung function, reduced physical activity, lower quality of life, an increased risk of some co-morbidities and increased mortality (115). Therefore, the burden of patients with COPD on the NHS service is high.

Treatment recommended for maintenance of COPD includes inhaled corticosteroids (ICS), combined with inhaled long-acting β_2 agonists (LABA), however many patients still have exacerbations despite this treatment (117).

High doses of oral theophylline have also been used to treat COPD, but the high doses involved cause unpleasant side effects in patients resulting in treatment with other, better tolerated bronchodilators. However, recent preclinical trials have indicated that oral theophylline at a low dose is beneficial in treating COPD in patients who are also treated with ICS; the low dose theophylline changes biological mechanisms in the patient allowing ICS to potentially reduce numbers of exacerbations more efficiently.

TWICS was a randomised, pragmatic, double-blind, placebo trial comparing usual care plus low dose theophylline to usual care plus a placebo for patients with COPD. The primary outcome of the trial was the number of moderate or severe exacerbations (requiring treatment/a change in management) during the 1-year trial follow-up period. The main inclusion criteria for participants included: people with a diagnosis of COPD; over 40 years old; currently treated with ICS, and with a history of two exacerbations needing treatment in the 12-month period prior to randomisation. Participants were randomised to either usual care plus low dose theophylline or usual care plus a placebo for 52 weeks. The dose of theophylline was either 200mg or 400mg; the specific dose was determined by participants' ideal body weight and smoking status.

3.2.2 Clinical results

Participants were recruited from 121 UK primary and secondary care sites between February 2014 and August 2016. 1,578 people were recruited and randomised to a treatment arm, 11 were excluded post-randomisation, leaving 1,567 participants in total; 788 were randomised to the theophylline intervention arm and 779 were randomised to the placebo control arm. Analysis of baseline characteristics found that participants in the two arms of the trial were evenly balanced for demographic and disease characteristics.

Primary outcome data (number of exacerbations) was available for 1,536 participants (Table 1): 772 in the theophylline arm and 764 in the placebo arm. 633 (82.0%) participants in the theophylline arm and 609 (79.7%) participants in the placebo arm reported one or more exacerbations during the trial follow-up period. There was a total of 1,727 exacerbations reported by participants in the theophylline arm, a mean of 2.24 (standard deviation (SD) 1.99), and a total of

1,703 reported exacerbations in the placebo arm, a mean of 2.23 (SD 1.97). Although the mean number of exacerbations was slightly higher in the theophylline arm, statistically there was no difference between arms in the exacerbation rate. The incidence rate ratio (IRR) was 0.99 (95% confidence interval (CI) 0.91 – 1.08); because the confidence interval crosses '1' the IRR is not a statistically significant result and the trial concluded that there was no difference in treatment effect between treatment arms.

106 (13.7%) participants in the theophylline arm and 130 (17.0%) participants in the placebo arm reported one or more exacerbations needing hospital treatment during the trial follow-up period. This amounted to a total of 134 hospital admissions in the theophylline arm and a total of 185 in the placebo arm; there were 51 more exacerbations resulting in a hospital stay in the placebo group compared to the theophylline group. The mean number of exacerbations requiring hospital treatment was 0.17 (SD 0.49) in the theophylline arm and 0.24 (SD 0.66) in the placebo arm, resulting in an adjusted IRR of 0.72, (95% CI 0.55 to 0.94), this is a significant result suggesting that theophylline decreases the severity of exacerbations. Further investigation of these results showed that most of these exacerbations were a result of more participants in the placebo arm reporting three or more exacerbations than in the theophylline arm. 39 exacerbations needing hospital treatment were linked to 10 participants in the placebo group (who reported three or more exacerbations needing hospitalisation during the follow-up period), compared to three participants in the theophylline arm reporting 12 exacerbations needing hospital treatment. This additional investigation concluded that theophylline does not reduce the number of exacerbations needing hospitalisation (and therefore does not decrease the severity of exacerbations) compared a placebo, furthermore the trial was multicentred and it might be that some sites have different care pathways in terms of deciding what severity of exacerbation is treated in the community or in hospital.

Table 1: TWICS primary outcome (exacerbations) summary

	Theophylline	Placebo	Adjusted /unadjusted IRR	Estimate (95% confidence interval)
Exacerbations				
Participants included in analysis	772	764		
Participants with at least one exacerbation	633	609		
Total number of exacerbations	1,727	1,703	Unadjusted	1.00 (0.92 to 1.09)
Mean number of exacerbations (SD)	2.24 (1.99)	2.23 (1.97)	Adjusted	0.99 (0.91 to 1.08)
Exacerbations needing hospital treatment				
Participants included in analysis	772	764		
Participants with at least one exacerbation	106	130		
Total number of exacerbations	134	185	Unadjusted	0.72 (0.55 to 0.95)
Mean number of exacerbations (SD)	0.17 (0.49)	0.24 (0.66)	Adjusted	0.72 (0.55 to 0.94)

IRR – incidence rate ratio, SD – standard deviation

In summary there was no statistically significant differences between arms in the total number of exacerbations, therefore, there was no difference in the treatment effect between arms on the rate of exacerbations.

3.2.3 Economic evaluation methods

The aim of the economic evaluation was to estimate the cost-effectiveness of adding theophylline to usual care in reducing exacerbations needing treatment, compared to usual care plus placebo. The economic evaluation followed the existing guidance presented in Chapter 2 and the order in which the methods are presented below are: 'Data identification and collection - Resource use', 'Data identification and collection - Health benefits' and 'Analysis'.

Chapter 3

Data identification and collection - Resource use

The first step was to ascertain relevant resource use for each of the participants. The guidance described in Chapter 2 sub-section 2.32.5.1 recommends using the perspective of the economic evaluation to guide the identification of the type of resources used by participants. This guidance also recommends considering the following categories to identify relevant resource use groups for inclusion in the economic evaluation; resources linked to the treatments and disease; resources likely to differ between arms, and resources linked to treatment side effects. Based on these recommendations the following types of resource were collected in TWICS:

Perspective – the perspective of the economic evaluation was the NICE recommended NHS and personal social services perspective (63); so only healthcare resource use that was paid for by the UK NHS (the healthcare payer) was included.

Treatment – the intervention of the trial was low dose theophylline; resource use was collected for the participants in the intervention arm based on their assigned dose. The comparator was placebo so no resource was assigned to this.

Disease – COPD is managed in this population with routine maintenance therapy; the resources used for this therapy were collected for all participants with input from clinicians to identify possible therapies.

Between arms difference **-** the primary outcome of the trial was the number of exacerbations needing treatment reported by participants during the trial follow-up; resource use relating to the treatment, length of treatment and location of treatment (the location of the treatment was specified as 'home', 'care by services to prevent hospitalisation' and 'admitted to hospital') of the exacerbations was collected.

Side effects/disease – resource use that could potentially be linked to either COPD itself or the side effects of the intervention was also collected; this included resource use from inpatient stays, outpatient attendances, primary

care use, emergency hospital admissions (not related to COPD) and regular (non-COPD maintenance) medication.

Resource use collection

Resource use was collected from each participant at three timepoints during the trial; baseline, six months after randomisation and 12 months after randomisation using a modified version of the Client Service Receipt Inventory (CSRI) (118), this is a research questionnaire used for retrospectively collecting health and social care related resource use from participants. The CSRI was included in the trial case report form and the timing of the resource use collection coincided with face-to-face assessments for the collection of clinical outcomes and safety data from trial participants, this was deemed frequent enough to minimise recall bias.

At the baseline visit resource use categories collected were: COPD maintenance therapy; the number of exacerbations in the previous 12 months; the number of exacerbations resulting in hospitalisation in the previous 12 months, and medication.

At the six and 12 month visits the resource use categories collected were: treatment for exacerbations; COPD maintenance therapy; other health service use (including inpatient stays, outpatient attendances, primary care use and non-COPD related emergency hospital admissions), and non-COPD maintenance medication, at each follow-up visit this resource use was retrospectively collected for the previous six months. The dose of the theophylline intervention for participants in the theophylline arm was collected in the trial database and the trial health economist was given access to this data.

Unit costs

Unit costs were obtained from different sources depending on the type of resource use category. Medication costs for the intervention (theophylline) and COPD maintenance therapy were obtained from the British National Formulary (BNF) (119), exacerbation costs, inpatient stays, outpatient attendances, primary care costs and non-COPD emergency admissions were obtained from five

sources: NHS reference costs (120); Information Services Division (ISD) (121); Personal Social Services Research Unit (PSSRU) (122); BNF (119), and published literature (123, 124). Unit costs were all adjusted to 2016 prices using the UK Health Service Cost Index (122) and valued in pounds sterling (£).

Data identification and collection - Health benefits

The economic outcome used for measuring the health benefits experienced by participants in the trial was the preference-based QALY, this outcome was measured using the EQ-5D-3L questionnaire (74); a self-reported quality of life questionnaire, this outcome is the preferred measure of heath benefit described in the guidance in Chapter 2 sub-section 2.5.2. The EQ-5D-3L questionnaire measures quality of life using five domains, these are: mobility; self-care; usual activities; pain/discomfort, and anxiety/depression. Participants choose from three levels of severity in each domain; for example in the mobility domain options include 'I have no problems in walking about', 'I have some problems in walking about' and 'I am confined to bed'. Results from the questionnaire were valued using the UK value set (125), resulting in utility values of between -0.59 and 1 where -0.59 represents a state worse than death, 0 represents death and 1 represents full health. The quality of life utilities were combined with length of life (12 months follow-up or less if the participant died during the trial) to calculate QALYs using standard area under the curve methods. The changes in the participant utilities between follow-up visits were assumed to be linear when calculating the QALYs.

Designing the economic evaluation to reduce 'noise'

At the design stage several decisions were made to limit the potential 'noise' in the results by limiting non-treatment resources likely to differ between arms and linked to the disease. These included the perspective of the economic evaluation being restricted to the NHS and personal social services. The main focus of resource use was healthcare directly relating to COPD: maintenance treatment and treatment of exacerbations. Other resource use that could potentially relate to COPD included: inpatient stays, outpatient visits and primary care use. Two further resource use categories were collected, which

may not have been related to COPD: emergency hospital admissions not related to COPD and non-COPD medication.

Analysis

As recommended by the existing guidance for conducting economic evaluations, an analysis plan was produced, and an early summary was included in the published trial protocol (111).

The analysis was based on the intention-to-treat population; participants' results were analysed according to the treatment arm to which they were randomised. Where possible missing data was replaced based on plausible assumptions agreed with a clinical expert, where this was not possible missing data was replaced using the multiple imputation by chained equations method (84). Three analyses were conducted: a complete case analysis; unadjusted multiple imputation, and adjusted multiple imputation using a generalised linear model (to address heterogeneity in the participants). The data collected from participants who were lost to follow-up (censored) was included in the analysis, and the time that they had spent in the trial was used to adjust total costs and QALYs using regression techniques.

The total costs and total QALYs were summed for each participant for both treatment arms, then divided by the number of participants in each treatment arm to give a mean cost or mean QALY per participant in each arm, the measures of variability for these means were standard deviation. The difference in mean costs and QALYs between treatment arms were calculated and presented with 95% confidence intervals for the measure of variability. The summary measure was the ICER, to assess cost-effectiveness this was compared to the current NICE willingness to pay threshold of £20,000 (55); if the ICER was below £20,000 the intervention would be deemed to be cost-effective.

As well as including the 95% confidence intervals as described, recommendations in the guidance for representing uncertainty were followed by creating bootstrapped samples using non-parametric techniques; these were plotted onto a cost-effectiveness plane with a 95% confidence ellipse and were also used to produce a cost-effectiveness acceptability curve (2). As previously mentioned,

uncertainty driven by the heterogeneity of participants was adjusted for using regression techniques based on the baseline participant characteristics as covariates. For the costs these covariates were: medication count at baseline; EQ-5D-3L outcome at baseline; offset time (time spent in the trial); age; number of hospitalisations for exacerbations in the 12 months prior to randomisation; number of exacerbations in the 12 months prior to randomisation, and a cluster command was used for the trial site. For QALYs the covariates were: baseline EQ-5D-3L data; medication count at baseline; offset time; age; sex; hospitalisation for exacerbations in the 12 months prior to randomisation; exacerbations in the 12 months prior to randomisation, and a cluster command was used for the trial site.

3.2.4 Economic evaluation results

The economic evaluation results are presented in full in the NIHR report (109). A summary of the results from this original treatment arm-based analysis are presented below, with additional results calculated specifically for this thesis (treatment and non-treatment cost point estimates, treatment and non-treatment costs cost-effectiveness planes and value of information - EVPI).

Economic resource use data was available for 1,470 participants: 743 in the theophylline arm and 727 in the placebo arm and quality of life data from the EQ-5D-3L questionnaire was available for 1,243 participants: 635 in the theophylline arm and 608 in the placebo arm. There was more missing data in the placebo arm compared to the theophylline arm; 4.8% v. 3.8% for resource use, and 20.4% v. 17.7% for EQ-5D-3L data (Table 2).

Table 2: TWICS missing economic data

Data	Theophylline (n=772)	Placebo (n=764)	Total (n=1,536)
Resource use	29 (3.8%)	37 (4.8%)	66 (4.3%)
EQ-5D-3L at baseline/6 or 12 months	137 (17.7%)	156 (20.4%)	293 (19.1%)

Baseline characteristic covariates included in the adjusted multiple imputations are presented in Table 3, the baseline variables were either significant predictors of costs or QALYs or both. There were no significant differences between arms for the baseline resources.

Table 3: TWICS baseline participant characteristics

	Theophylline Mean (SD)	Placebo Mean (SD)
Medication count at baseline * ^	4.65 (3.64)	4.41 (3.54)
EQ-5D-3L outcome at baseline * ^	0.629 (0.280)	0.643 (0.279)
Number of hospitalisations for exacerbations in the 12 months prior to randomisation * ^	0.404 (0.840)	0.358 (0.918)
Number of exacerbations in the 12 months prior to randomisation * ^	3.63 (2.22)	3.52 (2.08)
Age * ^	68.3 (8.2)	68.5 (8.6)
Sex ^		
Male	425/788 (53.9)	418/779 (53.7)
Female	363/788 (46.1)	361/779 (46.3)

*covariates adjusted for in cost model, ^covariates adjusted for in QALY model

Results for the complete case and multiple imputed (unadjusted and adjusted) analyses, are reported in Table 4.

Table 4: TWICS original treatment arm-based economic evaluation results

	Theophylline mean (SD)	Placebo mean (SD)	Difference	95% confidence interval
Complete case analysis (unadjusted)				
Total costs	£2,684 (£2,882)	£3,136 (£4,851)	-£452	-£771 to -£133
Total QALYs	0.626 (0.259)	0.637 (0.263)	-0.011	-0.040 to 0.018
Unadjusted multiple imputation results				
Total costs	£2,702 (£110)	£3,141 (£148)	-£439	-£846 to -£32
Total QALYs	0.617 (0.010)	0.621 (0.010)	-0.004	-0.031 to 0.024
Adjusted multiple imputation results				
Total costs	£2,784 (£125)	£3,006 (£167)	-£222	-£472 to £27
Total QALYs	0.621 (0.006)	0.616 (0.007)	0.005	-0.015 to 0.025

SD – standard deviation, QALYs – quality adjusted life-years

Complete case results for costs show that the theophylline arm was less costly than the placebo arm; £2,684 compared to £3,136, a statistically significant difference of £452 (95% CI £133 to £771). Complete case results for QALYs show that participants in the placebo arm reported higher QALYs than the participants in the theophylline arm; 0.637 compared to 0.626, a difference of 0.011 (95% CI -0.018 to 0.040). The complete case results suggest that the ICER falls into the southwest quadrant where decision makers would need to assess what level of cost savings per loss in health benefit that they would be willing to accept in order to adopt the theophylline intervention, this is clearly illustrated by plotting the bootstrapped samples on a cost-effectiveness plane (Figure 4). The cost-effectiveness plane includes a confidence ellipse to illustrate 95% confidence; we can be 95% confident that the true difference between costs and QALYs lies within this ellipse.

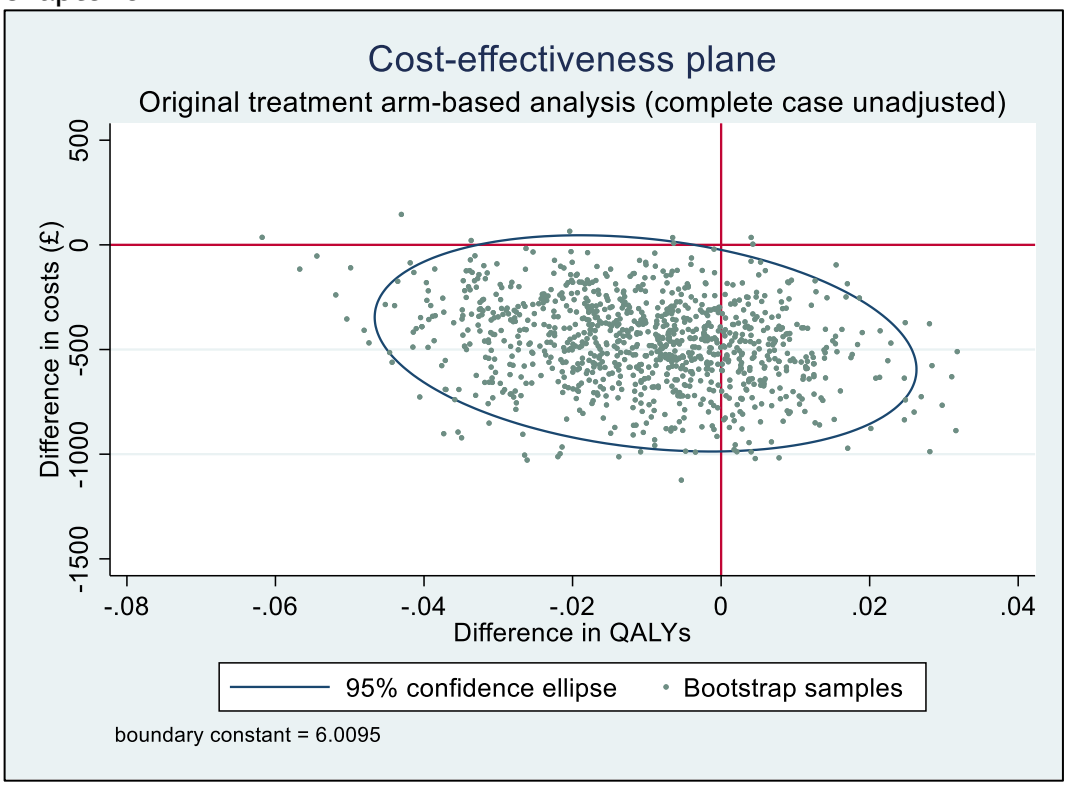


Figure 4: TWICS cost-effectiveness plane – original treatment arm-based analysis complete cases (unadjusted)

The bootstrapped samples were also used to plot a CEAC, which showed that at the UK willingness-to-pay threshold of £20,000 there is a 75% chance of the theophylline arm being considered cost-effective (Figure 5).

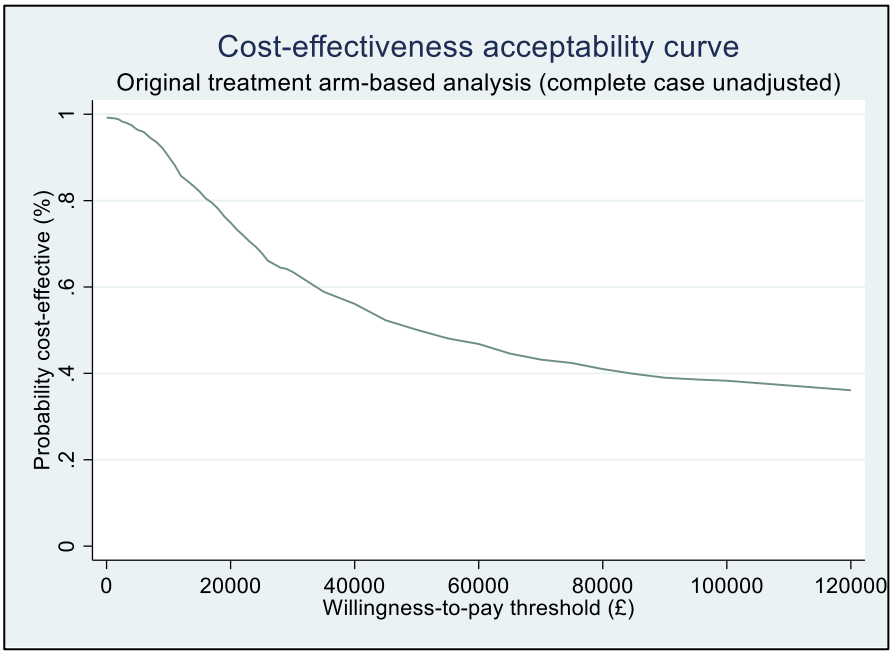


Figure 5: TWICS cost-effectiveness acceptability curve - original treatment arm-based analysis complete cases (unadjusted)

Once the missing data was replaced using multiple imputation the difference in costs decreased to £439; still higher in the placebo arm, however although this difference was still statistically significant the uncertainty around the cost difference was wider; 95% CI £32 to £846. Imputed results for the difference in QALYs between arms mirror the complete case results, with higher QALYs reported in the placebo arm compared to the intervention arm; 0.621 compared to 0.617, a smaller difference of 0.004 (95% CI -0.024 to 0.031). The multiple imputation results also suggest that the ICER falls into the southwest quadrant where decision makers would need to assess what level of cost savings per loss in health benefit that they are willing to accept in order to accept the theophylline intervention.

When the imputed results were adjusted for heterogeneity using baseline participant characteristics and offset time, costs remained higher in the placebo arm compared to the intervention arm, a smaller difference of £222 but this difference was no longer statistically significant, shown by the 95% confidence interval around the difference; -£27 to £472. In this analysis the direction of the QALY results reversed; QALYs were marginally higher in the theophylline arm; 0.621 compared to 0.616, a difference of 0.005, again this wasn't a significant result shown by the 95% confidence interval of -0.015 to 0.025. This change in direction of QALY difference is likely, in part, due to a small imbalance of EQ-5D-3L scores at baseline, however the difference in QALYs is not statistically significant in any of the analyses. These multiple imputed adjusted results suggest that the ICER falls into the southeast quadrant where theophylline is considered to dominate placebo; it is less costly and produces more health benefit compared to the placebo arm, however it can be seen from Figure 6 how wide the spread of QALY differences are, and that they straddle the vertical axis of the cost-effectiveness plane. An ICER was not calculated as planned due to the dominance in these results.

Bootstrapped results taken from the multiple imputed and adjusted analysis and plotted on the cost-effectiveness plane show that the majority of the bootstrapped samples do fall into the southeast quadrant where the intervention would be considered cost-effective, however this is not a significant result,

which is indicated by the 95% confidence intervals reported in Table 4 crossing zero and the 95% ellipse crossing into all four quadrants (Figure 6).

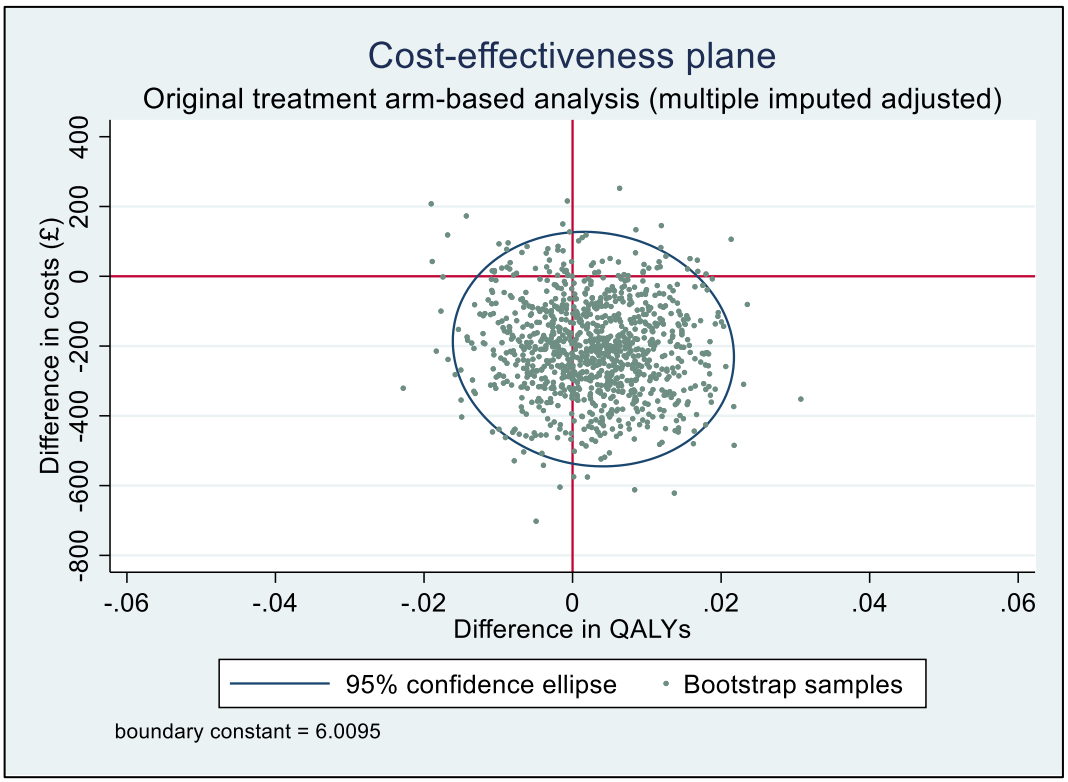


Figure 6: TWICS cost-effectiveness plane – original treatment arm-based analysis multiple imputation (adjusted)

The bootstrapped samples were used to plot a CEAC, which suggests that the UK willingness to pay threshold of £20,000 there is a 90% chance of theophylline being cost-effective (Figure 7). Although the results from the multiple imputed and adjusted analysis, for costs and QALYs, are statistically insignificant the probability of cost-effectiveness at £20,000 threshold is higher (90% compared to 75%) than the complete case analysis which had statistically significant results for the difference in costs, this is driven by the cloud of bootstrapped samples moving from the southwest quadrant (where decision makers decide at what level the cost savings, or compensation for QALY lost is acceptable), towards the southeast quadrant where theophylline is considered to be dominant – cost-effective.

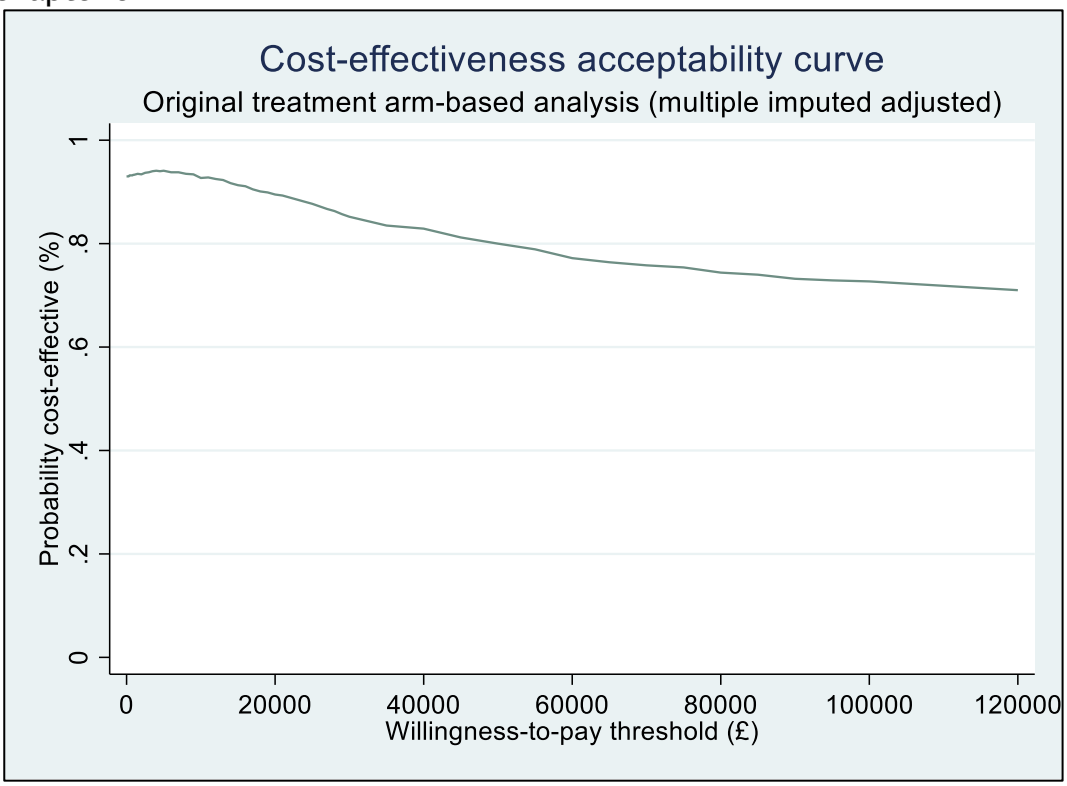


Figure 7: TWICS cost-effectiveness acceptability curve - original treatment arm-based analysis multiple imputation (adjusted)

An exploratory analysis was carried out to further understand the higher number of exacerbations needing hospital treatment in the placebo arm compared to the theophylline arm (discussed in sub-section 3.2.2). Mean total exacerbation costs were higher in the placebo arm compared to the theophylline arm, a difference of £447 (95% CI £186 to £709), a statistically significant result. Exacerbation costs were broken down into the location of the exacerbation treatment and treatment costs for the exacerbation. Taking location of treatment costs first, these were higher in the placebo arm compared to the theophylline arm, a difference of £422 (95% CI £171 to £673), again a statistically significant result. When the location costs are broken down into the different locations ('treatment at home', 'care by services to prevent hospitalisation' and 'admitted to hospital') only the costs of exacerbations treated in hospital showed statistical significance between arms; costs were higher in the placebo arm compared to the theophylline arm, a difference of £416 (95% CI £177 to £655). The treatment costs were higher in the placebo arm compared to the theophylline arm, a difference of £25 (95% CI £8 to £41). This difference was driven by high oxygen costs for seven participants, six of whom were in the placebo arm.

To further investigate exacerbation costs the mean cost of an exacerbation treated in hospital in each arm was calculated; the costs of exacerbations in the placebo arm were higher compared to the theophylline arm (£3,613 compared to £2,671), a difference of £941 (95% CI £140 to £1,743). This difference was driven by the placebo arm recording the 10 most expensive exacerbations in the trial. Interpreting this exploratory analysis, and because of the lack of treatment effect, the trial team believed that the difference in exacerbations requiring hospital treatment to be a chance finding.

For the purposes of this thesis further analysis was carried out using complete case data to investigate treatment and non-treatment costs and EVPI.

As the comparator treatment was a placebo, this arm had no treatment costs, resulting in a highly statistically significant difference in treatment costs between arms of £22 (95% CI £22 to £22). Bootstrapped samples from unadjusted complete cases are presented on a cost-effectiveness plane (Figure 8), showing the difference in treatment costs and no difference in QALYs

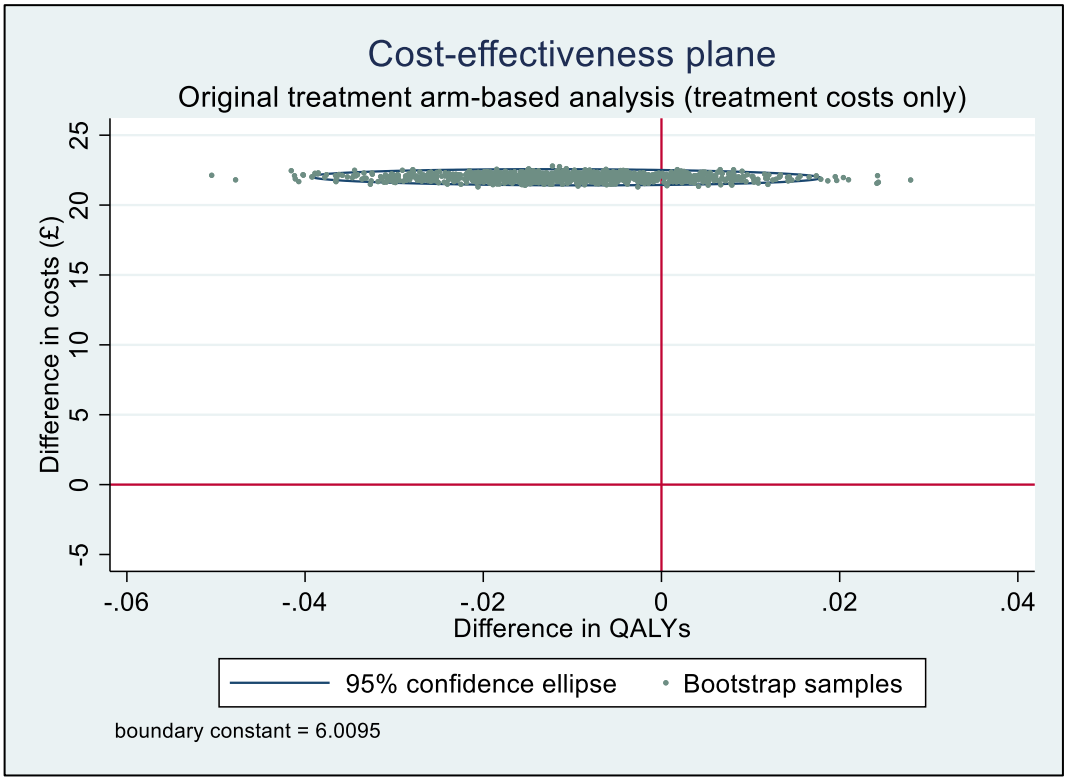


Figure 8: TWICS cost-effectiveness plane – original treatment arm-based analysis treatment costs only

The difference in non-treatment costs between arms was £474 (95% CI £155 to £793), statistically higher in the placebo arm. As discussed above exacerbations treated in hospital were statistically significantly higher in the placebo, and this result is driving the difference in non-treatment costs. Bootstrapped non-treatment costs are presented in a cost-effectiveness plane in Figure 9, this shows that there is no difference in QALYs but the difference non-treatment costs between arms favours the theophylline arm.

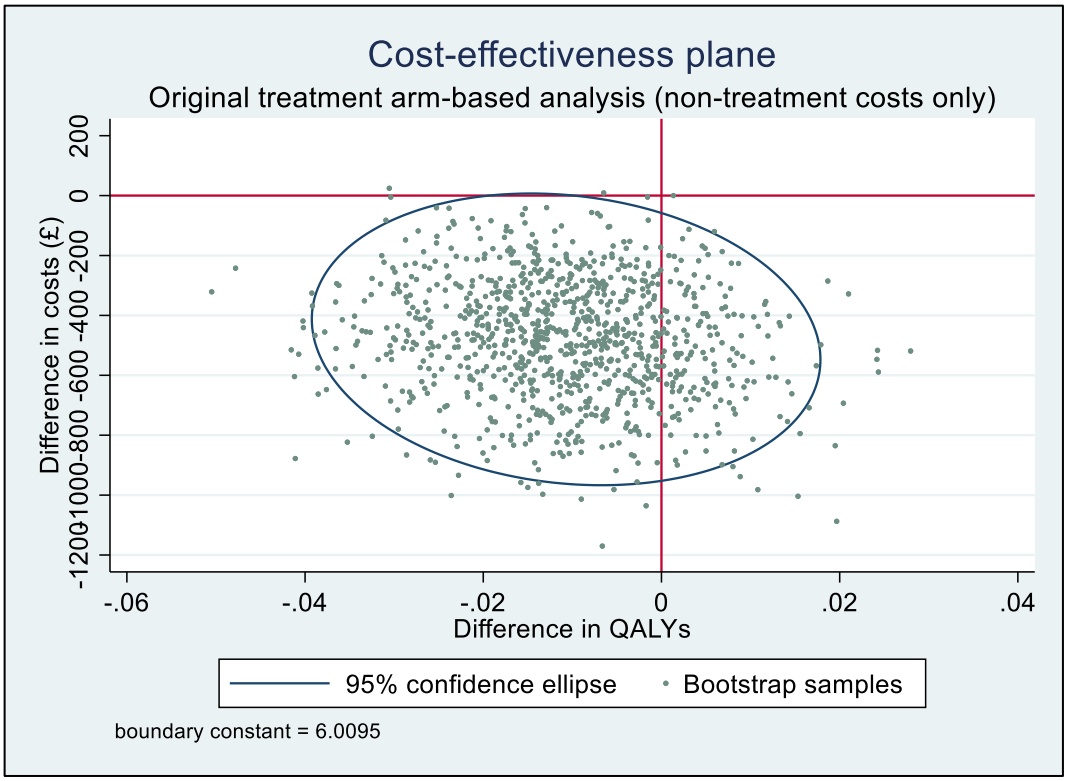


Figure 9: TWICS cost-effectiveness plane – original treatment arm-based analysis non-treatment costs only

If the exacerbation costs are also removed from the non-treatment costs, leaving only the costs of COPD maintenance therapy, other health service use (including inpatient stays, outpatient visits, primary care use and non-COPD emergency hospital admissions) and non-COPD medication, the difference between arms is £26 (95%CI -£181 to £234), a non-statistically significant result with a wide confidence interval. A cost-effectiveness plane including the difference in non-treatment, non-exacerbation costs and QALYs is presented in Figure 10, this clearly replicates a Scenario 9 where there is no difference between costs and QALYs.

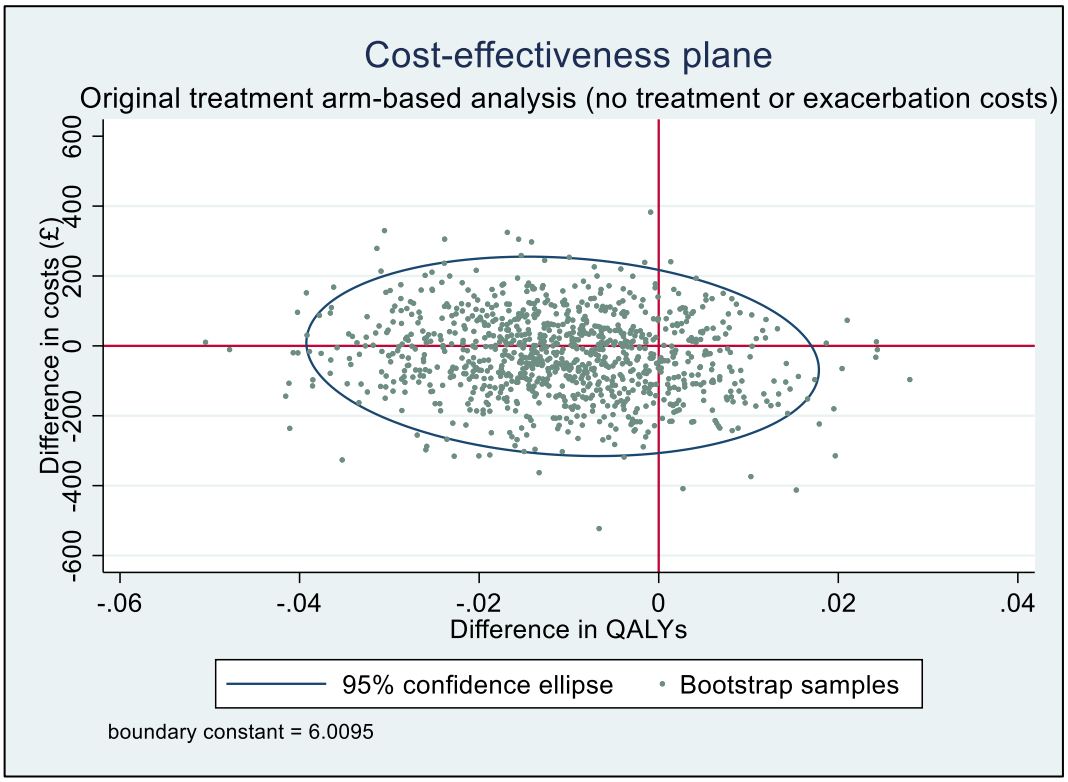


Figure 10: TWICS cost-effectiveness plane – original treatment arm-based analysis non-treatment, non-exacerbation costs only

EVPI results estimate that the value of perfect information per participant is £41, therefore it is unlikely further research is worthwhile in terms of value for money as the cost of extra investigation is likely to be greater than this. This result is driven by the difference in complete case total costs that shows that the total costs in the theophylline arm are statistically significantly lower than total costs in the placebo arm, however the additional analysis above shows that this is driven by a small number of participants in the placebo arm reporting exacerbations requiring hospital treatment, and when exacerbation costs are removed there is little difference between arms.

This additional analysis suggests that the difference in treatment costs is known and clear, however the difference in non-treatment costs, particularly when exacerbation costs are removed is not clear. This highlights the fact that whilst treatment costs are often known and precise, non-treatment costs contain noise and are less easy to interpret.

Whilst the total cost results suggest that theophylline dominates; is both cheaper and has higher QALYs than the placebo arm, the results should be interpreted with caution as the difference between arms for non-treatment, non-exacerbation costs is not statistically significant. There is not a statistically significant difference in QALYs, a result echoed by the clinical effectiveness results: theophylline is not clinically effective in reducing exacerbations.

3.3 Case study #2 BeatIt

3.3.1 Background

BeatIt was a multicentre UK based clinical trial which aimed to establish the effectiveness and cost-effectiveness of the BeatIt intervention compared to the StepUp intervention for adults with intellectual disabilities and a diagnosis of depression. The trial protocol and full details of the trial results are reported in detail in published articles (126-128).

Globally, depression is common; it affects around 300 million people (129) and places a huge burden on healthcare services in terms of costs of prescription and treatment, and on society with time off work. In 2007 the estimated number of people with depression in England was 1.24 million, and costs were £1.7 billion for healthcare services and £5.8 billion in lost earnings, a total of £7.5 billion (130). In 2017/18 the cost of antidepressants prescribed in Scotland was £44.8m, and over 900,000 patients were prescribed at least one antidepressant drug (131).

Research suggests that depression is highly prevalent in adults with intellectual disabilities, with a point prevalence of 5% (132). Depression is also more enduring in adults with intellectual disabilities compared with the general population; a cohort study found that in adults with intellectual disabilities

chronic depression was five times more common than in the adult general population (133).

Psychological therapies for treating depression in the general population are well established, but for adults with intellectual disabilities techniques such as cognitive behavioural therapy (CBT), which are essentially talking therapies and rely on verbal communication, are less accessible (126). While research into the use of CBT for adults with intellectual disabilities and depression has been encouraging (134), behavioural activation therapy, which is less reliant on verbal and cognitive skills, maybe more suited to this population.

Behavioural activation therapy focuses on the link between mood and activity; it aims to increase participation in purposeful and motivating activities, thereby bringing the individual into contact with positive experiences and helping to lift their mood (135). Research into members of the general population with a diagnosis of severe depression has shown that behavioural activation therapy is as effective as antidepressant medication and more effective than CBT (136, 137), and the positive treatment effects have been shown to last as long as those for CBT (138). Behavioural activation therapy is also recommended in NICE (2009) guidelines for the treatment of depression in the general population (139).

There is little evidence on the effectiveness and cost-effectiveness of behavioural activation therapy in the population of adults with intellectual disabilities for treating depression; previous research is limited to a feasibility study undertaken prior to the clinical trial described here. The feasibility study evaluated the feasibility and acceptability of a behavioural activation therapy for treating adults with intellectual disabilities and symptoms of depression (140). A manual to guide the delivery of the behavioural activation therapy was developed and piloted for this population in the feasibility study. Twenty-three adults were recruited into the feasibility study, only two participants dropped out and a further two were lost to follow-up; the high completion rate indicated that the intervention was acceptable. Results showed a significant reduction in self-reported depressive symptoms post-treatment, this reduction was maintained at 3-months follow-up suggesting that the behavioural activation

therapy intervention was feasible, and the authors concluded that a full randomised controlled trial was needed to establish effectiveness.

This full trial was BeatIt; a multicentre, single-blind (participants and therapists were not blinded but the assessors of measures were), randomised controlled trial. The aim of the trial was to determine the clinical effectiveness and cost-effectiveness of BeatIt (behavioural activation therapy - intervention) compared to StepUp (guided self-help therapy - comparator) in reducing self-reported depressive symptoms in a population of adults with mild to moderate intellectual disabilities and a diagnosis of depression. Participants were recruited from sites in Scotland, England and Wales. Key inclusion criteria for participants were mild to moderate learning disabilities; adults aged 18 years old or older; clinically significant depression; ability to give informed consent, and had a supporter (family member or carer) who was able to complete screening and baseline visits and accompany them to follow-up visits and therapy visits, plus provide a minimum of two hours support per week.

Participants were randomised to either BeatIt or StepUp in a ratio 1:1. Therapists were trained to deliver one therapy only; BeatIt was delivered over 12 sessions on a weekly to fortnightly basis, and StepUp was delivered over 8 sessions also on a weekly to fortnightly basis.

The main outcome measure was the Glasgow Depression Scale for Learning Disabilities (GDS-LD) score at 12-months post-randomisation (141). This is a self-reported measure designed to assess depressive symptoms in this population, it has 20 questions with a choice of three answers: never, sometimes, always. Results range from a score of 0 to 40, with lower scores indicating fewer depressive symptoms.

Participants and supporters were assessed separately at 4- and 12-months post-randomisation during face-to-face assessments for the collection of data on resource use and outcomes, an additional data collection point at 8-months was included for the supporter only, via telephone.

3.3.2 Clinical results

In total 161 participants were recruited and randomised to a treatment arm; 84 to the BeatIt arm and 77 to the StepUp arm. Six participants were lost to follow-up and 14 withdrew during the trial; this left a total of 141 participants who completed the trial (68 in the BeatIt arm and 73 in the StepUp arm). The primary outcome was available for a total of 135 participants; 65 in the BeatIt arm and 70 in the StepUp arm. At the primary endpoint of 12-months there was no statistically significant difference in treatment effect between the arms; the mean GDS-LD score at 12-moths post-randomisation in BeatIt was 12.03 (SD 7.99) and in StepUp it was 12.43 (SD7.64), a difference of 0.40 (95% CI -2.26 to 3.06).

However, a statistically significant reduction in the GDS-LD score was witnessed in both arms of the trial at both face-to-face follow-up points; 4- and 12- months post-randomisation. The reduction in GDS-LD scores for BeatIt and StepUp at the 12-months follow-up was 4.20 (95% CI -6.0 to -2.40) and 4.46 (95% CI -6.21 to -2.70) respectively, the difference between arms for this change was 0.26 (95% CI -2.18 to 2.70), a statistically insignificant result.

A similar pattern is seen in measures of activity: The Index of Community Involvement (ICI) measures participation in social and community-based activities, and the Index of Participation in Domestic Life (IPDL) measures changes in participation in domestic tasks. Neither measure found a statistically significant difference between treatment arms.

A summary of clinical results is provided in Table 5.

Table 5: BeatIt primary and secondary (activity) outcomes summary

	BeatIt	StepUp	Intervention effects (95% confidence interval)
	Mean (SD) n	Mean (SD) n	
Glasgow Depression Scale for people with a Learning Disability (GDS-LD)			
Baseline	16.60 (7.91) n=84	16.90 (6.73) n=77	
Post intervention (4 months)	11.91 (7.43) n=68	12.94 (7.77) n=72	-1.03 (-3.58 to 1.51)
Follow-up (12 months)	12.03 (7.99) n=65	12.43 (7.64) n=70	-0.40 (-3.06 to 2.26)
Change in mean scores from baseline (95% confidence interval)			
Post intervention (4 months)	-5.15 (-6.70 to -3.60)	-4.40 (-5.89 to -2.91)	-0.75 (-2.80 to 1.31)
Follow-up (12 months)	-4.20 (-6.0 to -2.40)	-4.46 (-6.21 to -2.70)	0.26 (-2.18 to 2.70)
The Index of Participation in Domestic Life (IPDL)			
Baseline	19.07 (8.46) n=84	18.19 (9.03) n=77	
Post intervention (4 months)	18.34 (9.24) n=68	18.46 (8.45) n=72	-0.12 (-3.08 to 2.84)
Follow-up (12 months)	16.47 (8.01) n=66	15.86 (7.97) n=69	0.61 (-2.11 to 3.34)
Change in mean scores from baseline (95% confidence interval)			
Post intervention (4 months)	0.31 (-0.94 to 1.57)	0.82 (-0.38 to 2.03)	-0.51 (-2.16 to 1.14)
Follow-up (12 months)	-0.83 (-2.22 to 0.56)	-1.58 (-2.94 to -0.22)	0.75 (-1.11 to 2.61)

SD – standard deviation

In summary, there was no statistically significant difference between arms in GDS-LD scores at 12 months, however both arms had a statistically significant decrease in GDS-LD scores at 4- and 12-months suggesting that both therapies

were successful in reducing depressive symptoms in this population, but neither was more effective than the other.

3.3.3 Economic evaluation methods

The aim of the economic evaluation alongside the BeatIt trial was to estimate the cost-effectiveness of the BeatIt therapy compared to the StepUp therapy for reducing depressive symptoms in adults with mild to moderate intellectual disabilities and a diagnosis of depression. The economic evaluation followed the existing guidance presented in Chapter 2, and the methods are presented below in the following order: 'Data identification and collection - Resource use', 'Data identification and collection - Health benefits' and 'Analysis'.

Data identification and collection - Resource use

The first step in the economic evaluation was to identify resource use categories relevant to the participants and trial. The guidance described in Chapter 2 recommends considering: the perspective of the economic evaluation; resources linked to the treatments and disease being studied; resources likely to differ between arms, and resources linked to treatment side effects, when ascertaining what resource use categories to collect. The resource use categories identified using these considerations are described below:

Perspective – the perspective was the one set out in the NICE reference case (63); an NHS and social services perspective taking into account direct health service use paid for by the NHS and consumed by the participants.

Treatment – the two treatments being compared were therapies delivered by a therapist, who had the support of a supervisor. Resource uses identified and collected were: training of therapists for their allocated therapy; therapist time spent preparing for and delivering the therapy; the time and mileage spent travelling to the participants, and supervisor time. The additional cost of the therapy manuals and materials used during therapy were also collected.

Disease/side effects

Resource use linked to depression and side effects of the therapies could potentially cover a wide range of categories, so healthcare resource use from hospital-based services, community-based services and medications (prescribed and over the counter) were collected.

Between arms difference

The aim of the BeatIt intervention was to increase activity levels in the participants randomised to this arm. Daytime activities considered to be representative of types of activity encouraged by therapists were collected, examples include day centres, sheltered work, and drop-in centres. It was expected that the StepUp therapy would not increase activity levels, thereby resulting in a difference between arms in this resource use category.

Resource use collection

Therapist and supervisor resource use was collected on an ongoing basis throughout the trial using time sheets, these were completed by the therapists and supervisors for each participant. Materials for the BeatIt and StepUp manuals, and materials used during therapy, were established based on expert opinion.

Hospital-based, community-based, and daytime activity resource use was collected using an adapted CSRI form (118, 142). The CSRI was completed at baseline, 4- and 12-months post-randomisation by both the participant and supporter, and at 8-months post-randomisation by the supporter alone. At each resource use collection point the resources used for the previous 4-months (17 weeks) were collected. Medication use was collected on a separate medication inventory which was collected at the same time as the above.

All resource use collected was part of the trial case report form.

Unit costs for resource use categories came from several sources: treatment costs were taken from Personal Social Services Research Unit (PSSRU) (143) for therapist and supervisor time and from expert opinion for the materials for the therapies; daytime activity costs were taken from literature (142); hospital-based resource use costs were sourced from NHS reference costs (144); community-based resource use costs were taken from PSSRU (143), and prescription costs were sourced from the BNF (119). All units costs were adjusted to 2015 prices using the Hospital and Community Health Services Index and reported in pounds sterling (£) (143).

Data identification and collection - Health benefits

The economic outcome measuring health benefits was the preference-based QALY, measured using the EQ-5D-Y questionnaire (77); this outcome is the preferred measure of heath benefit suggested in the guidance reported in Chapter 2. The EQ-5D-Y is a simplified youth version of the EQ-5D-3L questionnaire, it is aimed at children and adolescents aged 8 and above and the language used in it is simple and straightforward but not childlike, so this measure is suitable for adults with intellectual disabilities.

Participant responses extracted from the EQ-5D-Y were valued using the UK adult value set; as yet Euroqol have not provided a UK child value set (145). The resulting utility values were combined with length of life (12 months follow-up) to convert into QALYs using standard area under the curve methods. Changes in utilities between follow-up visits were assumed to be linear when calculating QALYs. Participants completed the EQ-5D-Y questionnaire at baseline and the 4- and 12-month trial follow-up timepoints.

Designing the economic evaluation to reduce 'noise'

As depression could potentially affect many healthcare resources the resource collection included a wide range of healthcare resources. The perspective of the economic evaluation was restricted to the NHS and personal social services. Healthcare resource use covered hospital-based services, community-based

services and medications were collected. The treatment was hypothesised to increase activity so daytime resource use was collected as this could reasonably be expected to differ between arms.

Analysis

As recommended in the existing guidance for economic evaluations an analysis plan was developed, and an early summary of this was included in the published main trial protocol (128).

The analysis was based on the intention-to-treat population, missing data was replaced using the multiple imputation using chained equation method, and results from both complete case and multiple imputation (unadjusted and adjusted) analyses were presented. Mean costs and QALYs in both arms were calculated by summing costs and QALYs for all participants in each arm to estimate total costs, and then dividing by the number of participants in each arm, these treatment arm specific means were presented with standard deviations to illustrate uncertainty. Differences between arms were estimated using the recommended generalised linear models, presented with 95% confidence intervals as a measure of uncertainty. The planned summary measure was the incremental cost per QALY ICER, presented with a 95% confidence interval, cost-effectiveness was assessed by comparing the ICER to the current NICE threshold of £20,000 (55).

Uncertainty was further explored using non-parametric bootstrapping; 1,000 samples were used, and results were presented on a cost-effectiveness plane and a cost-effectiveness acceptability curve. The uncertainty resulting from heterogeneity was adjusted for by using a generalised linear model and including covariates. Covariates used in the cost model were baseline costs and therapist (both significant predictors of costs). Covariates in the QALY model were baseline EQ-5D-Y score and baseline GDS-LD score (both significant predictors of QALYs). Three sensitivity analyses were carried out to explore parameter uncertainty; one looking at the effects of some participants having a mix of 16- and 17-weeks resource use collected compared to only 17 weeks; and two exploring the training costs for therapists; one was the most efficient situation

when training therapists (the maximum possible number of therapists trained), and no training costs.

3.3.4 Economic evaluation results

The economic evaluation results are presented in full in the NIHR report(127), a summary of the results from the original treatment arm-based analysis is presented below, with additional analysis carried out specifically for this thesis comprising treatment and non-treatment costs, additional cost-effectiveness planes and EVPI results.

The amount of missing data was less than 5% for all data categories, and one participant in the BeatIt arm had missing therapist data. Resource use data was available for 58 (85.3%) participants in the BeatIt arm and 68 (93.2%) in the StepUp arm. Quality of life data was available for 61 (89.7%) participants in the BeatIt arm and 68 (93.2%) in the StepUp arm. Missing data was highest in the BeatIt arm compared to the StepUp arm for both resources and quality of life data (Table 6).

Table 6: BeatIt missing economic data

Data	BeatIt (n=68)	StepUp (n=73)	Total (n=141)
Resource use	10 (14.7%)	5 (6.8%)	15 (10.6%)
EQ-5D-Y	7 (10.3%)	5 (6.8%)	12 (8.5%)

Baseline characteristics used in the generalised linear model for the adjusted multiple imputations are presented in Table 7.

Table 7: BeatIt baseline characteristics

Covariates in generalised linear model	BeatIt Mean (SD)	StepUp Mean (SD)
Baseline costs *	£11,980 (£13,686)	£11,582 (£15,753)
Baseline EQ-5D-Y score ^	0.474 (0.438)	0.638 (0.374)
Baseline GDS-LD score ^	16.6 (7.91)	16.9 (6.73)

* covariates adjusted for in cost model; ^ covariates adjusted for in QALY model; SD – standard deviation

Results for treatment arm specific total mean costs and QALYs, and differences between arms for complete cases, multiple imputation (unadjusted) and multiple imputation (adjusted) are presented in Table 8.

Table 8: BeatIt original treatment arm-based economic evaluation results

	BeatIt	**StepUp**	**Difference**	
	Mean (SD)	**Mean (SD)**	**Mean (SE)**	**95% confidence interval**
Complete case analysis				
Total costs	£27,158 (£35,751)	£26,786 (£38,232)	£371	-£12,689 to £13,432
Total QALYs	0.628 (0.361)	0.691 (0.287)	-0.063	-0.178 to 0.052
Multiple imputation (unadjusted)				
Total costs	£27,223 (£4,562)	£26,021 (£4,359)	£1,201	-11,299 to £13,702
Total QALYs	0.617 (0.044)	0.693 (0.034)	-0.076	-0.185 to 0.033
Multiple imputation (adjusted)				
Total costs	£26,369 (£2,382)	£27,962 (£2,347)	-£1,593	- £5,194 to £2,008
Total QALYs	0.657 (0.031)	0.655 (0.029)	0.002 (0.043)	-0.082 to 0.085

SD – standard deviation, SE – standard error

The complete case analysis shows that costs were lower in StepUp compared to BeatIt; £26,786 compared to £27,158, a difference of £371 (95% CI -£12,689 to £13,432), this difference was not statistically significant. Complete case results for QALYs were higher in the StepUp arm compared to the BeatIt arm; 0.691 compared to 0.628, a difference of 0.063 (95% CI -0.052 to 0.178), again a statistically insignificant difference. These results suggest that StepUp dominates BeatIt (less costly and more effective), but there is uncertainty in this result as the cost and QALY differences are not statistically significant. These results suggest that the ICER would fall into the northwest quadrant of the cost-effectiveness plane where decision makers would reject the BeatIt therapy.

As described in Chapter 2 it is good practice to replace missing data in a clinical trial; multiple imputation techniques were used to replace missing data. Unadjusted multiple imputation results confirmed the complete case results; StepUp dominates BeatIt, but these results are still uncertain as the 95% confidence interval crosses zero. Again, the results suggest that the ICER would fall into the northwest quadrant of the cost-effectiveness plane where decision makers would reject the BeatIt therapy.

After adjusting the multiple imputation results for heterogeneity uncertainty using baseline participant characteristics and allocated therapist as covariates, total costs for BeatIt were lower than for StepUp; £26,369 compared to £27,962, a difference of £1,593 (95% CI -£2,008 to £5,194). Adjusted QALYs were higher in BeatIt than in StepUp; 0.657 compared to 0.655, a difference of 0.002 (95% CI -0.082 to 0.085). In this analysis the results have changed direction compared to the previous analyses, suggesting that BeatIt now dominates StepUp. The imbalance at baseline in EQ-5D-Y scores is likely to have driven the reversal of total QALY results. Total QALYs were higher in the StepUp arm for complete case and multiple imputation results, but higher in the BeatIt arm for multiple imputation adjusted results; baseline EQ-5D-Y was significantly higher in the StepUp arm compared to BeatIt. However, it is important to note that despite this change in direction of results, it is still uncertain as neither cost nor QALY differences are significant, with the 95% confidence crossing zero. This result suggests that the ICER would fall into the southeast quadrant of the cost-effectiveness plane where decision makers would accept the BeatIt therapy.

Parameter uncertainty was explored in three sensitivity analyses (Table 9); total mean cost differences in all three scenarios favour StepUp mirroring the complete case and unadjusted multiple imputation results indicating that the unit costs estimates are robust.

Table 9: BeatIt original treatment arm-based economic evaluation - sensitivity analysis results

	BeatIt mean costs (SE)	**StepUp mean costs (SE)**	**Difference (SE)**	**95% confidence interval**
17 week/4 month period for resource use collection	£26,478 (£4,354)	£25,489 (£4,362)	£989 (£6,181)	-£11,238 to £13,216
Nine therapists trained in each training session and nine participants per therapist	£26,428 (£4,354)	£25,451 (£4,362)	£977 (£6,181)	-£11,250 to £13,204
No therapist training costs included	£27,172 (£4,396)	£25,989 (£4,382)	£1,183 (£6,225)	-£11,133 to £13,499

SE – standard error

The three main analyses (Table 8) show no statistical significant differences between treatment arms for mean total costs or QALYs. Whilst there is no evidence of BeatIt being more cost-effective than StepUp, there is also no evidence that it is less cost-effective, this uncertainty is explored further below.

The cost-effectiveness plane presented in Figure 11 illustrates the 1,000 bootstrap samples from the complete case analysis, showing the mean difference in costs and QALYs between StepUp and BeatIt, the samples fall into all four quadrants of the cost-effectiveness plane reflecting the uncertainty in the results, and indicating that there is no evidence to show that one treatment is more cost-effective than the other.

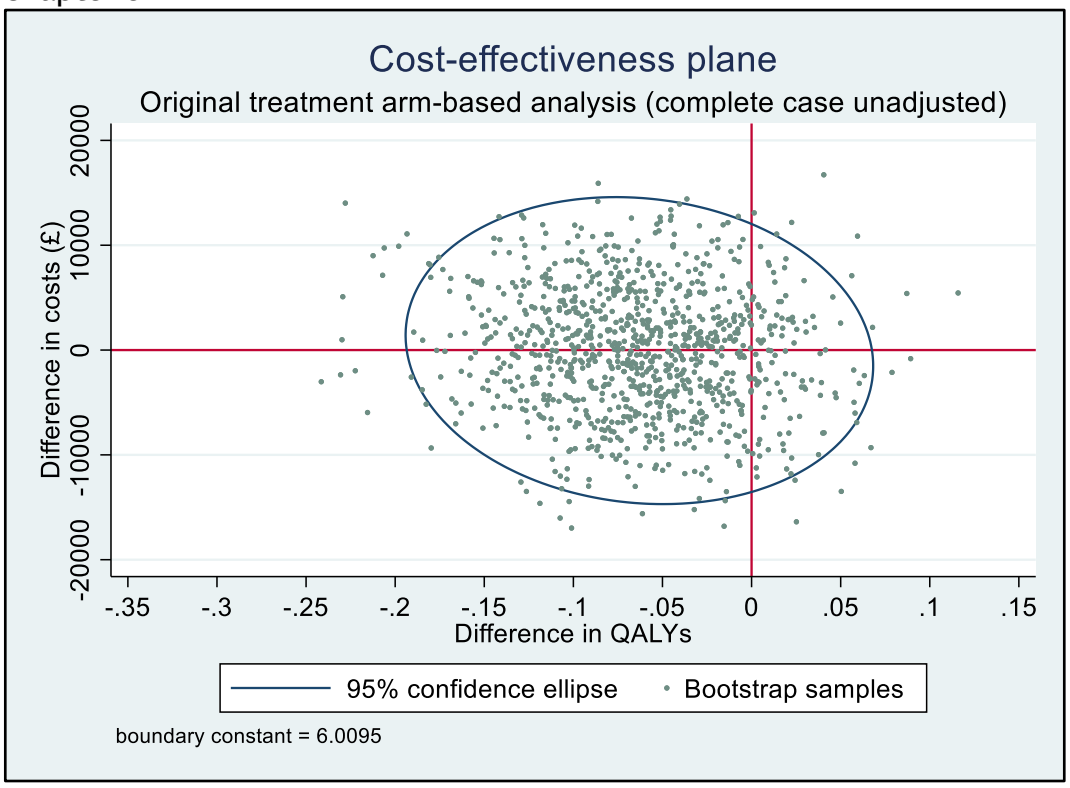


Figure 11: BeatIt cost-effectiveness plane – original treatment arm-based analysis complete cases (unadjusted)

These results are mirrored in the CEAC in Figure 12, illustrating that at the current NICE £20,000 threshold there is less than a 50% chance of BeatIt being cost-effective compared to StepUp. BeatIt is not likely to be cost-effective compared to StepUp at any threshold; at £120,000 the likelihood of BeatIt being cost-effective compared to StepUp is 44.1%.

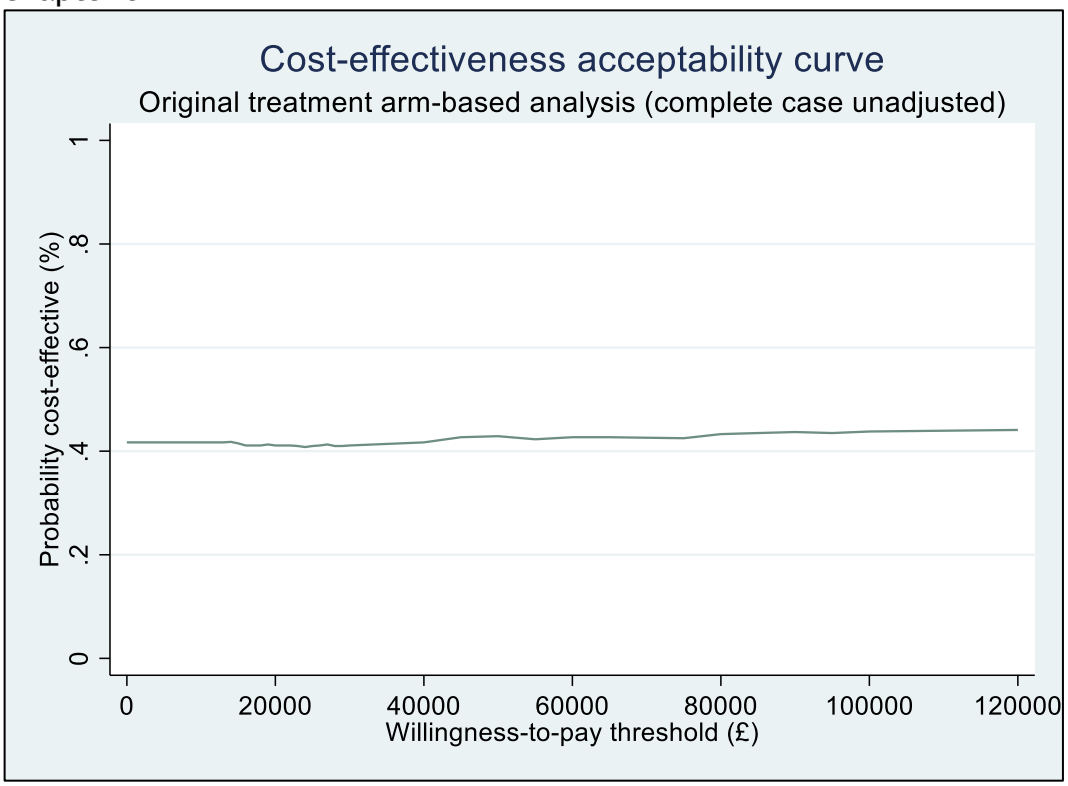


Figure 12: BeatIt cost-effectiveness acceptability curve - original treatment arm-based analysis complete cases (unadjusted)

For the purposes of this thesis additional analyses were conducted on the complete case unadjusted clinical trial data; treatment and non-treatment costs were separated out and the difference between arms estimated for these cost categories, the results were bootstrapped and samples were plotted on cost-effectiveness planes, and EVPI was calculated.

When looking at treatment costs only it is clear that the BeatIt intervention is more costly than the StepUp comparator, there is a difference between arms of £738 (95% CI £586 to £890), this is a statistically significant result; there is minimal uncertainty in the difference between arms for treatment costs. This is illustrated by the cost-effectiveness plane in Figure 13. It can also be seen that whilst BeatIt is more costly than StepUp to deliver, the difference between arms in QALYs is uncertain as the bootstrapped samples cross the vertical axis. This uncertainty in the economic outcomes is reflected in the primary outcome of GDS-LD score.

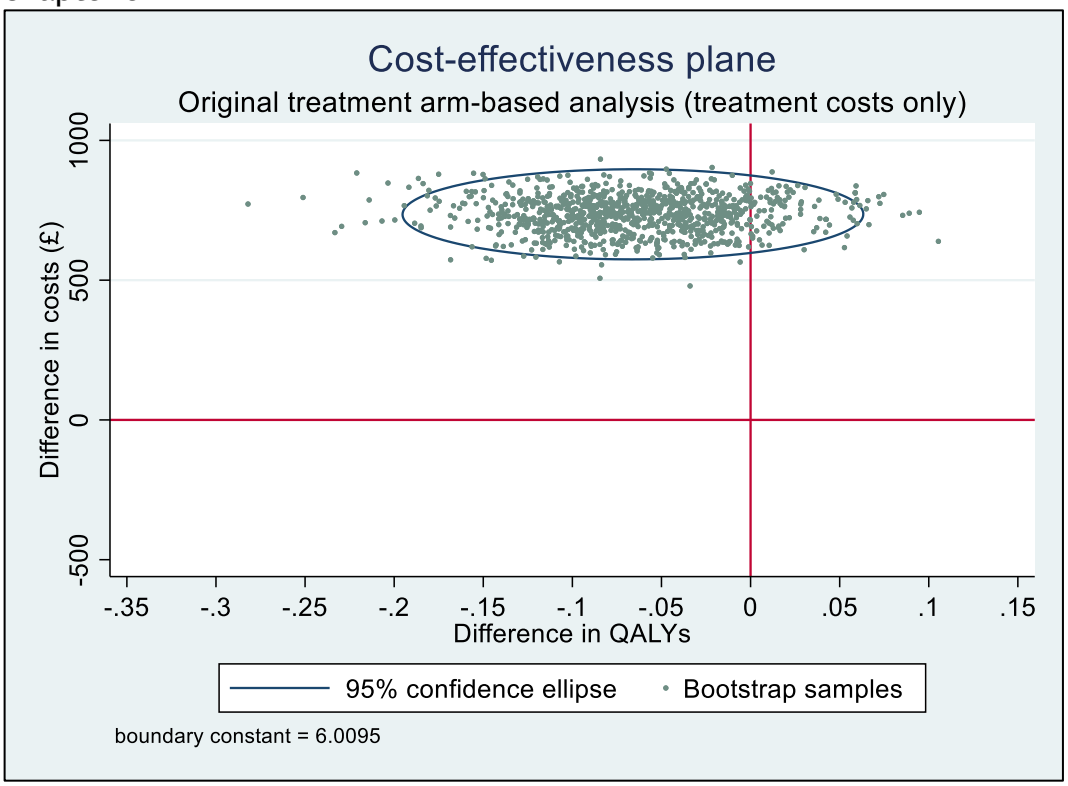


Figure 13: BeatIt cost-effectiveness plane – original treatment arm-based analysis treatment costs only

When the non-treatment costs are assessed they are marginally higher in the StepUp arm compared to the BeatIt arm, a difference of £367 (95%CI -£12,684 to £13,418), however the wide confidence interval suggests that there is little certainty that non-treatment costs are higher in the StepUp arm. This result is illustrated in the cost-effectiveness plane in Figure 14.

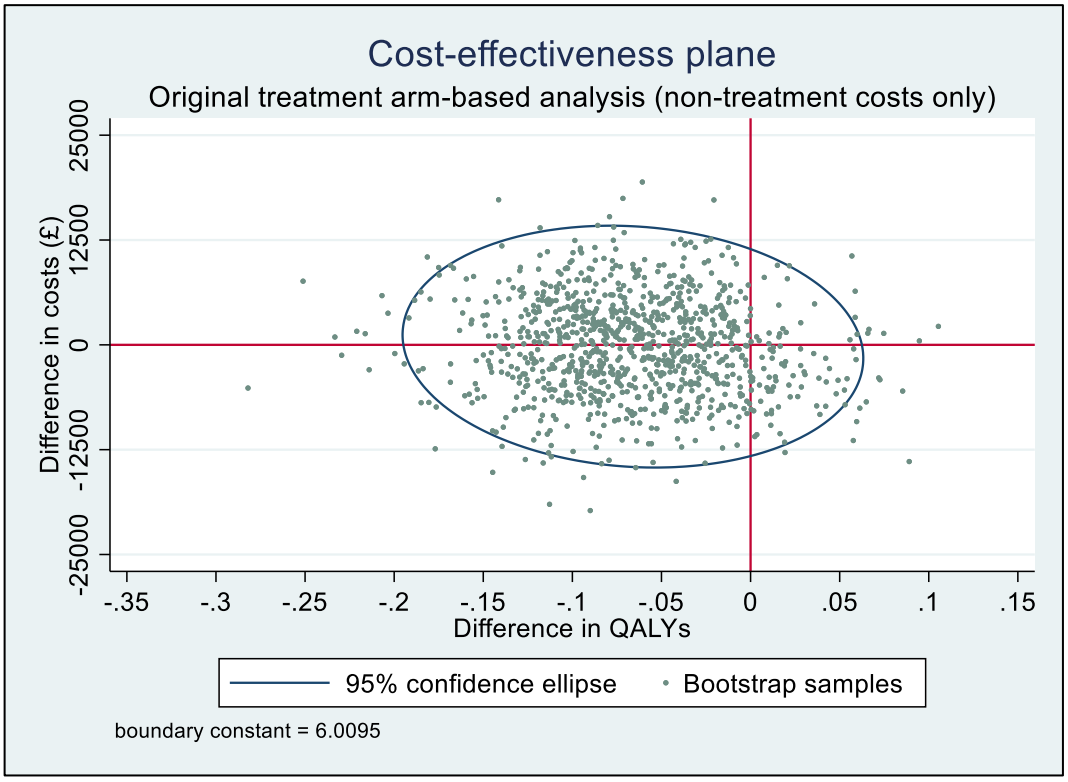


Figure 14: BeatIt cost-effectiveness plane – original treatment arm-based analysis non-treatment costs only

This additional analysis confirms that the treatment costs are known and certain, however non-treatment costs are uncertain, spreading widely either side of the horizontal axis on the cost-effectiveness plane.

The EVPI estimate is £1,679 per person suggesting that further research may be worthwhile; additional research would be worthwhile if its cost is less than the population EVPI cost. This result is driven by the uncertainty in the results, as illustrated on the cost-effectiveness planes in this section; only the cost-effectiveness plane for treatment costs shows a clear difference in costs.

The results of the economic evaluation indicate that there was no difference in outcomes between arms, which was mirrored by the lack of difference between

arms in the primary outcome of change to GDS-LD in the trial (126). Although there was no significant difference in mean total costs between arms, this masks a statistically significant difference in treatment costs.

3.4 Summary

This chapter introduced the two case studies which will be used throughout this thesis. In chapter 2 the limitations of using clinical trials to assess the cost-effectiveness of a new technology were discussed, and the uncertainty in economic evaluation results as a consequence of these limitations was highlighted. The purpose of the case studies in this chapter were to illustrate how the existing guidance for a standard treatment arm-based economic evaluation is applied in practice, and to demonstrate the uncertainty which can result from using trial data in an economic evaluation. In these case studies the recommended methods for showing uncertainty were followed: 95% confidence intervals were presented; bootstrapped samples were used to populate cost-effectiveness planes and cost-effectiveness acceptability curves; value of information was calculated; heterogeneity was adjusted for using baseline characteristics, and sensitivity analyses were conducted. The case studies both show uncertainty around the ICER, and because of this uncertainty there is limited evidence for decision makers to base a decision on.

Additional analysis conducted specifically for this thesis found that in both case studies the treatment costs were known and precise, however the other costs (except exacerbation costs in TWICS) were unclear and this uncertainty was masking the certain treatment costs.

In the next chapter suggestions for developing a methodological framework are reported, these are then applied in Chapter 5 to develop a methodological framework for developing conceptual models.

Chapter 4 Scoping review of methodological framework development

4.1 Chapter overview

The aim of this chapter is to provide suggestions for developing a methodological framework, this is the first output of the thesis and partly fulfils the second objective.

This thesis introduces a new role for conceptual models to provide additional understanding for decision makers to inform funding decisions. While guidance exists for developing conceptual models in the field of economic evaluation, this focusses on using conceptual models to guide the development of the structure of decision analytic models. New guidance is needed to develop conceptual models in the proposed new role and the chosen format of this guidance is a methodological framework. At the outset of this thesis there was no consensus on recommendations, or established methods, for developing a methodological framework (39), and due to this lack of guidance an extra work strand was added to the thesis; to compile practical suggestions for developing a methodological framework.

The objectives of this chapter were to:

1) identify existing methodological frameworks

2) extract reported approaches used to develop the methodological frameworks

3) collate and summarise the approaches into suggestions for developing methodological frameworks, and

4) explore terminology used for methodological frameworks.

A scoping review was used to achieve the objectives set out above, methods of the scoping review are described in Section 4.2, results of the scoping review are set out in Section 4.3, a discussion of the work in this chapter is included in Section 4.3.44.4, and a summary is presented in Section 4.5.

Chapter 4

4.2 Methods

4.2.1 Scoping review methodology

A scoping review was chosen over a systematic review to provide the suggestions for developing methodological frameworks. Systematic reviews answer a clearly defined research question, and are typically used to inform current practice with robust and reliable estimations of treatment effects (146). Scoping reviews ascertain the extent of information in a given subject area, particularly when evidence on a topic has not previously been comprehensively reviewed, and in emerging areas (147). Whilst scoping reviews can be used to determine whether a full systematic review would be worthwhile, they can also be an exercise in their own right, as in this piece of work, where the purpose is to map the body of evidence on a topic and explore the extent of this evidence (148, 149). Scoping reviews identify the nature and range of the evidence, and summarise, using charts or maps, the research findings, providing a descriptive overview of emerging evidence (149, 150). Systematic reviews follow rigid inclusion and exclusion criteria decided a priori and critically appraise evidence identified (151). A scoping review includes a narrative review using an analytic framework or thematic review, without critical appraisal or synthesis (150). Unlike systematic reviews, which synthesise evidence from studies, scoping reviews ascertain the extent of information on a given subject area (149). The purpose of this chapter was not to answer a precise clinical research question, either by providing new evidence or refuting existing evidence; the purpose of this chapter was to map new and emerging evidence on methods used to develop methodological frameworks, from a wide range of sources, and to summarise this evidence.

Guidance for conducting scoping reviews was first published by Arksey and O'Malley in 2005 (150). This was updated by Levac et al. in 2010 (152), and again in 2014 by Colquhoun et al. who published clarification on definitions, methods and reporting (153). In 2015 Peters and colleagues at the Joanna Briggs Institute, published updated guidance on conducting scoping reviews (149). In 2016 Tricco et al. carried out a scoping review on the methodology and reporting of scoping reviews, comparing the methodology used in the scoping reviews to the guidance presented by the Joanna Briggs Institute in Peters et al. (147).

For my scoping review I followed guidance by Arksey and O'Malley, supplemented by the most recent guidance on conducting a scoping review in Levac et al. (152), Armstrong et al. (148), and Peters et al. (149). The remainder of this section reports a summary of this guidance and sub-sections 4.2.2 through to 4.2.6, explain how this guidance is applied in my scoping review.

Identifying the research question

The first step is to identify the research question; the research question guides the search strategy and it should be broad enough to ensure a wide range of coverage (148). The purpose and rationale of the scoping review should also be considered at this stage (150).

Identifying relevant studies

The search should be as broad and comprehensive as possible, including published and unpublished studies and reviews (150). The strategy adopted by Arksey and O'Malley involved searching various sources: electronic databases, reference lists, hand-searching of key journals and existing networks, relevant organisations and conferences (150). Levac et al. advise using the research question and purpose of the review to guide the scope of the search (152). Peters et al. divide this stage into three steps: a limited initial search to identify key text words and index terms to use in the search; a second comprehensive search using key words and index terms identified in the initial step, followed by searching the reference lists of record (149).

Study selection

Inclusion and exclusion criteria are developed on an ongoing basis, not prior to the search, based on increasing familiarity with the literature (150). Levac et al. recommend that the study selection should be an iterative process including searching the literature, refining the search strategy and assessing articles for inclusion. Both Arksey and O'Malley and Levac et al. recommend using two researchers to independently review titles/abstracts and full articles.

Chapter 4

Charting data

Arksey and O'Malley describe this stage as 'charting' key items of information from the included studies (150). These key items should be entered onto 'data charting forms'; Arksey and O'Malley used Excel software for developing their forms and suggest collecting both general and specific information from the study. Armstrong suggests using a spreadsheet or database to chart the data, and using the research question as a focus, this helps the researcher(s) in identifying commonalities, gaps and themes; the focus of the scoping review will determine what data is charted during this stage (148).

Examples of the type of data to extract include: author/s, year of publication, location of study, intervention, comparator, duration of intervention, study population, aims of the study, methodology, outcome measures and important results.

Levac et al. add that charting should be an iterative process where the researchers continually extract data and update the data charting form (152). Peters et al. also state that refinement of the charting forms may be needed and that the results should be presented as a map in a logical, diagrammatic, tabular and/or descriptive form, emphasising that the charted results chosen should align with and illustrate the aims and purposes of the review.

Collating, summarising and reporting the results

Peters et al. recommend a narrative description of the search process and identified records, plus a flowchart illustrating the process which depicts decisions on exclusion and inclusion and includes the final number of items of included. The narrative review should include a numerical analysis of the included records, including the extent, nature and distribution of records. The numerical analysis can include tables and graphs.

Levac et al. developed this stage further by suggesting three distinct steps: analysing the data, reporting results, and applying meaning to the results. In analysing the data using thematic analysis, Levac et al. argue that this stage resembles qualitative data analytical techniques; researchers should think about

using qualitative analytical techniques and qualitative software to enable this process. In the second stage, 'reporting results' the researchers should consider the best way to report the results, for example using themes, a framework or a table. In the third stage the researchers should consider the implications of the results and apply meaning to the findings of the scoping review within the broader research, policy and practice context.

Discussion

A detailed discussion of the results and limitations of the sources, if relevant, should be included. Results should be considered in the context of current literature, practice and policy.

Conclusions and implications for research and practice

Conclusions should match the review question/objective, starting with an overall conclusion based on the review results. Recommendations for future research based on gaps identified in the review should be discussed, however, recommendations for future practice may not be able to be developed depending on the aim and focus of the review.

The following sections describe the methods used in this present scoping review, based on the methods just explained.

4.2.2 Identifying the research question

The primary research question for this scoping review was 'What reported approaches are being used to develop methodological frameworks?' The working definition of a methodological framework for this scoping review was 'a structured practical tool for guiding the user through a stepwise process, enabling/facilitating a standardised approach to the given task'. As there is no formal terminology for a methodological framework (45) a secondary research question of 'What terminology is used to define methodological frameworks?' was included.

The purpose of the review was to identify reported approaches used to develop methodological frameworks to inform the development of my methodological framework for developing conceptual models later in the thesis. The rationale for conducting the review was that there is no formal guidance on developing methodological frameworks, nor any consensus on the terminology used for describing methodological frameworks.

4.2.3 Identifying relevant studies

The search strategy was developed with input from the College librarian, both for search terms and databases.

Following Peters et al. guidance a simple initial search was conducted. This initial search was conducted in Web of Science because this database covers a broad range of disciplines and topics, and it also allows the user to easily check citations of identified papers. The initial search, using the broad terms of 'framework' AND 'develop*', resulted in 11,120 potential records and the results of the initial search helped to inform the more comprehensive search. Index terms were scrutinised, this technique is particularly useful when there is no firm consensus or consistency on definitions and terms, as in this case.

In Tricco's review of scoping reviews, 93% of reviews searched more than one database, only 6% searched one database (147). The comprehensive search followed Tricco's research methods and Arksey and O'Malley's methods described in sub-section 4.2.1, two databases were searched to identify peer-reviewed publications, and the internet was searched to identify grey literature. Grey literature has been described as not being under the control of commercial publishers, and examples of grey literature include; government reports, PowerPoint presentations, best practise documents, newsletters and working papers (154). The search took place in September 2018.

After the initial database search returned 11,120 results, to ensure that a more manageable number of results were identified in the comprehensive database search, search terms were restricted to 'methodological framework' plus terms for development. Details of search terms used are included in 0Appendix 2: Scoping review search terms (*Chapter 4)*. The two databases searched were Web

of Science and Ovid Medline, and only titles rather than abstracts were searched to ensure that the search terms were the main focus of the article or paper.

There is no gold standard for searching grey literature, however, research has shown that using internet search engines can identify up-to-date grey literature (155). I conducted an internet search which was guided by Godin's methods; searching the first 10 pages (100 hits) of results, and bookmarking websites that looked relevant (156). Godin also targeted specific websites, but as this search was not restricted to specific disciplines, this more targeted aspect of Godin's methodology was not used.

The internet search was carried out using the 'Google' web browser, the search term used was 'Methodological framework development'. As methodological frameworks are often presented as a diagram I also conducted a separate search in Google Images; this replicated methods presented by Rivera et al. (39), and based on methods from Rivera et al. I screened the first 50 items.

Following the three-step method recommended by Peters et al., the final stage was to search the references and citations of the methodological frameworks which met the inclusion criteria, for any additional relevant methodological frameworks.

4.2.4 Study selection

For the purposes of this search broad inclusion criteria were used, these included: a methodological framework which should be presented in the article and the approaches used in developing the methodological framework should be described; only English records were included, and only records published in the last ten years were included - from 2008 onwards. Titles and abstracts were screened, then either discarded or kept and read in full, and the full texts that met the inclusion criteria were selected and included in the review results.

4.2.5 Charting the data

In line with the guidance described in sub-section 4.2.1, Excel was used to extract study characteristics, these were: author/s; title; date of publication; type of record (for example published journal/conference proceedings);

discipline; the country of origin of the study; reported approaches used in developing methodological framework; terminology used to describe the methodological framework, and any keywords used.

To extract reported methodological approaches an iterative process was used, as recommended by Levac et al., for charting the data; initially basing the extracted the themes reported in Rivera et al. (39). These original themes were: original source of the methodological framework; literature review; stakeholder involvement; incorporating stakeholder views, and piloting phase. These themes were added to and updated as approaches were extracted and I became more familiar with the methodological frameworks and the range of approaches taken.

4.2.6 Collating, summarising and reporting the results

The search results, including the inclusion/exclusion process, were illustrated in a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram, and the characteristics of the studies identified and included in the results were presented narratively and graphically. Reported approaches were examined in detail and presented descriptively with counts and percentages. Meaning was applied to the results by considering their implications and further categorising themes into phases and interpreting to make suggestions for developing methodological frameworks. Finally, terminology used for the methodological frameworks was described narratively, numerically and visually, and keywords used for the studies were summarised.

4.2.7 Discussion and conclusions

The results were summarised and compared to existing literature as far as was possible, strengths and limitations were discussed, and recommendations for future research were suggested.

4.3 Results

The results are presented in four sub-sections: the first reports the literature search results (sub-section 4.3.1), the second describes the study characteristics (sub-section 4.3.2), the third presents the suggestions for developing

methodological frameworks (sub-section 4.3.3), and the fourth describes the terminology used in the methodological frameworks (sub-section 4.3.4).

4.3.1 Literature search

The results of the search are presented in a PRISMA diagram in Figure 15. In total 320 records were identified, 54 of these were duplicates, leaving 266 titles and abstracts to screen. 87 papers were excluded after reading titles and abstracts, and a further 149 were excluded after reading full texts; 30 papers were included in the final search results.

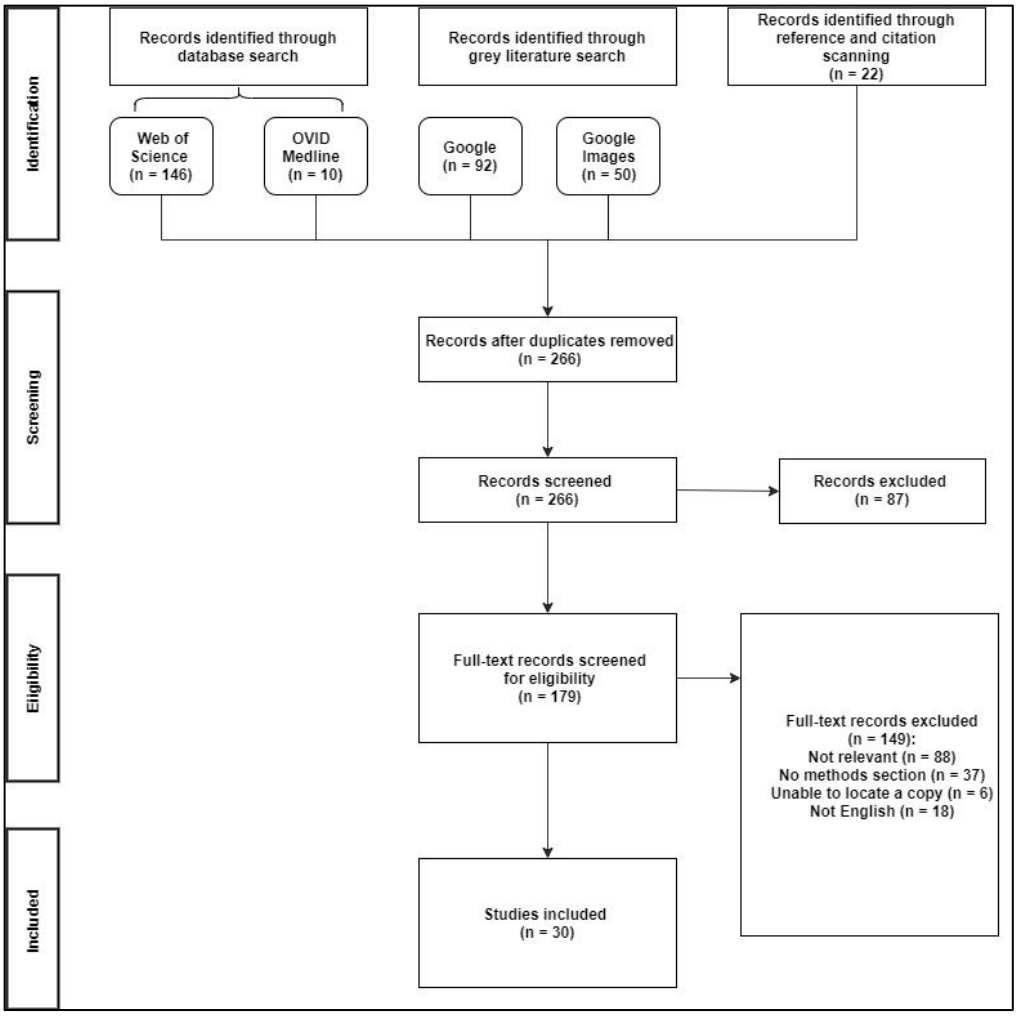
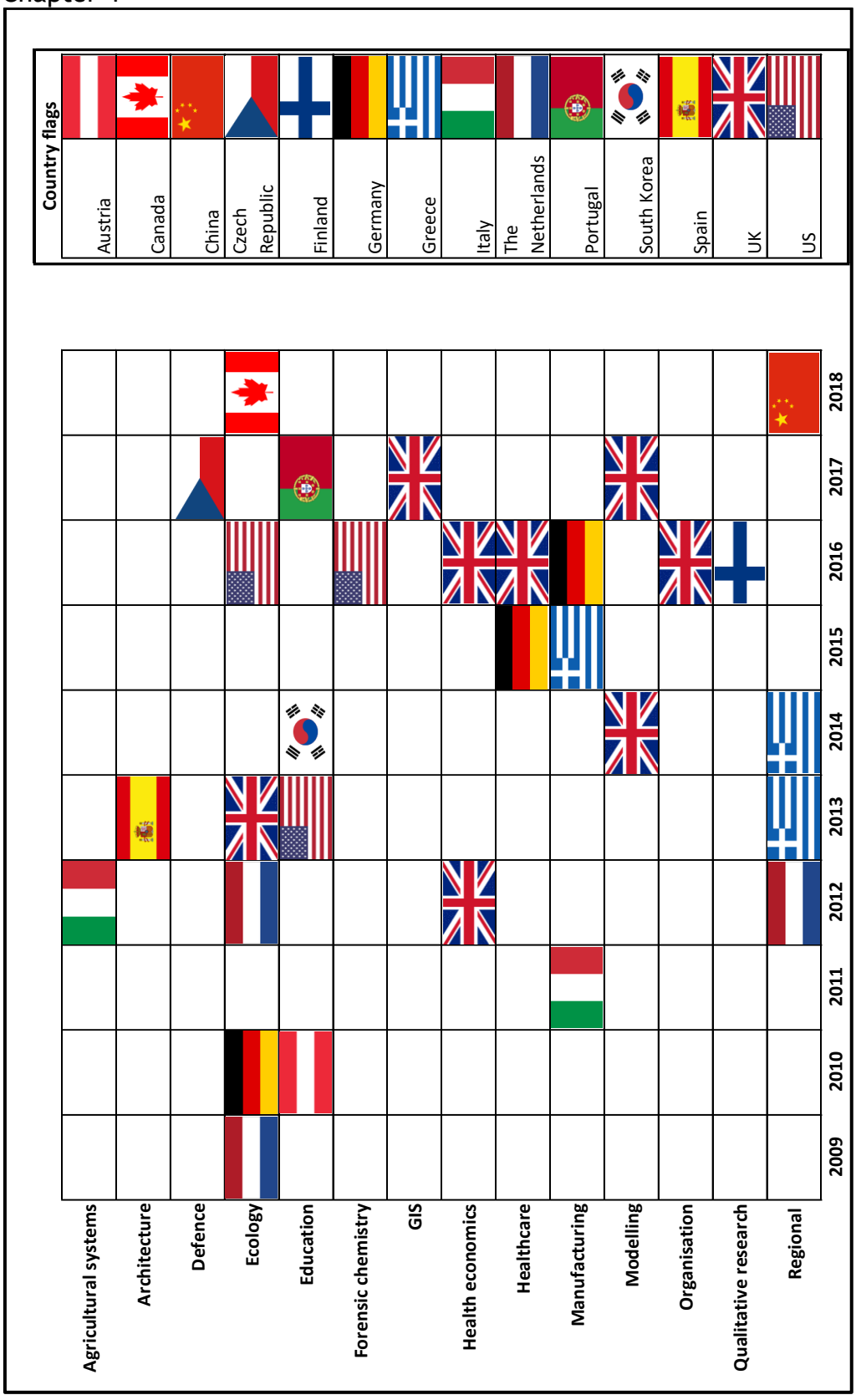


Figure 15: PRISMA diagram, scoping review results for methodological frameworks (duplicated from McMeekin et al. 2020 (45))

4.3.2 Study characteristics of included methodological frameworks

Study characteristics are presented in full in Appendix 3: Results of scoping review – study characteristics (*Chapter 4*) and summarised in Figure 16. Most studies were published in recent years; years of publication ranged from 2009 to 2018, 12 were published between 2009 and 2013, and 18 were published between 2014 and 2018. Most studies were journal articles (n=26), three were conference proceedings and one a book chapter. The studies represent a wide range of subject areas; 20 fields were identified, the most frequent was ecology (n=6), then education and regional (n=4), next manufacturing (n=3), and healthcare, architecture and health economics (n=2). The was one study each for agricultural systems, architecture, defence, forensic chemistry, geographic information system, organisation and qualitative research. The papers originated from 14 countries; the country where most were published was the UK (n=8), then Greece, Germany, The Netherlands and the United States (US) (n=3) and finally Italy (n=2). The remainder of the countries had one paper each. Most of the papers originated from Europe (80%), then America/Canada (13.3 %) and lastly Asia (6.7%).



GIS - geographical information system, UK – United Kingdom, US – United States

Figure 16: Summary of study characteristics (year of publication, field, and country of origin)

4.3.3 Extracted reported approaches

Following the iterative process of extracting and categorising reported approaches discussed in sub-section 4.2.6, eight reported approaches were extracted, results are presented in Table 10.

Table 10: Methods reported as being used in methodological framework

Reported methods	Number (n=30)	Percentage (%)
Based on existing methods	20	66.7
Refined and validated	10	33.3
Experience and expertise	9	30.0
Literature review	8	26.7
Data synthesis and amalgamation	7	23.3
Data extraction	3	10.0
Iteratively developed	2	6.7
Lab work results	1	3.3

The most frequently reported method was ‘based on existing methods’ (n=20, 66.7%), this included the use of existing methods and guidelines as the foundations of the methodological framework being developed. Existing methods and guidelines included previous methodological frameworks or guidance and published methodology. Whilst some studies did not explain how the existing methods and guidelines formed the foundations of the methodological framework being developed, several did expand on this further: adapting the methods (31, 157), integrating methods (158, 159), building on the existing methods (37), based on the methodological framework, combining well established guidelines which comprised the same stages (34), and the methodological framework was the basic inspiration (36). Only one study reported how the existing methods were identified; Squires et al. used a literature review (22).

Ten studies (33.3%) reported ‘refined and validated’ as an approach. There were two aspects to this approach: refining the methodological framework and validating the methodological framework. Specific approaches taken were:

piloting the methodological framework (160), trialling identified stages and using the results of the trial to further develop the methodological framework (161), using a case study to refine the methodological framework (162), using a case study or Delphi panel to evaluate and refine the methodological framework (22, 42, 162), using a case study to validate to methodological framework (28, 33), and testing the methodological framework (32). Two studies did not report details of the case study (157, 163).

Nine (30%) studies reported using 'experience and expertise' to develop the methodological framework. This included experience at a personal (35), school/university (161) and country level (36). Experience and expertise was collated during meetings (163), using consultations (37) and in collaboration (162). Two studies did not specifically mention experience but used surveys and interviews (164), and focus groups for extracting expertise (22).

Eight (26.7%) studies reported conducting a 'literature review'. Two papers used purposeful sampling (38, 165), sources used for searches included databases, dissertations (30), library catalogues, contacting key authors, databases, websites and citations (42, 166). Six studies included a literature review but did not report specific methods used (22, 28, 30, 42, 160, 162).

Seven studies (23.3%) reported using 'data synthesis and amalgamation'. Specific methods included: identifying phases (38), themes (22, 164) and dimensions (30), analysing and grouping or categorising themes, and thematic analysis (30, 38, 42, 165, 167).

Three studies (10%) reported using 'data extraction' as an approach. This included extracting data from interviews and focus groups using transcribing methods (22, 164), and extracting key information from published literature (38).

'Iteratively developed' was reported in two studies (6.7%), one study had no further details on this (32), the other explained that the methodological framework evolved and developed as items were extracted, synthesised and revised.

The least frequently mentioned approach was ‘lab work results’; one study from the field of explosives reported this approach; the results of lab tests were used to inform the methodological framework (168).

In applying meaning to, and considering the implications of, the results, a pattern emerged whilst reviewing the approaches and they were split into three categories. The first category related to identifying evidence or data to inform and form the foundations of the proposed methodological framework. This evidence came from: existing methods, literature reviews, lab results and experience/expertise. The second category related to developing the proposed methodological framework, comprising: extracting data, and synthesising and amalgamating this data iteratively. The third and final category was refining and validating the proposed methodological framework using: trialling the framework with pilot or case studies and or Delphi panels. These categories are illustrated in Figure 17, and described below presenting suggestions for developing a methodological framework.

Phase 1 – Identifying evidence to inform the methodological framework

This phase describes two approaches for identifying evidence: identifying existing evidence and identifying new evidence.

Existing frameworks or guidance can be identified using purposeful literature searches and evidence includes previous methodological frameworks, guidance or published methodology, and can inform the proposed methodological framework by adapting, integrating and building on the existing methods.

New evidence comes from several sources: the experience and expertise of experts at a personal, organisational or country level; qualitative research (surveys, interviews and focus groups), and collaboration and consultation with interested experts. If qualitative research is used it is preferable that experts in the field of the proposed methodological framework are used for this research rather than restricting the evidence to the experience of the developers of the methodological framework.

Phase 2 – Developing the methodological framework

In this phase the evidence identified in ***Phase 1*** is analysed, adapted, combined and built upon to create the foundation of the proposed methodological framework. Firstly, key information is extracted using appropriate methods: transcribing qualitative data; populating predesigned tables and entering quantitative evidence into piloted data extraction forms. Then the extracted data is analysed by synthesising, grouping or amalgamating it into categories, these categories then inform the proposed methodological framework. This phase is iterative; the proposed methodological framework evolves after it is presented to key experts and study team members who will review it and revisit evidence from ***Phase 1*** until a consensus is reached.

Phase 3 – Evaluate and refine

In this final stage the proposed methodological framework is evaluated and refined. Evaluation techniques include case studies to pilot the methodological framework and Delphi panels. If appropriate, the results from this evaluation can be used to refine the proposed methodological framework; updating it with changes identified during the evaluation and then presenting the revised proposed methodological framework to the key experts and study team members.

Figure 17: Three categories of reported approaches for framework development

4.3.4 Terminology used for methodological frameworks

Terminology used to describe the methodological frameworks included in the scoping review was extracted and the results are presented below.

The studies included seven different terms to describe methodological frameworks, in both the titles and the main text of the studies: ‘methodological framework’; ‘framework’; ‘conceptual framework’; ‘sequential framework’; ‘theoretical framework’; ‘governance framework’, and ‘problem-oriented framework’.

Most studies (n=24, 80%) used the term ‘methodological framework’ in the title. Of the remaining six studies one included ‘methodological’ and ‘framework’ separately in the title, four included only ‘framework’ in the title and one used the term ‘conceptual framework’.

Within the main text of the study most studies used the seven identified terms interchangeably, this is illustrated in Figure 18. A combination of terms was used in most studies; 19 used a combination of ‘methodological framework’ and ‘framework’, three used ‘framework’ only, two used a combination of ‘conceptual framework’, ‘methodological framework’ and ‘framework’, one used ‘methodological framework’ only, the remaining five all used a combination of two or four terms to describe the methodological framework.

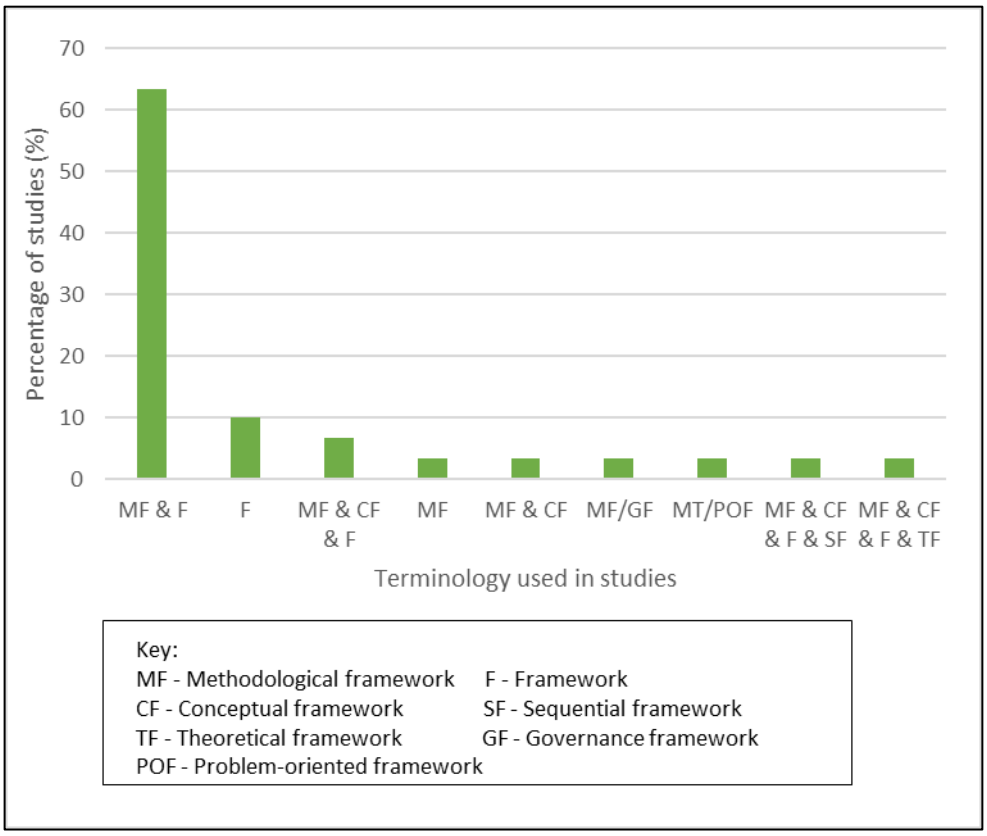


Figure 18: Terminology used in studies

Keywords extracted from the studies which are relevant to a methodological framework, are presented in Table 11. 15 (50%) studies had no keywords relevant to methodological frameworks, using only keywords relevant to the actual subject of the methodological framework. Studies that included keywords relevant to methodological framework used; 'methodology' (4/30), 'methodological framework' (3/30), 'design methodology' (2/30), 'simulation methodology' (1/30), 'methods' (1/30) and 'guidance' (1/30). Ten studies included one relevant keyword, and one had two relevant keywords ('methods' and 'guidance').

Table 11: Keywords relevant to methodological frameworks

Keyword	**Number (percentage) (*n*=30)**
None	15 (50%)
'Methodology'	4 (13.3%)
No keywords	4 (13.3%)
'Methodological framework'	3 (10.0%)
'Design methodology'	2 (6.7%)
'Simulation methodology'	1 (3.3%)
'Methods'	1 (3.3%)
'Guidance'	1 (3.3%)

A PRISMA checklist is included in Appendix 4: PRISMA Scoping review checklist (Publication from Chapter 4).

4.4 Discussion and conclusions

4.4.1 Overview

The scoping review reported in this chapter identified 30 methodological frameworks published in the last 10 years and reported the approaches that were used to develop them. The studies covered 20 disciplines and originated from 14 countries; the reported approaches were extracted and grouped. Eight reported approaches of development were identified and extracted from the frameworks, however, not all of the methodological frameworks identified in the scoping review reported the approaches used in developing them; 179

potentially relevant methodological frameworks were scrutinised in full and 37 (20.7%) of these were rejected because the authors did not report the approaches used to develop them. Studies which did report the approaches used for development were often not clear, however there was enough consensus in the common approaches to allow me to apply meaning to the results. The reported approaches form the basis of the suggestions for developing methodological frameworks presented in Figure 17.

Of the studies which did report the approaches used to develop them the number of studies reporting one or more of the eight approaches ranged from 20 ('based on existing methods and guideline') to one ('lab work results'). The number of mentions of the remaining six approaches ranged from two to ten. Whilst 'based on existing methods and guidelines' and 'literature review' were presented separately in the results, it could be argued that these two approaches are closely linked: a literature review is carried out to identify previous methods and guidelines. However, if a literature review is not carried out, it is entirely feasible that the framework will be based on previous methods that were not identified through a literature review.

The number of terms used to describe methodological frameworks highlights the lack of clarity around terminology. Most studies used a combination of 'methodological framework' and 'framework', this is understandable when there are often word limits for journals and restricting the terminology to 'framework' can help to improve the flow of the discussion. However, two studies used a combination of four terms to describe the methodological framework further illustrating lack of clarity between and within studies. This confusion in terminology is likely to lead to difficulties in identifying methodological frameworks in a literature search as many methodological frameworks might not be identified.

In addition to the lack of clarity in terminology, half of the identified studies did not include keywords relevant to methodological frameworks. This is likely to be the result of the subject of the methodological framework being the focus of the study rather than the actual methodological framework. Again, this could lead to difficulties in identifying methodological frameworks in a literature search.

4.4.2 Comparison to existing evidence

As there was no existing guidance for developing methodological frameworks it was not possible to interpret the results of this scoping review considering what is already known. However, in the following paragraphs I compare the results to Rivera et al., this comparison is particularly useful as there was no duplication of studies between my scoping review and the Rivera et al. literature review.

Rivera et al. reported that the methodological frameworks identified in their research varied in their development, but that there were some common approaches. Only one paper (4%) did not report any approaches used in developing the methodological framework compared to 37 (20.7%) in my search. Rivera reported four key approaches reported: using a literature review, consulting with stakeholders using interviews or surveys etc, methods to incorporate stakeholder views (thematic analysis, refining and feedback) and piloting the methodological framework to refine or illustrate it.

One approach identified by Rivera et al. that was not identified in my search was one used during the validation stage; comparing the methodological framework to existing literature to assess consistency. Whilst this was not identified in the search or included in the results, I believe that comparing a version of the framework with existing literature in a discipline is a useful approach which would result in a methodological framework that has been evaluated and refined using tried and tested methods and will consequently be more robust.

The results from my scoping review also identified approaches not identified by the Rivera et al. research, these were: refined and validated; data synthesis; data synthesis and amalgamation, and iteratively developed. Overall, I was able to extract more details on ‘how to do’, rather than ‘what to do’.

My scoping review has moved understanding forward by adding to the basic approaches extracted by Rivera et al., whose aim was to identify the impact of research and the approaches identified were extracted as part of reporting study characteristics. The research is useful because whilst Rivera et al. reports approaches used to develop methodological frameworks there is no information

on how to apply these approaches; for example conducting a literature review but not reporting what to do with the information from the literature review, for example data extraction or synthesis.

Finally, Rivera et al. concludes that the lack of guidance on developing methodological frameworks should be addressed to ensure that best practice methods can be used in the future. This scoping review starts to address this lack of guidance in this area and provides a foundation for much needed future research to develop this area further; a systematic literature review with more resources should be carried out to identify more methodological frameworks, and guidance could be further developed by using expert feedback and piloting. Another recommendation for future research is to develop a standardised procedure for collecting qualitative data evidence in phase one of the suggestions.

4.4.3 Strengths and limitations

To the best of my knowledge this is the first research done with the aim of identifying approaches reported for developing a methodological framework. This research attempts to fill a gap where there is a need for guidance in developing methodological frameworks.

The methodological frameworks identified are from many different contexts and countries, these demonstrate natural variation and give the suggested approaches a robust and generalisable nature. This research also provides a contemporary overview of how methodological frameworks are being developed.

Limitations of the scoping review mainly relate to the search strategy; restricting the search terms to titles, not abstracts ensured that the search terms were the main focus of any identified articles, however this may have excluded other relevant articles. Additionally, some frameworks may have missed in the search due to inconsistent terminology, in fact, none of the methodological frameworks identified by Rivera et al. were identified in this scoping review, this is because I used the search term ‘methodological framework’, not ‘framework’. Rivera et al. was able to use ‘framework’ for a search term as their search was specifically for research impact papers which

would have reduced the number of potential hits compared to my broader search. Some methodological frameworks will have been missed in the search due to restricting the search terms to 'methodological framework' because of pragmatism; if the term 'framework' had been used this is likely to have resulted in an impractical number of results.

Limitations of internet searches include; the search reports thousands of hits, when in reality far fewer are found; the transient nature of the internet, where documents may disappear over time; and the 'personalised search feature' element of an internet search making the search difficult to replicate, although Adams et al. concluded that replicating an internet search was feasible (155).

Not all methodological frameworks identified reported approaches taken; out of 179 potentially eligible methodological frameworks (met all criteria excluding including approaches), 37 (20.7%) were rejected because they did not report the approaches taken in developing the methodological framework, this limited the amount of data I could extract and include in the scoping review. Linked to this not all approaches were clearly reported, perhaps because of word count, aim of the paper, or traditionally how different disciplines report, so I needed to interpret them.

Arksey and O'Malley and Levac et al. recommend using two researchers to independently review titles/abstracts and full articles, however as this was a piece of work for my PhD only one researcher was involved.

Scoping reviews do not assess the quality of included evidence, therefore there is a risk that the frameworks included in this review were not of high quality. However scoping review methodology was the correct choice for this review; scoping reviews are recommended for identifying key characteristics related to a concept, which is useful when giving an overview of the evidence. Furthermore, scoping reviews use rigorous and transparent methods ensuring that the results are trustworthy (146).

Finally, the second objective in my thesis was to develop a methodological framework to guide the development of a conceptual model in economic evaluation, and it is possible that not all approaches identified and extracted in

this scoping review are suitable for developing a methodological framework specifically for developing a conceptual model.

4.5 Summary

This chapter used scoping review methods to produce suggestions for developing methodological frameworks, this is the first output of this thesis. The scoping review found that there was consistency and unspoken consensus on approaches used and a three-phase suggestion was produced to inform the development of a methodological framework. Not only does the research from this chapter fill a previously identified gap, but it also feeds into the second objective of my thesis which is to develop a methodological framework for conceptual models in economic evaluation. Future research could update the results of this scoping review by using a systematic literature review to develop guidance, and evaluate this guidance using suitable methods to come to a consensus on the contents of the guidance. A standardised procedure for collecting qualitative data evidence in phase one should also be investigated. I would also recommend the use of the term 'methodological framework' as a minimum in the titles of future methodological frameworks and as a keyword, this would enable future searches to identify all relevant evidence.

The suggested approaches reported in this chapter feed into Chapter 5, where they are applied to develop a methodological framework for developing conceptual models.

Chapter 5 A methodological framework for conceptual models in economic evaluation

5.1 Chapter overview

The aim of this chapter is to apply the suggestions presented in Chapter 4 to develop a methodological framework. The second objective of this thesis is to provide guidance for developing a conceptual model for the new role for conceptual models proposed in this thesis; while there is existing guidance on how to develop conceptual models for informing the structure of a decision analytic model in economic evaluation, there is no guidance on how to develop a conceptual model for other purposes. This chapter describes how the methodological framework was developed and presents the final version; this is the second output of the thesis.

The layout of the chapter is as follows; three Sections, 5.2, 5.3, and 5.4 describe the methods used to develop the framework following the three phases of suggestions from Chapter 4. To recap, in Phase 1 existing methodological frameworks for developing conceptual models are identified, in Phase 2 data is extracted from these methodological frameworks and amalgamated into a draft methodological framework, and in Phase 3 the draft methodological framework is evaluated and refined. The result, in the form of the final methodological framework, is presented in Section 5.5, a discussion of the development of the methodological framework is included in Section 5.6, and Section 5.7 summarises the chapter.

5.2 Methods: Phase 1 - Identifying evidence to inform the methodological framework

In this first phase evidence was identified to inform the methodological framework using a literature review.

5.2.1 Literature search - methods

The aim of the literature search was to identify existing methodological frameworks for developing conceptual models, searching outwith the field of economic evaluation, to inform the methodological framework.

From an initial search, it was apparent that conceptual modelling methodological frameworks were available in different disciplines and from different sources; not restricted to peer reviewed academic articles. Therefore, I needed literature search methodology covering a broad range of different disciplines and sources, allowing me to identify a wide range of frameworks; scoping review methodology filled this requirement. Scoping reviews are suitable for exploring an emerging area to ascertain the extent, nature and breadth of available evidence (147, 150). The purpose of a scoping review is to map the evidence in an area, with a range of designs and methods, giving a descriptive overview of the evidence without critical appraisal or synthesis (169). Using a scoping review enabled me to identify frameworks from a wide range of disciplines, summarise the frameworks available in Phase 1 and to present the themes from these frameworks in Phase 2.

5.2.2 Scoping review methodology

As I have provided detailed information on scoping review methods in Chapter 4, in this chapter I present my methods without reference to published scoping review methods and guidance.

1 - Identifying the research question

My research question was 'What reported steps are used to develop conceptual models in methodological frameworks, and is there enough commonality in these steps to combine them into a single methodological framework for developing conceptual models in economic evaluation?'

2 - Identifying relevant studies

Following scoping review guidance I conducted an initial search to gain an overview of the conceptual modelling methodological frameworks, then the results of the initial search were used to inform and guide the main literature search (149).

The initial search comprised a general search using Web of Science and the Google search engine. Search terms were kept simple; 'conceptual model*' and

'framework' OR 'method' in the Web of Science search and 'steps for developing a conceptual model' in the Google search.

Published peer-reviewed and grey literature were found; grey literature included government guidelines and charity websites.

Disciplines identified in the search included; operations research, database management, software design, ontology, ecology, engineering, space missions, healthcare and military/defence. The methodological frameworks from database management, software design, ontology and space missions primarily focussed on the software used in the model at the conceptual model phase, rather than the representation of a decision problem, so at this stage seemed less likely to be suitable to inform the proposed framework.

Alongside methodological frameworks, records with no stepwise methodological framework, only methodology and were also identified, these mainly described the iterative process of developing a conceptual model and keeping the conceptual model simple. Whilst these points are not considered discrete steps, as they relate to more than one step in a conceptual model process, they are important in the development of a conceptual model.

The results of this initial search informed the main literature search in the following ways:

- Grey literature was searched in addition to published peer-reviewed literature.

- Only methodological frameworks with discrete steps/stages were included; whilst methodology only is useful, the aim of this thesis was to present a stepwise methodological framework, therefore papers containing only methodology were excluded from the search.

- All disciplines were included in the search ensuring that no relevant frameworks were omitted, however, special attention was given to any methodological frameworks identified from the fields of database

management, software design, ontology and space missions as these were unlikely to be suitable for inclusion.

The search strategy was developed with input from the College librarian, both for search terms and databases. The search terms were simple and not restrictive, and only titles were included rather than abstracts to ensure that the search terms were the main focus of the article or paper (search terms are included in Appendix 5: Scoping review search terms (*Chapter 5)*).

The Web of Science database was purposefully searched to identify conceptual modelling methodological frameworks published in peer-reviewed academic articles. The search took place in November 2016. An alert was set up for this search to identify relevant articles published after the search was carried out, I included the results of this alert in my search results up until October 2018. A second search was undertaken in the Scopus database, Scopus is a larger database than Web of Science, and although there is overlap between the two databases it is considered good practice to search both databases.

To ensure the search was as broad as possible in both database searches, there was no restriction on dates. References and citations were examined to identify any further relevant methodological frameworks. Where papers were not relevant (for example; not methodological frameworks but conceptual models of a specific subject), the methodology was scrutinised for any description of methodological frameworks used in the development of the specific conceptual model.

Alongside these database searches, two grey literature searches were carried out to identify non-peer reviewed methodological frameworks. As mentioned previously, during the initial search it became apparent that there was a rich source of non-peer reviewed methodological frameworks that could add to this research and I did not want to exclude this potential source. There is no gold standard for searching grey literature, however, research has found that using internet search engines is likely to result in identifying up-to-date grey literature (155).

The two grey literature searches comprised an internet search and a search of the grey literature repository 'OpenGrey'. For the internet search, methods developed and presented by Godin for reviews of grey literature (156) were adopted. The internet search was carried out using the 'Google' web browser, three search terms were used ('conceptual model steps', 'conceptual model development' and 'conceptual model guide'). OpenGrey is a grey literature repository with 700,000 pieces of grey literature produced in Europe. The types of grey literature include technical or research reports, doctoral dissertations, conference papers, official publications, and other types of grey literature. The disciplines covered by OpenGrey include: science; technology; biomedical science; economics; social science, and humanities. The OpenGrey search included two broad search terms (''"conceptual model*" NEAR framework' and ''"conceptual model*" NEAR methodology').

3 - Study selection

The inclusion criteria were: 1) methodological frameworks with discrete steps only; 2) no restrictions on dates were included so that the search was as broad as possible; 3) only English records were included, and 4) I included all methodological frameworks whether or not they were intended for developing conceptual models that were planned to be used for further development to mathematical models.

Duplicates were excluded, titles and abstracts were screened, and either discarded or kept and read in full.

4 - Charting the data

Excel was used to extract the following data from the included methodological frameworks: author/s; title; year; source and type of record (ie published article/government guidelines); discipline; rationale for developing the framework, and which search had identified the record. NVivo software was used to extract information on themes (NVivo qualitative data analysis software; QSR International Pty Ltd. Version 11, 2015). Thematic analysis is recommended in scoping reviews (150) and this technique was used to extract data as there was no prior hypothesis to test; the themes needed to be extracted from the

frameworks in an inductive way to best summarise the evidence. NVivo is compatible with thematic analysis (170); it is a data management tool useful for extracting themed data from a wide range of records. NVivo is also useful for an iterative extraction process where themes are updated and developed as the literature is understood in more detail.

5 - Collating, summarising and reporting the results

The characteristics of the methodological frameworks were summarised narratively and in a table, and the rationales for developing the methodological framework were also presented, grouping the methodological frameworks into disciplines to highlight any commonalities within disciplines.

The unit of analysis chosen for the thematic analysis was steps in the included methodological frameworks, these were coded and an iterative process was followed, as illustrated by the first four boxes in Jamieson's diagram in Figure 19, to allocate the units of analysis to the emerging themes (171). My rationale in choosing steps as a unit of analysis was that these are needed to answer the research question. To avoid confusion with the stages in the proposed methodological framework these steps are called 'themes' in this chapter. I went through three iterations of labelling emerging themes and allocating the units of analyses to these in a logical manner. I present my results as both numerical and qualitative data.

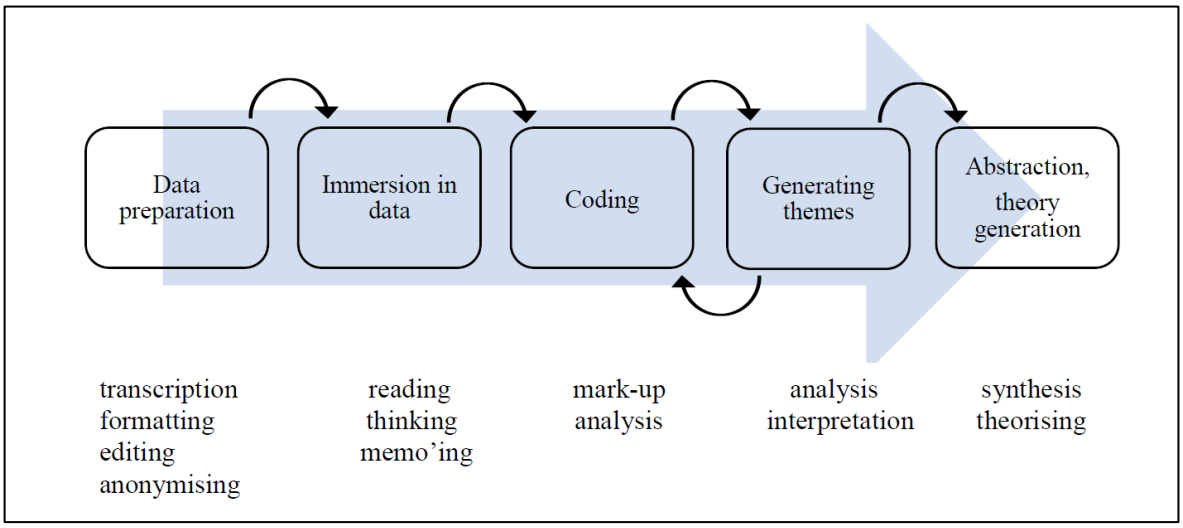


Figure 19: The process of qualitative data analysis from Jamieson (2016) (171)

Chapter 5

5.2.3 Literature search - results

Search results and selection of methodological frameworks

Results of the search and selection are presented in Figure 20.

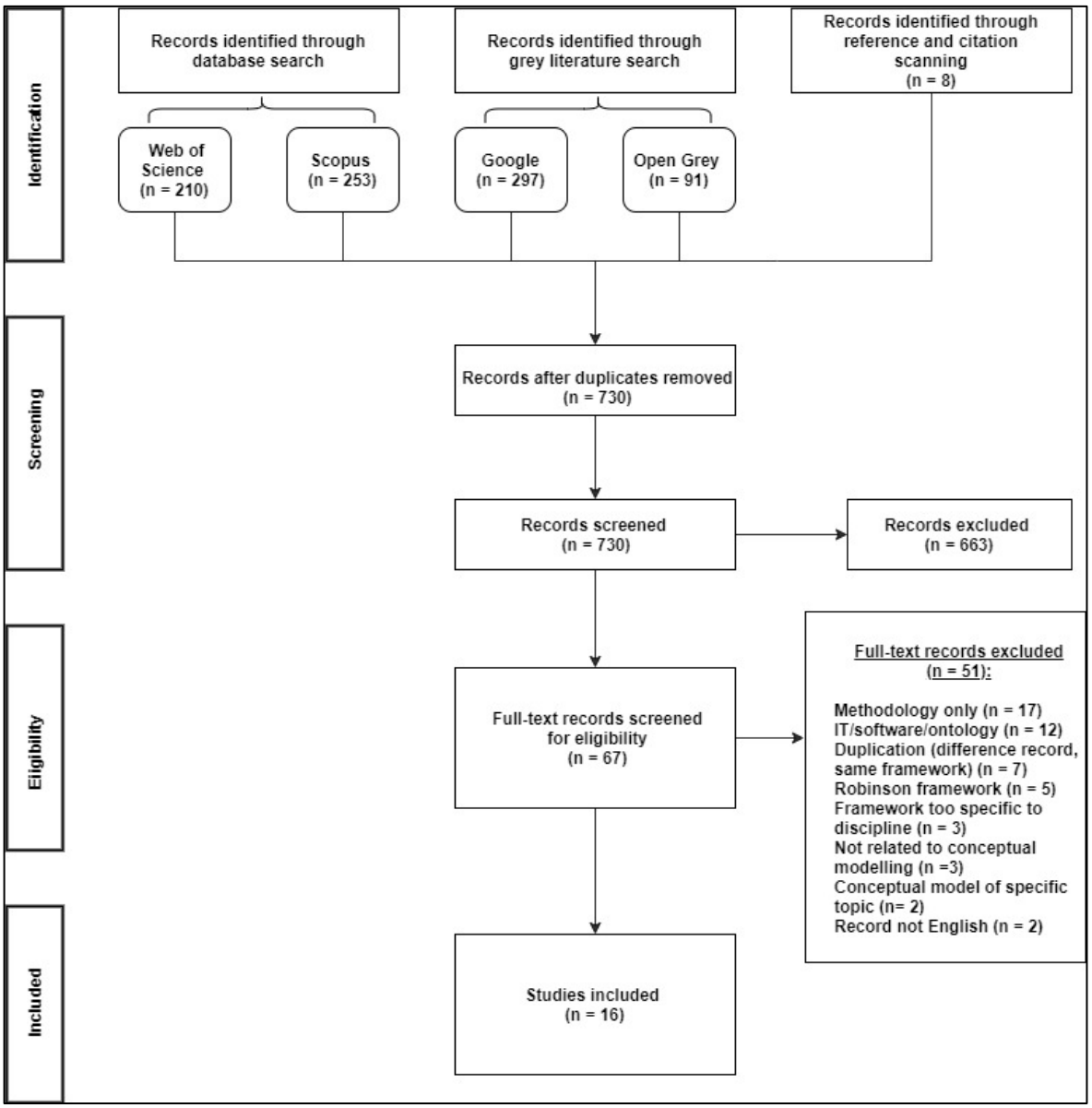


Figure 20: PRISMA diagram of included methodological frameworks

Records were identified from the Web of Science, Scopus, Google grey literature and Open Grey literature searches, and from searching the references and citations of included studies.

177 potential records were identified in the Web of Science search; articles, proceedings papers, meeting abstracts, book chapters, editorial materials, reviews and letters. A further 33 items were identified from the Web of Science alert. 253 potential records were identified in the Scopus search; articles, books, conference proceedings and reviews. The Scopus search included 120 items that had already been identified by Web of Science from a potential 253 records, confirming that there is overlap between the two databases.

The Google grey literature searches resulted in a total of approximately 60,650,000 hits, Godin's methodology was applied to reduce the number of potential records;

- The 'Conceptual model steps' search resulted in approximately 48,400,000 hits, I searched the first 10 pages, retrieving 99 potentially relevant records.

- The 'Conceptual model development' search resulted in approximately 7,780,000 hits, retrieving 99 potentially relevant records on the first 10 pages.

- The 'Conceptual model guide' search resulted in approximately 4,470,000 hits, again retrieving 99 potentially relevant records on the first 10 pages.

The OpenGrey literature search resulted in 81 potential records from the '"conceptual model*" NEAR framework' search and 10 from the '"conceptual model*" NEAR methodology' search.

In summary, 210 potential records were identified in the Web of Science search, 253 from Scopus, 297 records were identified in the Google grey literature search and 91 from the OpenGrey literature search. In total 851 potential records were identified from the literature search, and a further 8 from reviewing references and citations.

Of the 859 potential records 129 were duplicates and discarded, after the titles and abstracts of the remaining records were scanned a further 663 records were

discarded. 67 records were screened in full for eligibility, 51 were excluded at this stage.

In total 16 records (10 from the Web of Science search and 6 from the grey literature Google search) were identified as eligible to be included in the scoping review, these studies are reported in Appendix 6: Results of scoping review (*Chapter 5)*.

Methodological framework characteristics

The included methodological frameworks came from the fields of ecology, engineering, operations research, healthcare, physics, manufacturing and construction. They dated from 2000 to 2017 and the types of records identified comprised: articles; conference proceedings; government guides; a charity guide, and a presentation, illustrating the broad origins of the methodological frameworks (Figure 21).

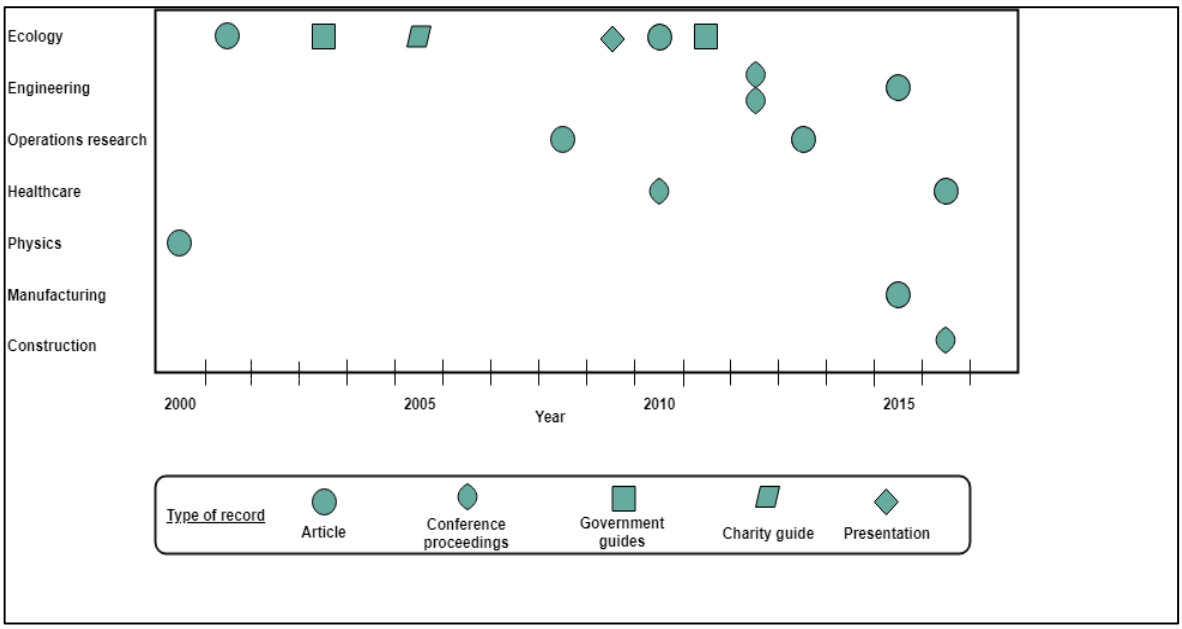


Figure 21: Summary of framework characteristics

In ecology the rationales for producing the methodological frameworks were either meeting statutory requirements, or the need for monitoring and risk assessments necessary in the discipline. In engineering and operations research the rationales were mainly because of a lack of guidance, either in specific

methodologies such as discrete-event simulation, using soft systems methodology (SSM), or linking to engineering process. In healthcare the rationales presented were mainly to increase the success of studies to inform change in healthcare. The rationale behind the physics methodological framework was to improve the quality and relevance of the simulation study, similarly in manufacturing the aim of the methodological framework was to use SSM to improve the simulation study. The construction methodological framework rationale was that having a methodological framework to develop conceptual models reduces the burden of producing a conceptual model, which is believed to be a complex task, and having the conceptual model would increase the acceptance of discrete event simulation in the construction industry.

In summary, reported rationales for developing the methodological frameworks can be split into three groups: first, to meet statutory requirements for risk assessments (15, 172) or to meet internal standards (14); second to produce guidance in areas where there is lack of methodology (18, 20, 21, 43, 173, 174), and third, to improve the conceptual modelling phase or quality of the conceptual model and therefore project outcomes or acceptance of modelling methods (12, 16, 175-177). There were two papers with no rationale (19, 178)

5.3 Methods: Phase 2 – Developing the methodological framework

In the second phase the methodological frameworks identified in Phase 1 were analysed and built upon; key data was extracted from the identified evidence and amalgamated into common themes to inform and produce the draft methodological framework.

5.3.1 Extraction of themes

18 themes were identified from the 16 methodological frameworks. Initially I extracted steps and allocated these to themes. Allocating the steps to themes was an iterative process; as I read more papers and used NVivo to identify steps in the frameworks, it became clear that most methodological frameworks included similar themes, although sometimes different terminology in the steps

was used for the same theme. A summary of the different terminology allocated to each theme is included in Appendix 7: Terminology allocated to themes (*Chapter 5*).

Themes identified in the included methodological frameworks are summarised with a count in Table 12, with the most frequently cited theme first. The descriptions of the steps and rationale for allocating the steps to each of the themes are summarised narratively. These descriptions feed into the guidance included in the draft methodological framework.

Table 12: Counts of themes in frameworks

Summary of themes in frameworks (n = 16)		Count (%)
Objectives	Yes No	13 (81%) 3 (19%)
Model scope	Yes No	10 (62%) 6 (38%)
Model content	Yes No	8 (50%) 8 (50%)
System behaviour	Yes No	8 (50%) 8 (50%)
Understanding the problem	Yes No	8 (50%) 8 (50%)
Model inputs	Yes No	8 (50%) 8 (50%)
Model outputs	Yes No	7 (44%) 9 (56%)
Documentation	Yes No	6 (38%) 10 (62%)
Assumptions and simplification	Yes No	6 (38%) 10 (62%)
Model detail	Yes No	6 (38%) 10 (62%)
Model diagram	Yes No	5 (31%) 11 (69%)
Entities	Yes No	5 (31%) 11 (69%)
Validation	Yes No	4 (25%) 12 (75%)
Stressors	Yes No	3 (19%) 13 (81%)
Review, revise, refine	Yes No	3 (19%) 13 (81%)
Model structure	Yes No	2 (12%) 14 (88%)
Team identification	Yes No	2 (12%) 14 (88%)
Use previous conceptual model	Yes No	1 (6%) 15 (94%)

Whilst in most cases it was obvious which themes a step should be allocated to, where it was not clear I used the intended outcome of the step to guide the allocation. For example, ‘Understanding the problem’ overlaps with ‘System behaviour’, these steps have been allocated to separate themes based on the intended outcome or result of the step. In ‘Understanding the problem’ the intention is for the modeller to get a good understanding of the problem

situation to lay the foundations of the conceptual model development and confirm that the modeller and stakeholders all have the same understanding from their different viewpoints. ‘System behaviour’ is how the entities within the system interact and the intention of this theme is to contribute to the development of the body of the conceptual model.

Some methodological frameworks had more than one step in them that only related to one theme, in these cases the separate steps were amalgamated. Interpretation of the less clear, more discipline specific steps, was necessary to assign them to a relevant theme. For example, the three ecology methodological frameworks; Brassington et al., Gross and Gentile et al. all have more than one step that relates to one theme, this is because ecology systems are complex with many components and stressors, so to develop an ecological conceptual model many elements need to be understood (15, 20, 172). The Brassington framework has the following steps relating to the ‘model content’ theme; ‘defining the topography and surface water drainage’, ‘defining the geology’, ‘defining the aquifer framework and boundaries’, ‘defining groundwater flow directions’, ‘defining the aquifer relationships’ and ‘water balance’ (15). Gentile has the following steps in the ‘stressors’ theme; ‘Inventory resource use and other human activities’, ‘describe sources of natural and anthropogenic stressors’, ‘identify the primary and secondary stressors of concern’ and ‘describe stressor mechanisms and routes of exposure’ (172).

Conversely, some methodological frameworks included more than one theme in a step, in these cases the steps were split into separate themes. For example, in Brassington et al. the final step of ‘Description of model’ Brassington describes two themes; documenting the conceptual model and producing a diagram (15). Robinson includes the themes ‘model scope’, ‘model content’, ‘assumptions and simplification’ and ‘model detail’ in his step ‘defining the model content’ (43).

Chapter 5
Objectives

Objectives are linked to the project problem definition (44), examples include; identifying problems such as low productivity, long queues and bottlenecks, or checking the workings of new systems or modifications to old systems given (21). The objective should never be the development of the conceptual model (21, 43).

Rationales for including objectives in the methodological frameworks comprise; objectives are important in defining the 'stressors, endpoints and management options' relevant to the conceptual model and setting these objectives is essential because the conceptual model needs to relate to the research question (172). Also, objectives are key; they drive the modelling process development and use of the conceptual model (43). However, the relative importance of setting goals will vary depending on the audience (20). Defining objectives also helps to manage stakeholders' expectations of the model (12).

Several methodological frameworks discussed two types of objectives; modelling and general objectives (12, 18, 43, 44). Modelling objectives are specific to the conceptual model and project, for example limiting throughput times (43) and can be broken down into three components:

1. Achievement: what the clients aim to achieve, such as increase throughput, reduce cost, improve understanding of the system, improve efficiency of resources.

2. Performance: quantifying performance, such as increasing throughput by 10% or reducing cost by £10,000.

3. Constraints: constraints on the modeller, such as budget, design options.

Modelling objectives may also change and develop, emphasising the iterative nature of conceptual modelling (43).

General objectives are requirements of the conceptual model linked to the nature of the project, examples include run-time, development time of the

conceptual model, visualisation requirements or re-usability of the conceptual model, with 'time-scale' described as especially important (43). These will likely be determined by the project. General objectives add to the modelling objectives and include; visualisation, interaction, flexibility, reuse and project time frame (44).

Model scope

The scope defines the 'extent' of the project, an example being a manufacturing system where only certain parts of a process are included (21). Deciding on the conceptual model scope can be described as 'bounding the decision problem'; choosing what is to be included in and excluded from the conceptual model (12), and what is outside the scope of the project (172).

Model content

There was only one description of conceptual model content; defining components of the conceptual model (15).

System behaviour

System behaviour relates to the components of the system, how these are linked (20), and the rules that control the conditional behaviour of the system (12). An example is how endpoints are linked to physical characteristics and stressors; developing hypotheses for explaining how human influences in the South Florida ecosystems are responsible for the deterioration of the ecosystems (172).

Understanding the problem

This theme relates to understanding the subject matter (44), however no descriptions were reported in the methodological frameworks.

Several methodological frameworks included a rationale; to develop a 'sound' conceptual model the system should be fully understood both as a whole and in individual sub-sections (13). Furthermore, the more complete and clear a problem area is the easier it is to understand (176) and this stage is important

because it lays the foundation of the study (175). Robinson states that the need for a simulation model should always be 'driven' by attempting to improve the problem situation. Therefore, the starting point for developing a conceptual model should be understanding the problem situation, however, problems can occur when stakeholders hold different views of the problem situation.

Model inputs

Inputs are the information to be entered into the conceptual model (21), that can change in different scenarios (18). For example, in the Furian et al. port problem case study, inputs denote the number of berths. Furian et al. explain that in healthcare input factors may not represent single measures, but policies, the example given is dispatching orderlies in a hospital where different policies include; dispatch the closest orderly, dispatch the orderly who could get to the task first, or use an algorithm to dispatch orderlies. Other examples of inputs are machines, work posts, reception areas, as well as individuals who work in these locations. The project objectives will affect how these inputs are assessed (13).

Conceptual model inputs (or experimental factors) can be either quantitative or qualitative (12). For example, quantitative inputs include the number of staff or speed of service, whereas qualitative inputs include changes to rules or the structure of the model (43). Overall, clients will have control over the inputs (for example staff numbers), however, it can be useful to vary the inputs with little control over them, which can help with understanding the system or help plan for future happenings.

Documentation

Documentation is a written description of the conceptual model (15), there are few descriptions or rationales of this theme.

Model outputs

Outputs are functions of 'targeted performance measurements', examples of which are; throughput, average time waiting and picker utilisations (21).

Outputs (or responses) can be either be numerical (for example means) or streamed data (for example time series) (18). An example of outputs from the Furian et al. port case study is throughput times of tanker and container ships.

An output is usually the output from the project system's transformation process, for example a patient who needs treatment becomes a treated patient. The purpose of the output is to assess whether the objectives of the model are met, and if they are not met, why not (12, 13, 43).

There are two categories of outputs. The first type are linked to conceptual model objectives; performance measures. The second type are outputs helping to locate bottlenecks in the system, for example, flow time performance might be explained by waiting times (44).

Diagram

A diagram visually links the process and components (172). The rationales for including a diagram are; the process of visualising the conceptual model enhances the thought process of developing the conceptual model (16), it can be useful for validation and assessing relevance (44), and it can also be used as a communication tool (13).

Assumptions and simplifications

Assumptions and simplifications 'define' the conceptual model scope and detail (12). They 'are a facet of limited knowledge or presumptions' (21), a result of uncertainty about the real world situation (43), and the more assumptions made in developing the conceptual model the less detail is included in the conceptual model (21). 'Simplifications are a facet of the desire to create simple models' (21), and are made to make the conceptual model quicker to develop and use (43).

Model detail

The level of detail is the depth that the conceptual model goes into, this is subtle, for example 'manufacturing process equipment' might be included in the

model, but this may only be the total processing time of that equipment(21). Deciding on the conceptual model detail includes basic choices about whether a system is represented at a higher, more broader level, or whether the modeller drills down into sub-sections to include the level of detail required(176).

Validation

Validation is 'determining that the theories and assumptions underlying the conceptual model are correct and that the model representation of the problem entity is 'reasonable' for the intended purpose of the model' (13).

The rationale for including this theme is that whilst the conceptual model may appear completely rational from the modeller's point of view, the modeller should view the conceptual model as a communication tool to engage with stakeholders in the project and confirm and agree on understanding. These stakeholders will have different perspectives of the research problem to the modeller and should test the conceptual model to ensure it is making sense. Gray et al. explain that this step is 'informal but important'. Validation ensures there are no errors in the conceptual model (179).

Entities

Entities are the core elements of a discrete event simulation; there are two types of entity; active and passive (12, 18). Active entities comprise resources and consumers, and other entities that can change their role. For example, in the tugboat case study these would be tugboats and tankers. Passive entities are not related to the flow of a system, they are fixed, for example a harbour waiting area or berthing area (18).

Entities are components that can be identified as a 'pressure, state or response' to aid model construction (178).

Pace's rationale for including this theme is that entities (and processes) are needed to achieve the project objectives and link to the detail and scope of the conceptual model (176).

Stressors

Stressors are human activities and natural stressors that affect the ecosystem (172), for example 'excessive hunting and illegal timber extraction' (14).

Review, revise and refine

The review, revise and refine theme is testing the conceptual model for its usefulness and suitability (16). The rationale behind including this theme is that all conceptual models are an 'incomplete abstraction of reality' and that most will require revision to 'accommodate new observations, information or changing goals' (20). Also, when the original conceptual model is first developed it will be based on the best-available knowledge, and as this knowledge is updated the conceptual model should be reviewed (178).

Model structure

The conceptual model structure is characterised by the entities and their aggregation (18). There are no examples or rationales for this stage.

Team identification

There were no descriptions of 'team identification'. The rationale given for including this stage is that choosing a project team results in a successful beginning to the project and a 'seamless execution' of the project (175).

Use previous conceptual model

There was no description of this theme. The rationale given was that previous conceptual models can be applied or adapted to the research problem in hand (16).

5.3.2 Ordering, categorising and amalgamating themes; developing the framework iteratively

In this section the 18 extracted themes were grouped and amalgamated as the draft methodological framework was iteratively developed. First the number of

themes in each included methodological framework and the sequence of those themes were summarised, then the sequence of the themes was assessed for common patterns, this enabled the themes to be categorised into broad phases. Similar themes within each phase were amalgamated into stages, and the descriptions and approaches used in the identified methodological frameworks were summarised into 'how to' advice in the draft methodological framework. Care was needed when the existing guidance related to mathematical model development, which is not an element of my methodological framework.

Ordering of themes

The number of themes extracted from each of the included methodological frameworks varied from three to eleven (Figure 22). Seven was the most frequent number of themes extracted (n=4), and the least frequent was ten (n=0).

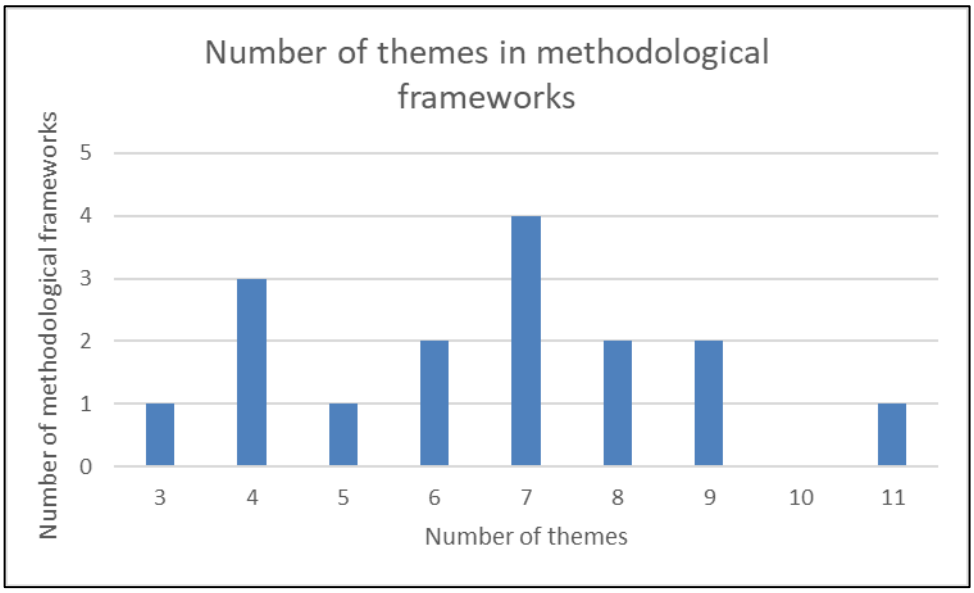


Figure 22: Histogram of number of themes in included methodological frameworks

Figure 23 sets out the sequence of the themes reported in each of the included methodological frameworks in the order that they were presented and includes the total number of themes identified. Themes highlighted in yellow represent where two or more themes were identified in one step within the framework, therefore it was not always possible to place these in an exact order.

Methodological framework	Total number of themes	Sequence of themes										
Abdelmegid	11	Identify team	Understanding problem	Objectives	Inputs	Outputs	Model structure	Assumptions & simplifications	Entities	Detail	Scope	System behaviour
Brassington	4	Objectives	Content	Documentation	Diagram							
Chwif	7	Objectives	Scope	Detail	Inputs	Outputs	Assumptions & simplifications	Documentation				
FEFLOW	4	Objectives	Inputs	Content	Validation							
Furian	9	Understanding problem	Objectives	Outputs	Inputs	Model structure	Entities	Scope	Detail	System behaviour		
Gentile	6	Objectives	Scope	Stressors	Entities	System behaviour	Diagram					
Gray	7	Understanding problem	Previous CM used	Diagram	Documentation	Validation	Content	Review, revise, refine				
Gross	5	Objectives	Scope	System behaviour	Stressors	Review, revise, refine						
Montevechi	7	Understanding problem	Objectives	Content	Inputs	Outputs	Diagram	Validation				
NSW	7	Objectives	Scope	Entities	System behaviour	Documentation	Assumptions & simplifications	Review, revise, refine				
Pace	6	Understanding problem	Detail	System behaviour	Entities	Content	Scope					
Pereira	4	Objectives	Content	Validation	Documentation							
Robinson	8	Understanding problem	Objectives	Outputs	Inputs	Assumptions & simplifications	Detail	Scope	Content			
Tako	8	Understanding problem	Identify team	System behaviour	Objectives	Assumptions & simplifications	Content	Inputs	Outputs			
van der Zee	9	Understanding problem	Objectives	Outputs	Inputs	Scope	Detail	Assumptions & simplifications	Documentation	Diagram		
WWF	3	Scope	System behaviour	Stressors								

Figure 23: Sequence of themes in included frameworks

Once the themes were ordered a general pattern emerged, and the themes were categorised as three phases (Figure 24). The first phase was understanding the problem and setting the conceptual model objectives. The second phase was the content of the model; what was included and excluded and at what level of detail, and how the components of the conceptual model are related to each other. The third phase was validating the completed conceptual model and documenting it. Full details and explanations of the ordering and sequence of themes in each phase is available in Appendix 8: Ordering and sequence of themes in each phase (*Chapter 5*).

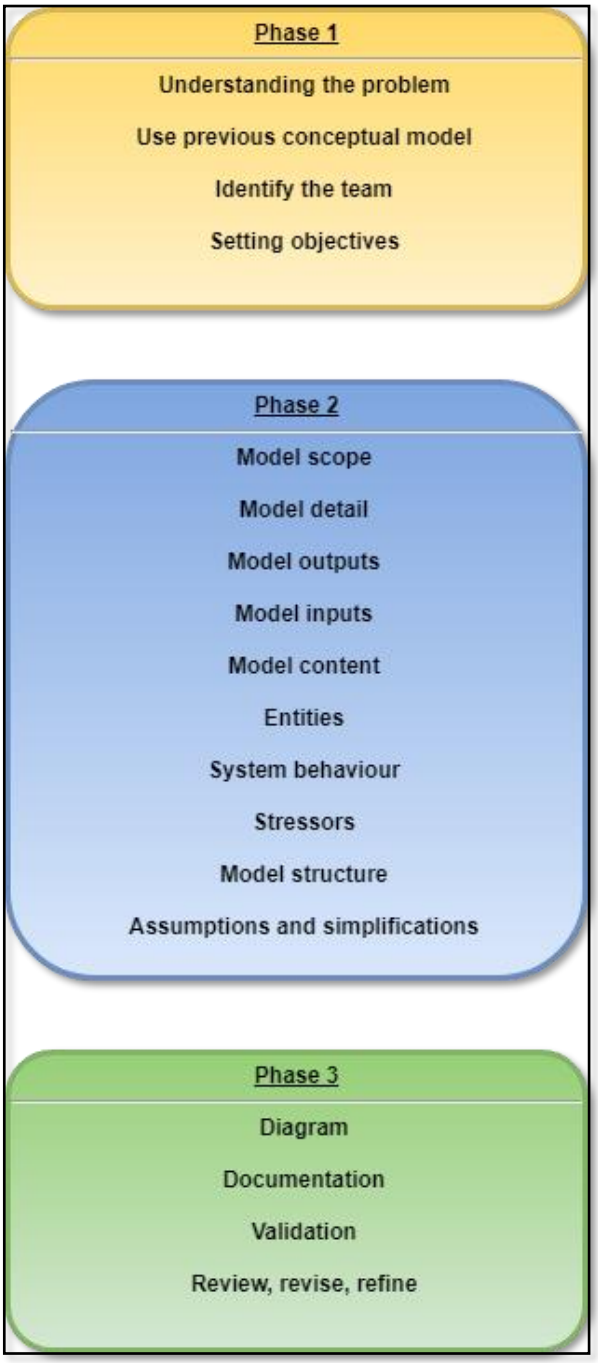


Figure 24: Initial sequence of stages and phases

The next step in developing the methodological framework was to combine similar themes within each phase into discrete stages. To do this I looked at the sequence of themes within each phase and the descriptions of the themes to inform the amalgamation of similar themes into stages. The extracted themes, stages, and phases are presented in Figure 25.

In Phase I ‘Context’, the ‘Understanding the problem’, ‘Use previous conceptual models’ and ‘Identify the team’ themes are combined into an initial stage of

'Understanding the problem'. The 'Setting objectives' theme is also included in this phase as a separate stage.

In Phase II 'Development', I grouped all themes except the scope and detail themes into an umbrella 'Determine model content' stage, the determine the scope and detail themes were left as separate stages. Assumptions and simplifications are an important aspect of developing the conceptual model (and are also used in understanding the problem), but I do not consider them to be a distinct stage, rather an concept that is crucial to the successful development of a conceptual model.

In Phase III 'Finalising', I combined the diagram and document themes into one 'Diagram and documentation' stage, and combined the 'Validation' theme with the 'Review, revise and refine' theme; validation is similar to reviewing and fits well into the 'Review, revise and refine' stage.

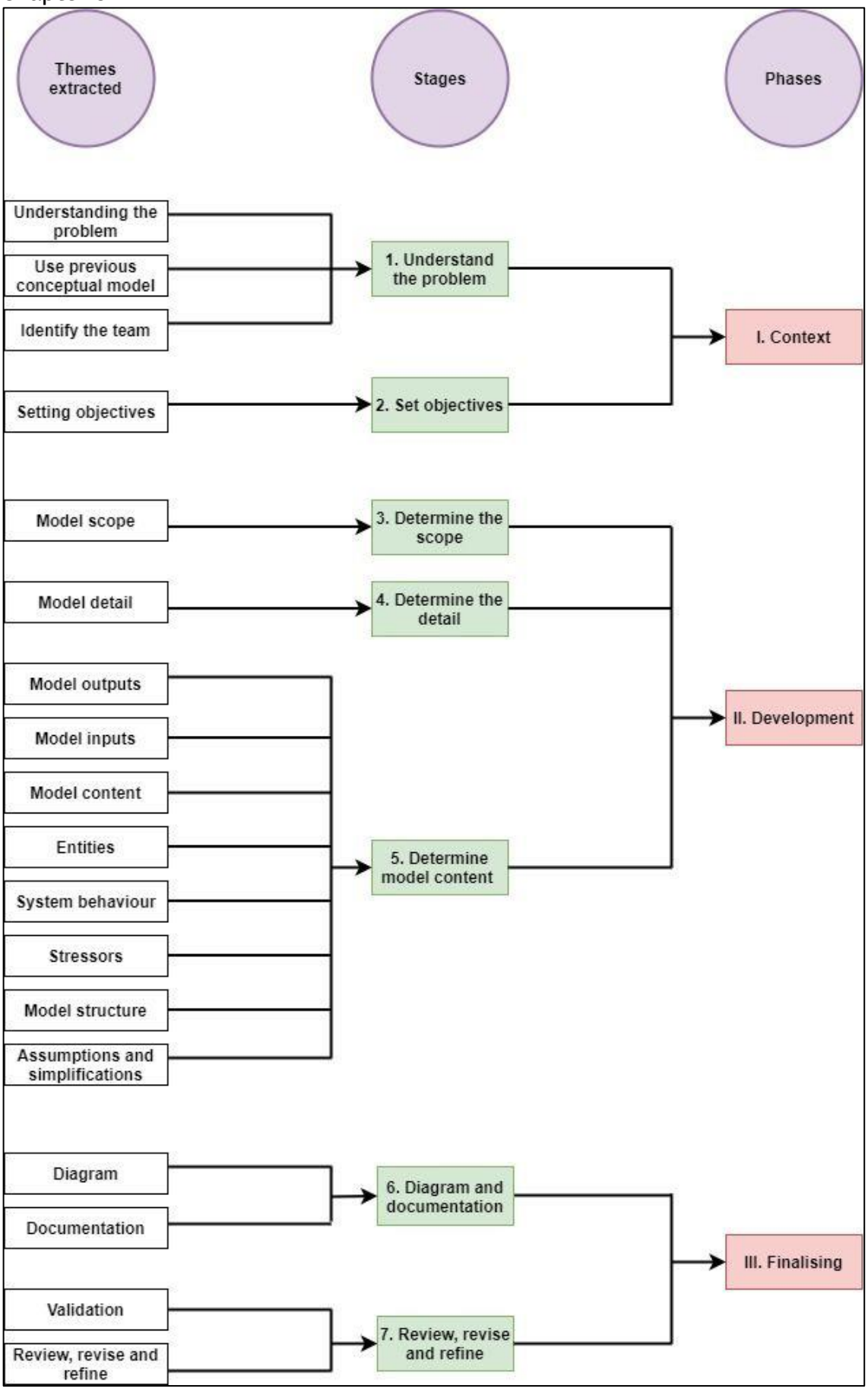


Figure 25: Themes, stages, and phases from scoping review results

The stages were populated with descriptions from the themes extracted from the included methodological frameworks to produce the draft methodological framework, I also supplemented the descriptions with economic evaluation specific descriptions to tailor it to developing conceptual models in economic evaluations. The draft methodological framework is presented in Appendix 10: Draft methodological framework.

5.3.3 Framework diagram

In this section the steps for developing the diagram for my methodological framework are described.

After reading the included methodological frameworks it was clear that including a diagram would be beneficial for my methodological framework; it would serve as a useful tool for the modeller, enabling them to see at a glance the stages involved in developing a conceptual model without being bogged down in the written detail. Also, a diagram is a useful tool for showing the flow of the stages in the methodological framework and the links between those stages without the need for wordy explanations in addition to the written description of the methodological framework.

To design a clear and understandable diagram I first looked at the included methodological frameworks to see if a diagram was included; eight methodological frameworks included a diagram (12, 15, 18, 19, 43, 175, 177, 180), these diagrams are presented in Appendix 9: Methodological framework diagrams from identified studies and early drafts of diagram to include in final methodological framework (*Chapter 5*), along with early versions of my diagram.

Whilst the Abdelmegid et al. diagram (Figure A9.70) showed the forward linear direction through the stages it read more like a list and had no indication of the iterative nature of a methodological framework (12). The diagrams in the Tako et al. (175) and Furian et al. (18) methodological frameworks (Figure A9.71 & Figure A9.72Figure A9.71) were simple and easy to understand and follow, showing the flow through the process but again there was no indication of the important iterative nature of the methodological framework. The layout of the

Robinson (43) and Pace et al. (176) diagrams (Figure A9.73 & Figure A9.74) was simple and clear, showing the flow of the process but also the iterative nature of the methodological framework, however my draft proposed methodological framework had more steps than these and I was concerned that using a similar diagram would look cluttered and busy and would be too difficult to understand at a glance. Brassington et al. (15) and Hesch (19) included diagrams (Figure A9.75 & Figure A9.76) that were complicated to follow, these were similar to flow charts and suitable for the complexity of the methodological frameworks, but not suitable for the draft proposed methodological framework. Finally, the circular nature of the Pereira et al. (177) diagram (Figure A9.77) indicated the flow of the process, but also illustrated the iterative nature of the methodological framework.

I wanted a diagram that was not simply a list, but which would be an easy to understand visual guide to the modeller, showing the order of the stages and phases, and also illustrating the iterative process of developing a conceptual model. I felt it was extremely important to produce a diagram that succinctly and correctly summarised the proposed methodological framework as it would be the part of the methodological framework that would be the most memorable for many users.

Based on the Robinson and Pace et al. diagrams I developed two drafts of linear diagrams, these showed the linear progression through the methodological framework but not the iterative nature (Figure A9.78 & Figure A9.79).

I also looked at the wider literature to explore different types of diagram, firstly investigating ‘waterfall models’. Waterfall models were introduced in 1970 by Winston W Royce and are primarily used for software development but can used for any step-by-step process (181). They are also known as ‘linear-sequential life cycle models’ and ‘process models’. This led me to explore process flow diagrams, these are often used in chemical and process engineering to illustrate the layout of major pieces of equipment to carry out a specific process (182).

Having looked at examples of waterfall model diagrams I considered that a waterfall diagram with the general downward flow of stages or tasks, but with the ability to revisit previous stages iteratively, may illustrate my proposed

framework well. However, on consultation with colleagues, it became clear that the iterative nature of the framework was not well communicated using a waterfall diagram (Figure A9.80). I decided that a circular diagram would illustrate the framework at a high level of detail, similar to the Pereira et al. diagram, an early draft version of the circular diagram is included in the (Figure A9.81). The final diagram is presented in the final methodological framework in Section 5.5, Figure 30.

5.4 Methods: Phase 3 – Evaluate and refine

In Phase 3 the draft methodological framework developed in Phase 2 was evaluated by comparing it to existing health economic evaluation conceptual modelling guidance and assessing any gaps (22, 23, 25, 26); refinements were made to produce the final methodological framework. Suggested approaches for evaluating the methodological framework from Chapter 4 did not include comparing the draft methodological framework to current guidance, however this method was reported by Rivera et al (39).

5.4.1 Comparison to economic evaluation methodology

In this section I compare my draft methodological framework to published guidance on developing conceptual models in the field of economic evaluation. This process helped me assess my draft methodological framework within the field of economic evaluation. I identified what was similar and what was different between the guidelines and my draft methodological framework. I also assessed whether the differences warranted a change in my draft proposed framework, particularly to translate guidance extracted from methodological framework in fields other than economic evaluation into economic evaluation specific guidance in my methodological framework, but without including guidance related to decision analytic modelling. All of these guidelines were presented in Section 1.2.

I have split this comparison into ‘Steps’ and ‘Best practice and guidance’.

Chapter 5
Steps

Squires et al. (22) comprises four discrete steps and advice and it is also described as a methodological framework by the author; the other guidance are presented as guidelines (23, 25, 26). Tappenden's guidance describes three broad steps, and the Roberts et al. guide and Chilcott research contains guidance rather than sequential steps. These latter guidelines were harder to compare to my draft proposed framework as they do not comprise discrete steps.

The four steps in the Squires framework are illustrated in Figure 26, they are: 1) 'Aligning the framework with the decision making process'; 2) 'Identifying relevant stakeholder'; 3) 'Understanding the problem' and 4) 'Developing and justifying the model structure'.

The first step 'Aligning the framework with the decision making process' is not a step that was included in any of the methodological frameworks from the scoping review in Section 5.2. The second step of 'Identifying relevant stakeholders' was included in two methodological frameworks from the scoping review, although it was implicit in many of the other methodological frameworks. Whilst I do not include this step as a discrete stage in my draft methodological framework, I do include it in the first stage of my draft methodological framework: 'Understanding the problem'. The third step 'Understanding the problem' is the first step in my draft methodological framework and was included in 50% of the methodological frameworks identified in the scoping review. The Squires et al. final step of 'Developing and justifying the model structure' is analogous to my Phase 2 'Development of the conceptual model'. Within the Squires et al. final step are several stages which I discuss in the 'Best practice and guidance' section.

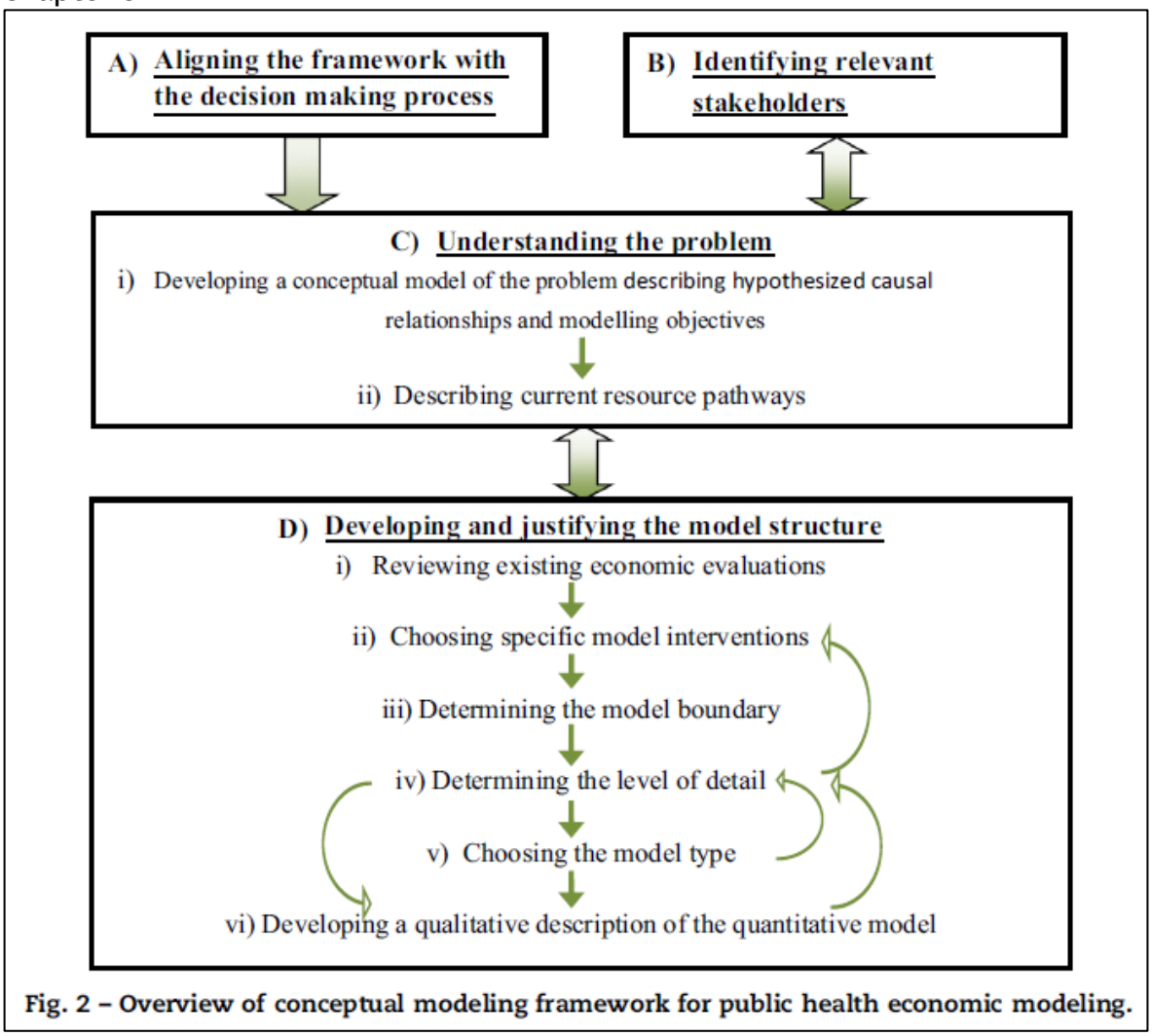


Figure 26: Squires framework diagram from Squires et al (2016) (22)

Although Tappenden does not produce guidance with discrete steps (illustrated in Figure 27), he suggests starting the conceptual modelling process by developing a problem-oriented conceptual model; this helps the modeller to understand the decision problem and the system relating to that problem. This is analogous to my first stage of 'Understanding the problem'. The second step in Tappenden's advice is the development of a design-oriented model. This type of conceptual model provides a platform to discuss and agree the structure of the decision analytic model and the evidence or data needed for it.

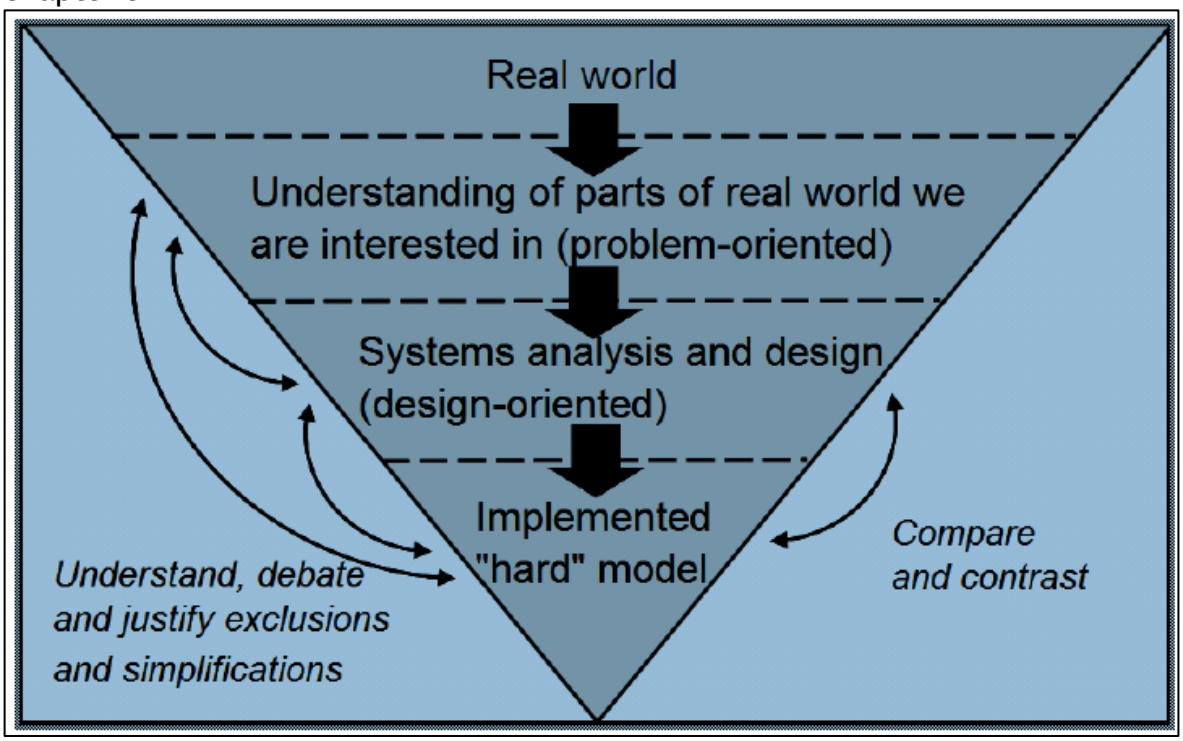


Figure 27: Hierarchy of models from Tappenden (2012) (23)

This is mirrored in the Roberts et al. guidance in Figure 28; the paper is split into two components; conceptualising the problem and conceptualising the model. The authors state that understanding of the nature of the problem and the project objectives should be clear before developing a model. The rest of the guidance is set out as ‘best practices’, these will be compared to my guidance in the next section.

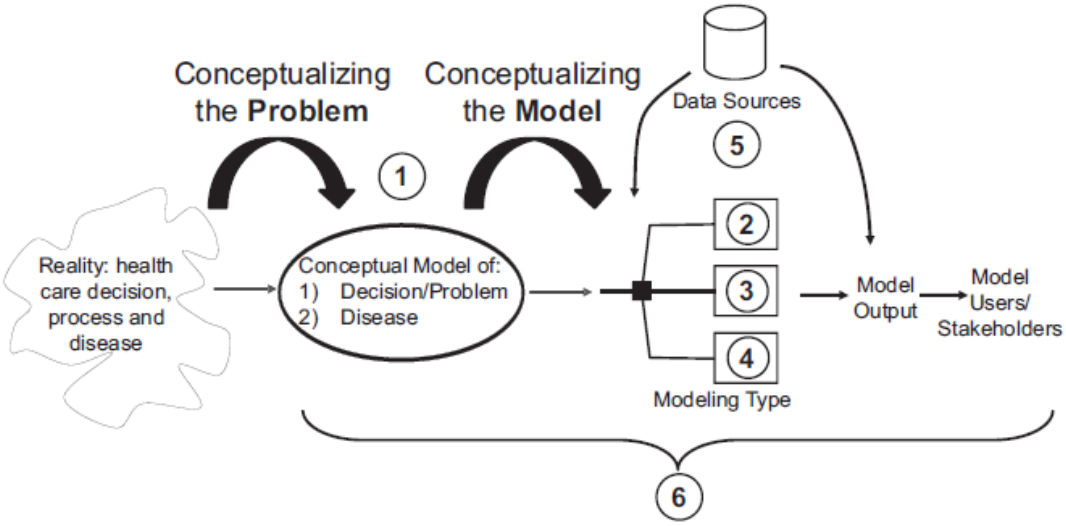


Fig. 1 – Development and construction of a model. The numbers in the figure represent the methods papers in this series: 1) the conceptualization paper, which describes the conceptualization of both the problem and the model; 2), 3) and 4) which describe the three main kinds of modeling methods addressed, including state transition model, discrete event and agent based models and dynamic transmission models; 5) parameter estimation used to calibrate the models, and 6) the transparency and validation of a model. See text for details.

Figure 28: Conceptualising a model from Roberts et al. (2012) (26)

Lastly, these stages are similar to the stages reported in the Chilcott paper (Figure 29): the first stage being 'Understanding the decision problem', and the second stage being 'Conceptual modelling' where the understanding from the first stage feeds into a decision analytic model.

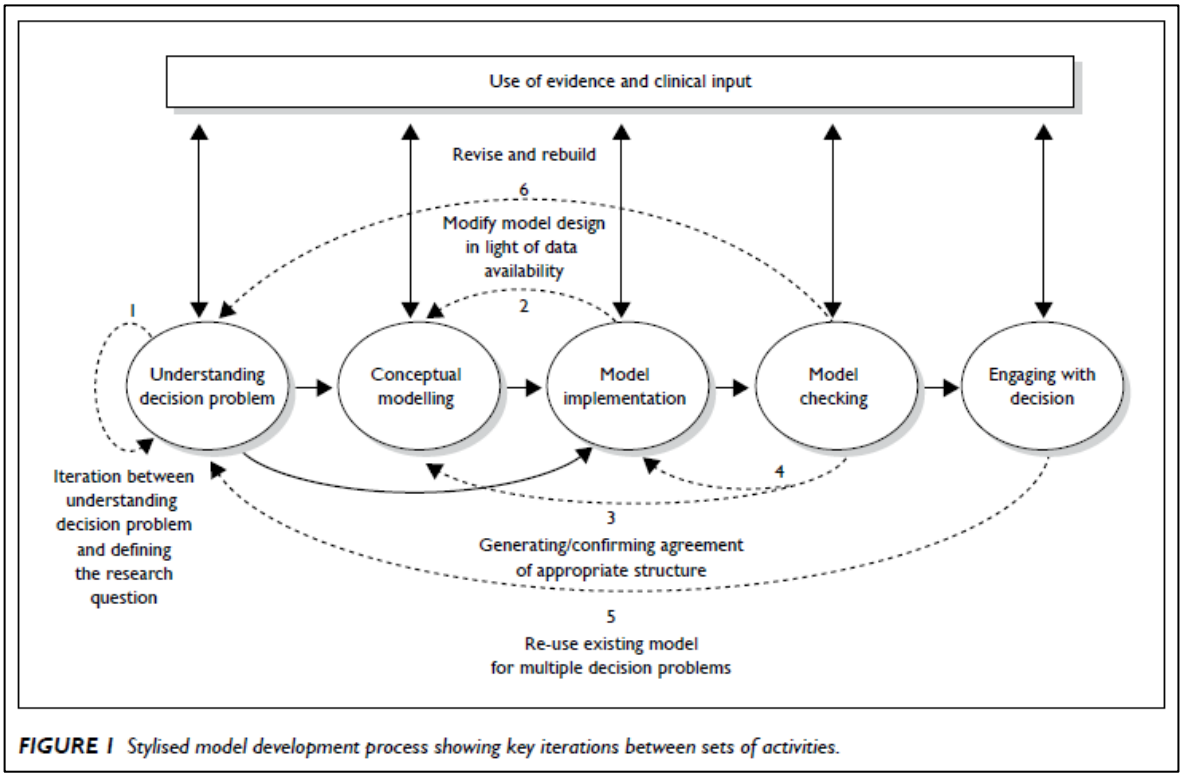


FIGURE 1 *Stylised model development process showing key iterations between sets of activities.*

Figure 29: Model development process from Chilcott et al. (2010) (25)

Best practise and guidance

Next, the draft methodological framework was compared to the best practise and guidance included in the existing economic evaluation conceptual modelling guidance. To do this I drilled down further into the guidance, breaking it down into component parts, these were not discrete steps followed in a linear or iterative way, but best practise, tips and guidance. I have split the guidance into sections that follow the layout of the draft methodological framework.

Before I present these comparisons, when reading through the existing guidance it became apparent that I needed to account for when a conceptual modelling process starts; does it begin with the research conception, or, in the case of a large trial or project where the economic evaluation is only one part, later in

the research process? This is important because the point when the process starts dictates whether or not the conceptual modeller will decide on the research question, or whether it will have been decided at project conception, and the modeller 'inherits' the research question. I have included determining the research question in Phase 1 in my draft methodological framework, but also made it clear that determining the research question is not always relevant. I have included this in the final methodological framework in the 'Understanding the problem' stage.

Phase 1

Squires et al. suggests developing a conceptual model of the problem, bearing in mind causal links and current resource pathways. Developing a conceptual model of the problem is implied in Tappenden with his suggestion of developing a problem-oriented conceptual model. I haven't specifically suggested this in my draft methodological framework, however, I do discuss the benefit of SSM techniques at this stage and creating a rich picture. I think that it would be very useful when understanding the problem and communicating this understanding with stakeholders to have a diagram to use as a communication tool and include this advice in the final methodological framework. Two of the papers identified in the scoping review, Gray and Robinson (16, 17), suggest using a conceptual model of the problem as a communication tool, indeed, even if the modeller does not consult with stakeholders to agree on their understanding of the problem, I still think setting out an understanding of the problem in a visual form would be beneficial.

Squires et al. also recommend deciding on the research question and identifying sources of evidence in this phase. I have discussed the research question in the section above. Using evidence to understand the problem is included in my draft methodological framework, however I do not suggest identifying the sources of this evidence, although it would be good practice to include this information in the document produced with the conceptual model.

Squires et al. recommend examining existing health economics models, comparing structures, variables, results and identifying key variables that

influence the results; I include advice to examine previous conceptual models, but do not specify previous health economics models.

Squires et al. recommend choosing interventions and comparators, whilst Roberts et al. state that the interventions should be clearly defined; I have assumed in my draft methodological framework that the intervention and comparators will have already been decided in the research question, and therefore will not need to be considered. However, as discussed before, if the research question has not been decided, then choosing interventions and comparators is an important step in developing a conceptual model. Roberts et al. also recommend identifying outcomes (which are linked to the research question), perspective (the outcome is consistent with the perspective), and population.

Phase 2

Squires et al. explain that the model boundary is different to the problem boundary, and that the following should be considered: population, subgroups, perspectives and outcomes. Chilcott also recommends applying scope to the conceptual model. Roberts et al. recommend that the scope and structure of the conceptual model should be consistent with and address the problem. Tappenden implies this stage by asking the modeller 'is the breadth of the conceptual model complete?' In my draft methodological framework my advice when deciding the scope is to ensure the research question is answered, I also recommend that the modeller considers the perspective and type of economic evaluation.

Squires et al. and Chilcott recommend that the level of detail should be assessed, Squires et al. recommends assessing the impact of including more detail on results; the bigger the impact on the results, the more detail should be included.

Tappenden recommends developing the structure of the problem-oriented conceptual model using clinical guidelines and health professionals, I recommend this to understand the problem in the first phase.

Squires et al. recommends presenting a qualitative description of the conceptual model, in this guidance she doesn't specifically mention a document, however, this is implied from the title of the guidance ('Developing a qualitative description of the quantitative model'), and from the fact that she mentions documenting the conceptual model elsewhere in her guidance. She adds that the conceptual model diagram is a communication tool. Tappenden explains that the precise graphical approach taken in developing the diagram of the conceptual model is only important in that the diagram should be easily understood. He adds that the diagram and accompanying text should use non-technical language, and that key decisions should be clearly documented. Roberts et al. also recommend including a clear written statement of the decision problem, modelling problem and scope and ensuring that the policy context is clearly stated.

Finally, Tappenden recommends including health professionals who were not involved in the development of the conceptual model to check their understanding of it and whether it is clear. Roberts et al. recommend consulting with experts and stakeholders to ensure that the model represents the disease process and addresses the decision problem. This advice is analogous to my 'Review, revise and refine' stage in Phase 3.

Roberts et al. also discuss the link between a simple model with the right level of complexity, I include this in my general advice section at the start of the framework.

The existing economic evaluation conceptual modelling guidance also included guidance that was not relevant to my framework as it related solely to the development of a decision analytic model. This advice included:

- Tappenden and Squires et al. explain that the conceptual model may need to be altered when developing the decision analytic model, as the process is iterative.

- Tappenden suggested assessing alternative model development choices to run through the model in order to assess each structure's impact on the results. Squires has a step in her framework to choose the model type, Roberts et al. add that several model types may be suitable; for simple problems (short time frame, or few outcomes), a decision tree may be appropriate; for problems involving a series of health states a transition state model would be appropriate, any interactions between individuals should be evaluated, resource constraints should be represented, and a combination of model types may be suitable for some problems.

- Roberts et al. also recommend that the time horizon should be long enough to capture relevant differences in outcomes and that sensitivity analysis can be used to assess the impact of key uncertainties in the model structure. Finally, Roberts et al. explain that is it an explicit process to convert the problem conceptualisation in a model structure.

Summary of comparing published economic evaluation methodology with my draft methodological framework

Table 13 summarises and compares which stages are in my draft methodological framework and are also included in the existing economic evaluation guidance. All of the stages in the draft methodological framework are included in two or more publications of the existing guidance. The stages 'Understanding the problem' and 'Determining the scope' are included in my draft methodological framework and in all existing guidance. Therefore, five stages in the draft methodological framework are not in all of the publications of the existing guidance: 'Setting objectives', 'Determine the detail', 'Determine the content', 'Diagram and documentation' and 'Review, revise and refine'. The methodological framework published by Squires et al. includes all but the 'Review, revise and refine' stage in my draft methodological framework.

Table 13: Summary of stages included in draft methodological framework and existing economic evaluation guidance on conceptual models

Stage	Draft methodological framework	Squires et al.	Tappenden	Roberts et al.	Chilcott
Understanding the problem	✓	✓	✓	✓	✓
Setting objectives	✓	✓	×	✓	×
Determine scope	✓	✓	✓	✓	✓
Determine detail	✓	✓	×	×	✓
Determine content	✓	✓	✓ not explicit	✓ structure only	×
Diagram and documentation	✓	✓	✓	✓ document only, not diagram	×
Review, revise and refine	✓	×	✓	✓	✓

The main difference between the published economic evaluation guidance and my draft methodological framework was the purpose; the purpose of the published economic evaluation guidance was to recommend methods for the development of a conceptual model that aids the development and

implementation of a subsequent decision analytic model. Whereas the purpose of my draft methodological framework is to develop a conceptual model for the new role proposed in this thesis.

The advice included in the existing economic evaluation guidance was more prescriptive and more focussed on economic evaluations and decision analytic models than the recommendations in my draft methodological framework. I have taken some of this methodology and guidance and incorporated into my methodological framework to refine it. Most of this additional guidance relates to the 'Understanding the problem' stage: patient representatives were added as suggested team members; developing a conceptual model of the bigger picture has been added; resource use has been emphasised; using clinical guidelines to gain understanding of the problem area and systems has been suggested, and questions to guide understanding the problem have been added. Additions in other stages of the methodological framework include: setting the scope to include components that are hypothesised to have an effect on the results; only including key assumptions in the document, and including specific aspects of understanding the problem (document stage).

5.5 Results – Final methodological framework

5.5.1 Introduction

General advice

This methodological framework is a guide to develop conceptual models for use in economic evaluations. The methodological framework is not designed to be exhaustive or prescriptive, rather a guide to developing a conceptual model which should be applied to each study in a pragmatic way and tailored to each research problem. Developing conceptual models is considered an 'art rather than a science' and each conceptual model will be different, dependant on the research problem and modeller.

Because the research theme and question may be predefined (as is often the case in clinical trials), some aspects of the methodological framework may not be relevant to all circumstances.

Chapter 5
How to use the methodological framework

The methodological framework is based on two key concepts fundamental in developing all conceptual models: an iterative process and keeping the conceptual model simple. The modeller should iteratively revisit stages in the framework until the conceptual model is complete; each stage can be revisited at any time and from any stage. The conceptual model should be kept simple; it should contain enough detail to answer the research question, but not too much, otherwise it becomes unwieldy and contains unnecessary information. Both concepts should be borne in mind by the modeller throughout the conceptual modelling process.

The methodological framework is split into three phases; Context, Development and Finalising, within these phases are discrete stages making up the stepwise structure of the methodological framework, these are illustrated in Figure 30. The diagram depicts the methodological framework, in a clockwise direction (starting from '12 o'clock'), in three circles; the inner circle comprises the phases, the middle circle comprises the stages and the outermost circle presents an descriptive overview of each stage. Each stage consists of a 'Recommendation' and an 'Explanation'; the Recommendation states an information request and rationale for each stage (except Stage 7), and the Explanation contains guidance for achieving the 'Recommendation'. The output from the information request should be included in the conceptual model document. The modeller should work through the stages, revisiting previous stages iteratively when needed. The outputs from this methodological framework are a conceptual model diagram and document depicting and explaining the conceptual model. Throughout the conceptual model development process the modeller will add information to the document, this is described in more detail in Stage 6 'Diagram and documentation'.

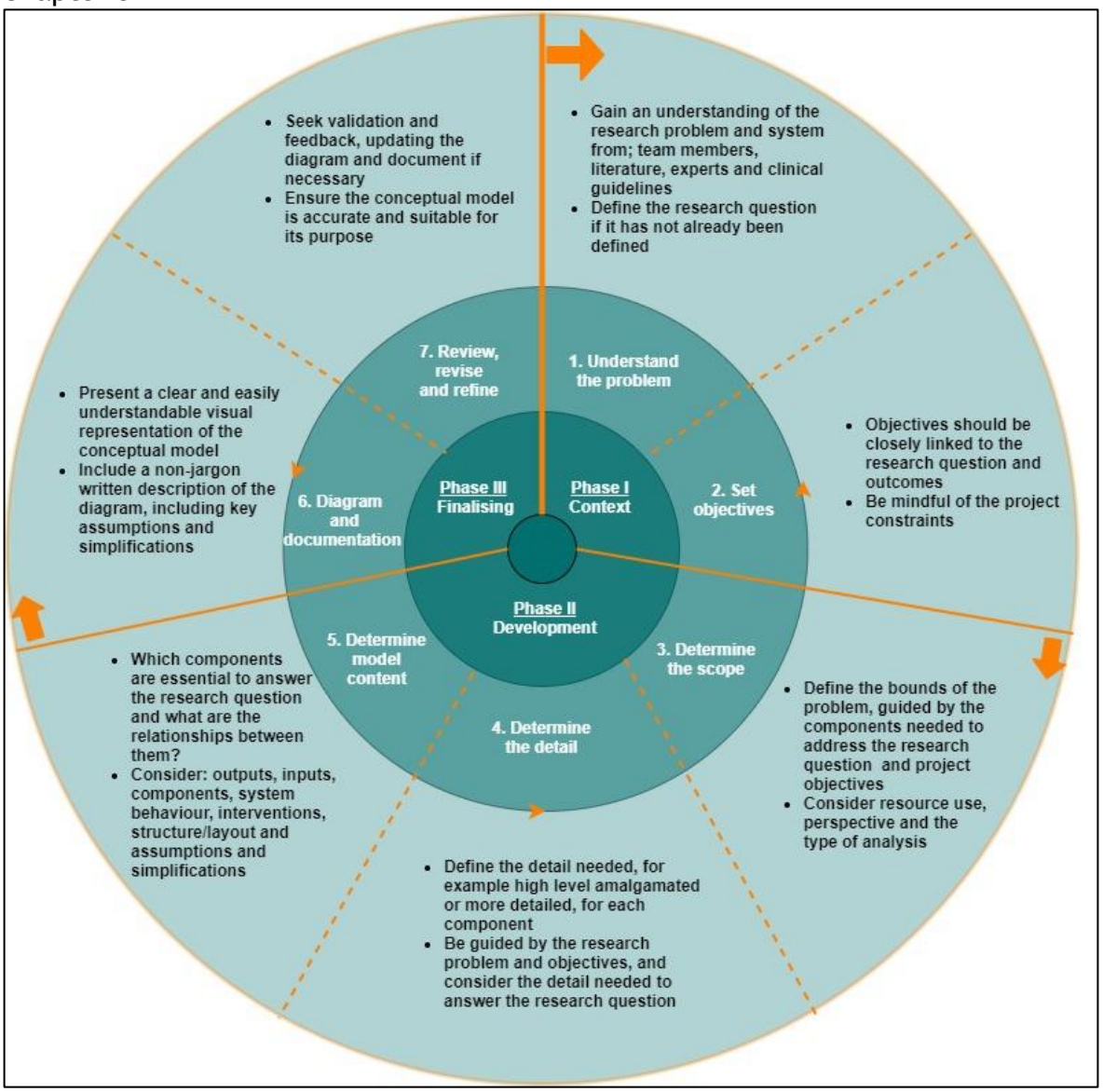


Figure 30: Methodological framework diagram

5.5.2 Phase I - Context

In this phase the modeller considers the context of the conceptual model, immersing themselves in the research problem, gaining an understanding of the problem system, assembling a project team, and setting the research question and objectives.

Stage 1: Understanding the problem

Recommendation

The information request for this stage is a description of the research system, problem, project team and question. The modeller should gain a good understanding of the bigger picture of the research problem, this involves looking at the wider systems and subject area in context to the research problem; without a sound understanding of the research problem it is not possible to develop an accurate conceptual model. A diagram of the problem system is a useful communication tool at this stage to check understanding of the problem and to use during the development of the conceptual model. The project team should be identified during this stage if they have not been identified already and if the research question has not already been defined it should be determined during this stage.

Explanation

Gaining an understanding may require identifying patient pathways, patient behaviour or disease pathways and particular attention should be given to identifying possible categories of resource use. Understanding can be taken from several sources: the literature, stakeholders, decision makers, experts, existing conceptual models, trial protocol (if relevant) and clinical guidelines. Questions useful to help the modeller understand the research problem are presented in Table 14.

A specific element of soft systems methodology can be used in this stage to come to an understanding of the problem systems; developing a rich picture,

this is an informal, often hand drawn, diagram of the problem area, examples include the disease area or treatment pathway.

The makeup of the project team will vary depending on the research problem and will typically comprise; the modeller, other health economists, clinicians with an expert knowledge of the disease area, patient representatives and study or trial team members (for example trial manager). Whilst it is good practise for the project team to be involved at this stage, it may be practical to only include them at the later 'Review, revise, refine' stage. However, if the project team are involved at this stage the project team and the modeller should come to a consensus on the research problem and question; the modeller can suggest alternative hypotheses to the project team at this stage to help reach an understanding of the research problem.

The modeller will need to make assumptions when coming to an understanding of the problem; any key assumptions made in reaching this understanding should also be recorded in the conceptual modelling document.

Table 14: Questions to help understanding the problem

Questions to guide understanding the problem
What is the problem?
Why is it a problem?
Who are the target population?
What are the interventions and comparators to be included?
What are the outcomes?
What perspective will be taken?

Stage 2: Setting objectives

Recommendation

The information request for this stage is a description of the objectives of the conceptual model; setting objectives is key in guiding the development of the conceptual model and defining the model content.

There are two types of objectives to consider: modelling and project.

The modelling objectives are closely linked to the research question and what the model should achieve, including the research problem outcomes. The outcomes may include hospital services and throughput, disease progression or patient behaviour.

The project objectives relate to the resources available to the modeller; constraints on time and budget in the project, the modeller should ensure that the project objectives are realistic, and expectations are properly managed.

The modeller should agree the objectives with the model team (if applicable) and record them in the conceptual model document.

5.5.3 Phase 2 - Development

In the second phase the modeller decides what to include and exclude in the model, which components best represent the research problem, the dynamics of these components and how they are linked. During this phase the modeller should keep in mind the objectives set in Stage 2 and the project hypothesis, if there is one, to ensure the conceptual model is kept relevant. The development of the conceptual model will be an iterative process, with the modeller revisiting stages until it is complete.

Stage 3: Determine the scope of the conceptual model

Recommendation

The information request for this stage is a description of the scope of the conceptual model. The scope bounds the research problem, limiting the conceptual model to the elements of the bigger picture needed to address the research question and achieve the objectives.

The modeller should see the 'Understanding the problem' stage as getting to know the 'bigger picture' of the problem; understanding the wider subject area that includes the research problem, whereas the 'Determine the scope of the conceptual model' stage is when the narrower scope of the conceptual model is considered. The scope guides the development of the conceptual model, helping the modeller decide what should be included and excluded from the bigger picture to answer the research problem.

When determining the scope, the modeller should consider the research question, project hypothesis (if there is one) and conceptual model objectives; only components important and relevant to these should be included. The modeller should also consider which components may influence the results and include these, however, care should be taken not to make assumptions prior to the analysis. If irrelevant components are included in the conceptual model it will become too complicated and unwieldy, creating 'noise', and making interpretation difficult.

The modeller should consider resource use, the perspective taken for the costs and outcomes and the type of economic evaluation (cost-utility, cost-effectiveness or cost-benefit), as all these considerations will determine the components to be included in the scope of the conceptual model.

Each of the components identified as within the scope of the conceptual model should be assessed by evaluating the relevance of these components in answering the research question.

The modeller and project team (if appropriate) should come to a consensus on the scope of the research problem and record this in the conceptual model document.

Stage 4: Determine the detail of the conceptual model

Recommendation

The information request for this stage is a description of the detail of the conceptual model. The detail determines how far the modeller will drill down into individual components of the conceptual model; depending on the objectives the detail may be high level, aggregated and simplified, or may include disaggregated levels of components.

Explanation

The modeller should determine the detail in relation to the research question, project hypothesis and conceptual model objectives. The modeller should consider the components included in the conceptual model, including health events and resource use data needed to answer the research question and the detail needed to capture this. For example, resource use may be amalgamated into one component or disaggregated into separate resource use categories such as treatment, resource directly related to disease area and other resource use.

Like the 'Determining the scope of the conceptual model' stage above, the detail of each component should be assessed on their relevance in answering the research question.

Stage 5: Determine the content of the conceptual model

Recommendation

The information request for this stage is a description of the trial mechanism, dynamics, and how the components in the conceptual model interact. This stage is arguably one of the most important and the most complicated in the methodological framework, with several aspects that the modeller should consider.

The **components** within the scope of the conceptual model should be determined, along with how they interact or are linked to each other (**system behaviour)**, these links can be causal or an association. Components are directly linked to the Scope and Detail stages and the decisions made in those stages will feed into this stage, they are also are closely linked to the research question and objectives. Components can either be passive (static) or active within the flow of the system, and can be classified into three groups; components which feed into the conceptual model **(inputs**), components in the body of the model illustrating the **system behaviour**, and components which are the product of the system behaviour (**outputs**). **Model inputs** are the components within the model which can be altered to represent different scenarios, examples include patients, health facilities and treatments. **Model outputs** can be used to check that the objectives of the conceptual model have been reached, key model outputs are costs and health benefits; examples include an untreated patient becoming a treated patient, quality of life measures, clinical effects (such as strokes avoided or cancer cases detected), or patient behaviour such as increase in physical activity.

The way in which the **interventions** alter, or are hypothesised to alter, the **system behaviour** should be established and included in the dynamics of the conceptual model. The overall **structure/layout** of the model should be decided to best illustrate the trial process or system based on the research question and modelling objectives. The **model structure** will be driven by the objectives and research problem, for example a simple patient pathway may be relevant, or disease progression. When determining the model structure, the modeller should consider resource constraints and capturing relevant resource use categories and outcomes. A conceptual model is a simplification of a real-world system, therefore assumptions and simplifications are important features of this stage; assumptions relate to limited knowledge of, or evidence of the research problem bigger picture. The more assumptions made; the less detail included in the conceptual model. Simplifications result from keeping the conceptual model straightforward and uncomplicated.

A description of the model content and assumptions and simplifications made to determine the model content, should be recorded in the conceptual model document, this section of the conceptual model document is key in explaining and describing the conceptual model, and particularly important for communicating with stakeholders.

5.5.4 Phase 3 - Finalising

In this phase of the methodological framework the conceptual model diagram and document are developed, then finalised with a 'Review, revise, refine' stage where the conceptual model is evaluated, and any new information is used to refine the existing conceptual model.

Stage 6: Diagram and documentation

Recommendation

The information request for this stage is the conceptual model diagram and an accompanying document containing background information on the research problem, and a description and explanation of the conceptual model. The diagram should be a clear, accurate and relevant visual representation of the research problem, and as a communication tool should not leave the reader having to make assumptions. The document, a non-jargon written description of the conceptual model, should help the reader to understand the model, and can be used as a communication tool, along with the diagram.

Explanation

The diagram should be clear and understandable, and it is likely that the modeller will need several iterations to develop it. The diagram will show the components and how they interact and interconnect with each other, causally or otherwise. If a rich picture has been used in the 'Understanding the problem' or 'Determine the content of the model' stages, it will be useful in developing the final conceptual model diagram.

The purpose of the document is to help the user to understand the model, it does not need to be extensive or overly comprehensive. The document can act as a living document to the modeller and a short version focussing on the model content is helpful as a communication tool. Information recorded at each stage of the conceptual model development should be included in the document, this is summarised in Table 15, and key assumptions and simplifications made at each stage should be reported.

The conceptual model diagram and document can be included in the Health Economics Analysis Plan, if applicable.

Table 15: Items included in the conceptual model document

Conceptual modelling stage	Output for diagram/document
Understanding the problem	Diagram of the problem system, with a written description of the research problem, problem system and the sources of evidence used in coming to this understanding Research question List key assumptions and simplifications made
Setting objective	Description of objective List key assumptions and simplifications made
Determining the scope of the conceptual model	Describe the scope of the conceptual model List key assumptions and simplifications made
Determining the detail of the conceptual model	Describe the detail of the conceptual model List key assumptions and simplifications made
Determining the content of the conceptual model	The model content (explanation of inputs/outputs, components and their relationships), and assumptions and simplifications made in developing the content of the conceptual model
Diagram and documentation	A diagram of the conceptual model, accompanied by a non-technical document
Review, revise and refine	Validation methods and update the diagram and document with any changes as a result of the validation

Stage 7: Review, revise, refine

Recommendations

In this stage the conceptual model should be reviewed, seeking agreement and feedback on the conceptual model from the project team, this is an important step without which the conceptual model may not be accurate or useful. Any suggested changes during this stage should be used to revise the model, strengthening it, and making it robust and suitable for its intended purpose.

The modeller and project team should check the conceptual model to ensure it accurately depicts and reflects the teams' understanding of the research system and question, it is good practice to involve experts who were not involved in the development of the conceptual model in this stage. The conceptual model, and any assumptions and limitations made, should also be checked in terms of clinical accuracy, logic, presentation and ease of understanding. Table 16 lists questions the modeller and project team can ask themselves during this stage to review the conceptual model. This is an iterative process where more than one version of the conceptual model is likely to be reviewed.

Table 16: Validation questions for conceptual model

Diagram:
Is the diagram well defined, logical and transparent? Does the conceptual model reflect the research question and subject area system?
Document:
Is there a clear description of the research problem and question, including background information to help understand the problem? Is there a clear objective reported? Are the scope and detail considered relevant to the research question and objectives, and are they justified? Is there a clear and understandable explanation of the content of the conceptual model? Are the key assumptions and simplifications made explicit? Does the user need to make assumptions about the conceptual model to understand it? Are changes made to the conceptual model during the 'Review, revise, refine' stage recorded in the document?

Any revisions made to the conceptual model after it is reviewed, should be reflected in the diagram and document. If new information comes to light after

the development of the conceptual model the conceptual model should be updated and refined if the new information is relevant.

5.6 Discussion

5.6.1 Overview

From 859 potential records identified in the literature search, 16 frameworks were included in the scoping review results. These records were from seven different fields, and 18 themes were extracted from the records and used to inform the draft methodological framework. The sequence of the themes reported in the included frameworks was extracted to ascertain whether there were patterns or consistency in these sequences, by looking at this evidence it was established that the themes were split into three broad phases: 1) An initial phase where the modeller immersed themselves in evidence about the research problem, identifying the project team and setting objectives for the model and project; 2) a middle stage which considered the relevant components to include in order to answer the research question in the detail needed, and 3) the final stage included producing a diagram and document to represent the conceptual model, plus a stage to assess the conceptual model and revise it if necessary.

I combined descriptions and recommendations from the stages identified in the included records in the development of the draft methodological framework. I then added economic evaluation specific advice and compared the draft methodological framework with published guidance for developing conceptual models in economic evaluations. This comparison was difficult to undertake as some the existing guidance did not contain discrete stages, and some of the guidance was implied and only mentioned when discussing other aspects of the guidance. However, all stages included in my proposed framework were included in two or more of the existing guidance. Where gaps were identified refinements were made to the draft methodological framework if necessary, to produce the final methodological framework presented.

Scoping review methodology was chosen for this literature review after considering and rejecting using a systematic review. Systematic reviews are traditionally used to answer a specific question according to a rigid set of a

priori factors, they have narrow parameters, the quality of identified studies is assessed formally and there is detailed data extraction, with synthesised quantitative and qualitative data. Whilst scoping reviews are often undertaken for 'reconnaissance' purposes to assess whether a full systematic review would be worthwhile, they can also be conducted as 'exercises in and of themselves' to summarise and present research results particularly where a research area is complex or has not been studied before and identify gaps in the research or recommend areas for future research (4). Scoping reviews have a broader approach than systematic reviews, mapping literature and addressing broader research questions (4). The objective of a scoping review is to map key 'concepts that underpin a research area' (5), assess the main types and sources of record available. This requires a broad range of records, but not a great depth of analysis. Scoping reviews are designed to give an overview of the records in a research area, without assessing the quality of the records. The objectives of a scoping review are a good match for the objectives of my literature review. The research question of the scoping review was 'What reported steps are used to develop conceptual models in methodological frameworks, and is there enough commonality in these steps to combine them into a single methodological framework for developing conceptual models in economic evaluation?' I found that although the included methodological frameworks came from different fields there were enough similarities in them that, when the themes (steps) were extracted, it was possible to combine them into a draft methodological framework.

The overarching rationale of the identified frameworks was to fulfil an unmet need for a framework to guide the conceptual modeller in developing good quality and relevant conceptual models. The rationale for this thesis is that there is a lack of guidance for developing conceptual models in the field of economic evaluation for the purpose of the proposed new role, this lack of guidance similar to the rationale reported in many of the included records.

Despite the frameworks being heterogeneous in their origin and form; from seven diverse disciplines and five different types of publication, there was consistency in the steps included in them and I was able to allocate these steps to 18 different themes, with only one theme having one mention. Out of the 16

frameworks identified, six were identified in the Google search (two from the 'development' search and four from the 'steps' search). This result suggests that whilst there is no gold standard for conducting an internet search in a literature review, it does produce relevant sources of evidence.

Setting objectives was the most frequently reported theme, followed by scope, content, system behaviour, understanding the problem and inputs, all these themes were reported in 50% or more of the frameworks. Amongst the least frequently mentioned themes were team identification, review, revise and refine, validation and diagram. The lack of frequency of including a diagram as a step in a conceptual modelling framework was surprising as one of the top benefits mentioned for conceptual models is as a communication tool, and eight of the 16 identified methodological frameworks included a diagram.

As far as I can ascertain, this is the first literature search completed for the purpose of identifying conceptual modelling frameworks in fields other than economic evaluation, with the aim of extracting data from those frameworks to inform a methodological framework for developing conceptual models in the field of economic evaluation. Previously in Squires et al. , a literature search was completed to inform a conceptual model framework for public health economic evaluations, however this was limited to frameworks where the aim of the conceptual model was to develop a decision analytic model (22). This search was also not restricted to stepwise methodological frameworks. Squires et al. identified eight frameworks, two of which were included in my results above, the rest did not include a framework, only methodology. Furthermore, Squires et al. did not methodologically extract steps from a framework, they focussed on the methods discussed and amalgamated those.

5.6.2 Strengths

The strengths of my literature review were: I consulted with a qualitative researcher to identify suitable methodology for the purposes of the literature search; I followed scoping review methods; I applied methodological rigour even when there was scarce guidance for suitable methodology when conducting the internet engine literature search, and I consulted with a librarian over the search terms and sources to search. I also followed the suggestions from

Chapter 4 to develop the methodological framework and evaluated it by comparing it to existing guidance.

5.6.3 Limitations

I needed a broad search for potential frameworks enabling me to identify them from different disciplines and from different sources. This involved using less established search techniques such as using an internet search engine. Whilst there is published methodology on this technique, it is not well established and there is little guidance on best practice and how to carry out and present these searches.

I limited my search terms to 'titles', although this may have excluded some frameworks it was practical and by including four sources in my search I made the search as comprehensive as possible without it being unwieldy.

A further limitation of the search was that specific terms for 'logic model' were not included, potentially decreasing the number of frameworks identified, and therefore the steps used in developing them. Logic models are types of conceptual models, typically used in public health complex interventions, (183, 184) and visually represent anticipated causal links between the intervention and outcomes. They are recommended and encouraged for use in public health interventions (184, 185). Logic models are based on programme theory or theory of change, showing how the intervention works and in what population (183). Logic models can be simple or complex, with the simplest examples showing linear relationships between intervention, short-term, medium-term and long-term outcomes (184). There is clear crossover between conceptual models and logic models, with logic models a specific type of conceptual model used in public health interventions. Extracting the steps from the frameworks and amalgamating them into stages was an iterative process and interpretation was needed to decide on which stage was suitable for some steps that were ambiguous. This made the research hard to reproduce as different researchers may have different interpretations. However, there were not many ambiguous steps and I have been open about the names of the steps and where they were allocated, so the reader is able to make their own judgement on my interpretation.

5.6.4 Further research

The methodological framework needs further evaluation, which was outside the scope of this PhD, this could be a Delphi panel to come to a consensus on the content of the methodological framework, a focus group to discuss the methodological framework and suggest improvements and refinements, or with further case studies to validate it and suggest refinements.

5.7 Summary

This chapter used the suggestions for creating methodological frameworks described in Chapter 4 to develop a methodological framework for producing conceptual models in economic evaluation, fulfilling the second objective and producing the second output of the thesis. A scoping review identified 16 existing methodological frameworks for developing conceptual models and the steps and approaches in these frameworks were extracted and synthesised to produce a draft methodological framework. The methodological frameworks identified in the scoping review were from seven different fields outwith economic evaluation, and although the purposes of the methodological frameworks differed 18 discrete themes were extracted with enough similar steps to produce the draft methodological framework. This chapter demonstrates that it is possible to produce a methodological framework based on disciplines outwith economic evaluation.

This draft methodological framework was compared to existing guidance for developing conceptual models in economic evaluation, the main difference was the purpose of the conceptual model, the comparison also highlighted gaps in the draft methodological framework which were filled when it was refined.

The final methodological framework was presented and is the first to provide guidance for the development of conceptual models in economic evaluation that do not precede a decision analytic model.

In the next chapter the methodological framework guidance is applied to the two case studies in this thesis, to develop conceptual models for use in the proposed new role.

Chapter 6 Applying the methodological framework to developing conceptual models

6.1 Chapter overview

In the previous chapter a methodological framework for developing conceptual models was introduced; the aim of this current chapter is to demonstrate the methodological framework by applying it to the thesis case studies, in illustrative examples, to develop two conceptual models, which are subsequently used to illustrate the new role in Chapter 7. The demonstration of the methodological framework partly fulfils the second objective of the thesis.

The layout of the chapter is as follows: Section 6.2 applies the methodological framework to the TWICS case study, Section 6.3 applies the methodological framework to the BeatIt case study, Section 6.4 discussion limitations and in Section 6.5 the chapter is summarised.

6.2 Case study: application of the methodological framework - TWICS

This section of the chapter describes applying the methodological framework to the TWICS case study which was introduced in Chapter 3. In sub-section 6.2.1 the conceptual model development process is described, and the final iterations are presented. The initial iterations of the conceptual model are presented in Appendix 11: Iterative development of TWICS case study conceptual model diagram (*Chapter 6)*, and the output of diagram and documentation is presented in sub-section 0.

6.2.1 Application of methodological framework

This sub-section describes the development process step by step, presenting each stage separately, a reminder of the stages of the methodological framework the diagram is included in Figure 31.

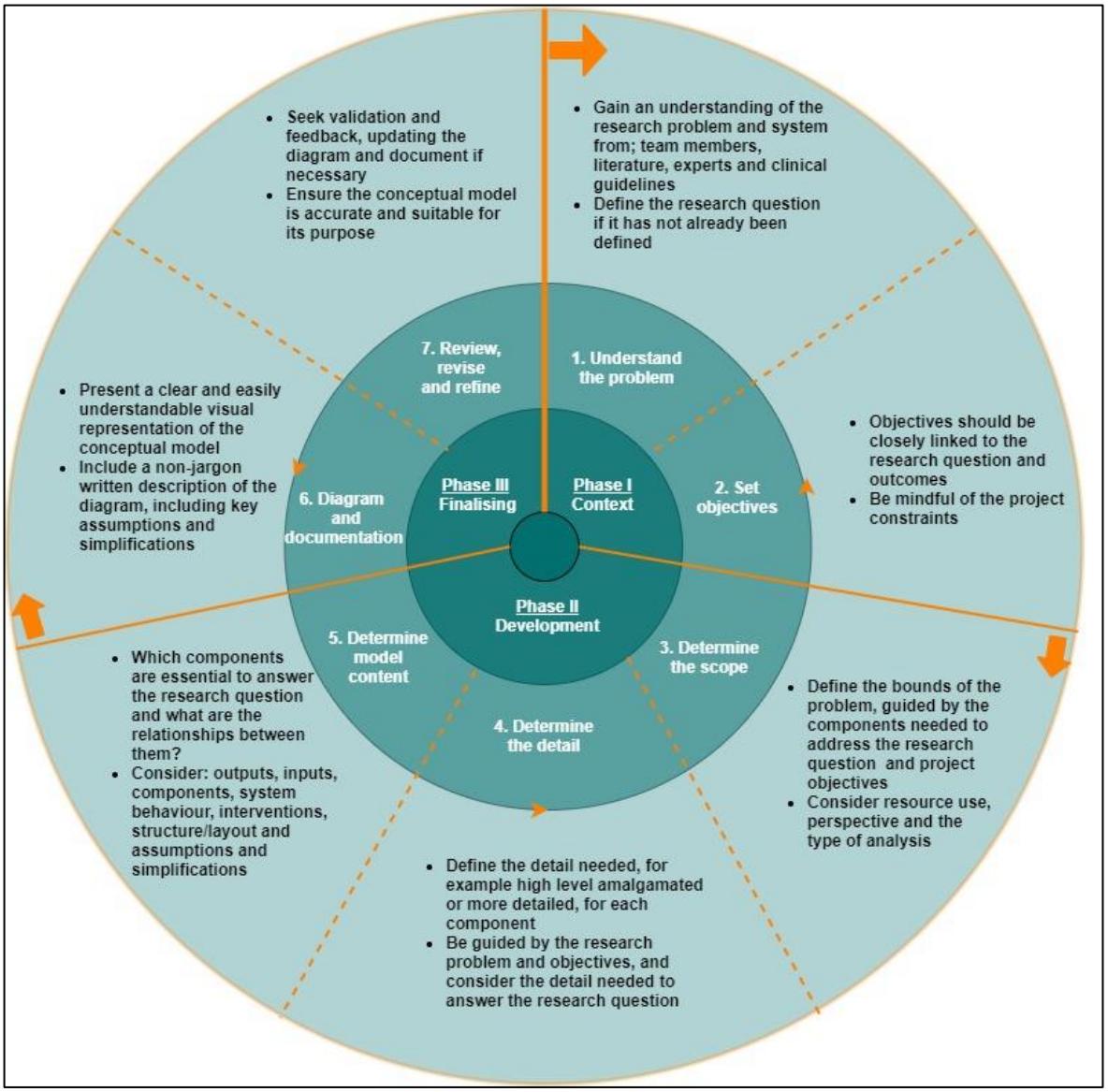


Figure 31: Methodological framework diagram

Stage 1: Understand the problem

The recommendation in this stage is to describe the relevant components of the decision problem, these include understanding the context of the research system and the problem. Other aspects of this stage include developing a rich picture, determining the project team and defining the research question.

To understand the research problem the questions presented in Table 14 (Chapter 5) were considered and answered in the conceptual modelling document. Evidence was taken from the TWICS protocol (111), and published literature.

The next step was to develop a rich picture of the disease area to help understand the disease area; a rich picture is an informal diagram of the problem area. To do this I took my understanding of COPD from the questions above, and from existing diagrams depicting COPD and disease progression. Disease progression is illustrated by Hoogendoorn et al. in Figure 32 (186).

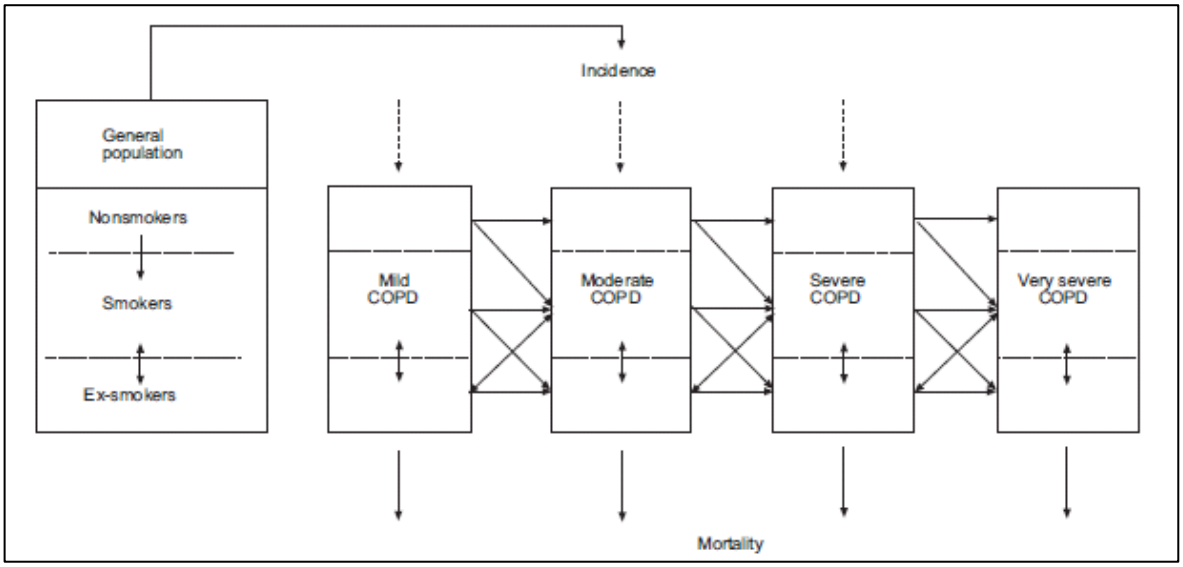


Figure 32: Dynamic population model of disease progression in COPD from Hoogendoorn et al. 2005 (186)

A diagram by Borg et al. (Figure 33) illustrates that exacerbations can occur during any stage of the disease, that COPD is progressive, and the severity of COPD increases as the number of exacerbations increase (187).

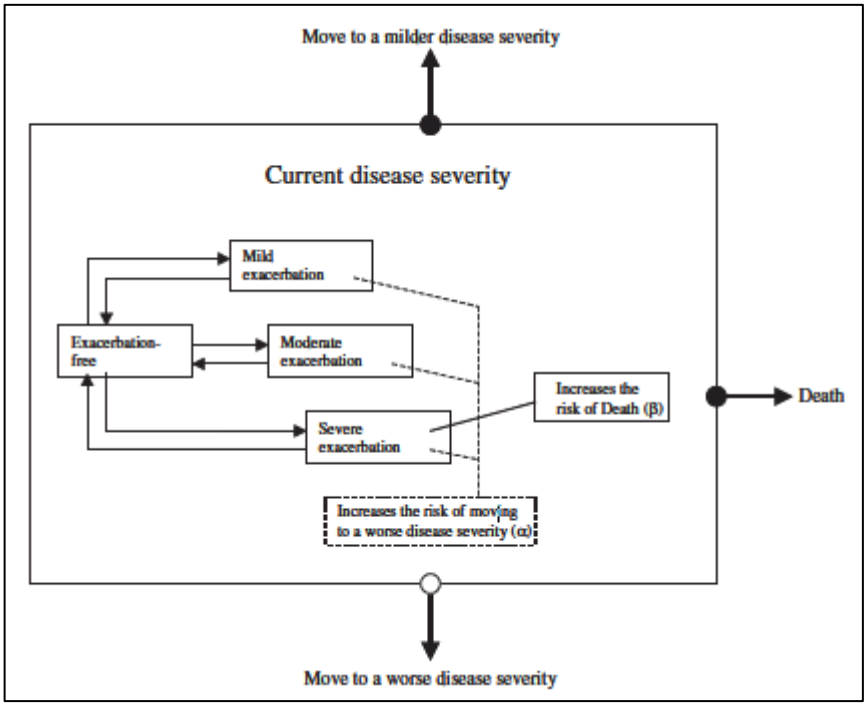


Figure 33: A computer simulation model of the natural history and economic impact of COPD from Borg et al. 2004 (187)

The rich picture is presented in the conceptual model document in Figure 42.

Finally, in this stage the research question for the conceptual model was defined and the team determined.

Stage 2: Setting objectives

The recommendation in this stage is to describe the objective of the conceptual model, this was closely linked to the research question.

Stage 3 Determine scope of the conceptual model

The recommendation in this stage is to describe the scope of the conceptual model. In this stage I narrowed down the elements included in the rich picture relevant to the research question and objectives. The trial mechanism did not

include any measure of disease severity so that was omitted; the inclusion and exclusion criteria of TWICS restricted participants to people with moderate or severe COPD and at a high risk of exacerbations requiring treatment, the aim of the trial was to assess the effect of theophylline on the number of exacerbations requiring treatment in this population. The scope was also driven by the data available from TWICS.

Stage 4 Determine the detail of the conceptual model

The recommendation in this stage is to determine the level of detail of the conceptual model. This was driven by the objectives and available data from the TWICS clinical trial.

Stage 5 Determine the content of the conceptual model

The recommendation of this stage is to describe the trial mechanism and how the key components of the mechanism interact.

The iterative nature of deciding components, system behaviour, structure and applying assumptions and simplifications was experienced in developing the conceptual model content and diagram and the initial interactions are described 0, the conceptual models developed at each iterative stage are also included. The final conceptual models were based on the disaggregated cost-effectiveness template presented in Figure 3; first a full conceptual model was built up to comprise all components in the template, then the components not considered as important for the conceptual model driven analysis were dropped from this interim conceptual model to produce the final conceptual model. Each step of the conceptual model content and diagram development is illustrated and explained in the following paragraphs, the final diagram and document is presented in the 'Stage 6: Diagram and document' section.

Components included in the model comprise: Model input of treatment arm; model outputs are closely linked to the research question and hypothesis as suggested – treatment and non-treatment costs and quality of life as measured by QALYs. Exacerbations act as a mediator linking the inputs and outputs.

The first step was to develop a conceptual model of COPD, this simplified the disease and included key components of: symptoms, lung capacity, exercise capacity and exacerbations (Figure 34), a combination of these components predict survival. The associations between these components are represented by arrows: COPD symptoms experienced by the patient with COPD predict lung and exercise capacity, and exacerbations, these three components in turn predict survival. Lung capacity predicts survival and exacerbations, with a backward association of exacerbations predicting lung capacity (exacerbations worsen a patient's lung capacity). Exercise capacity predicts survival and exacerbations.

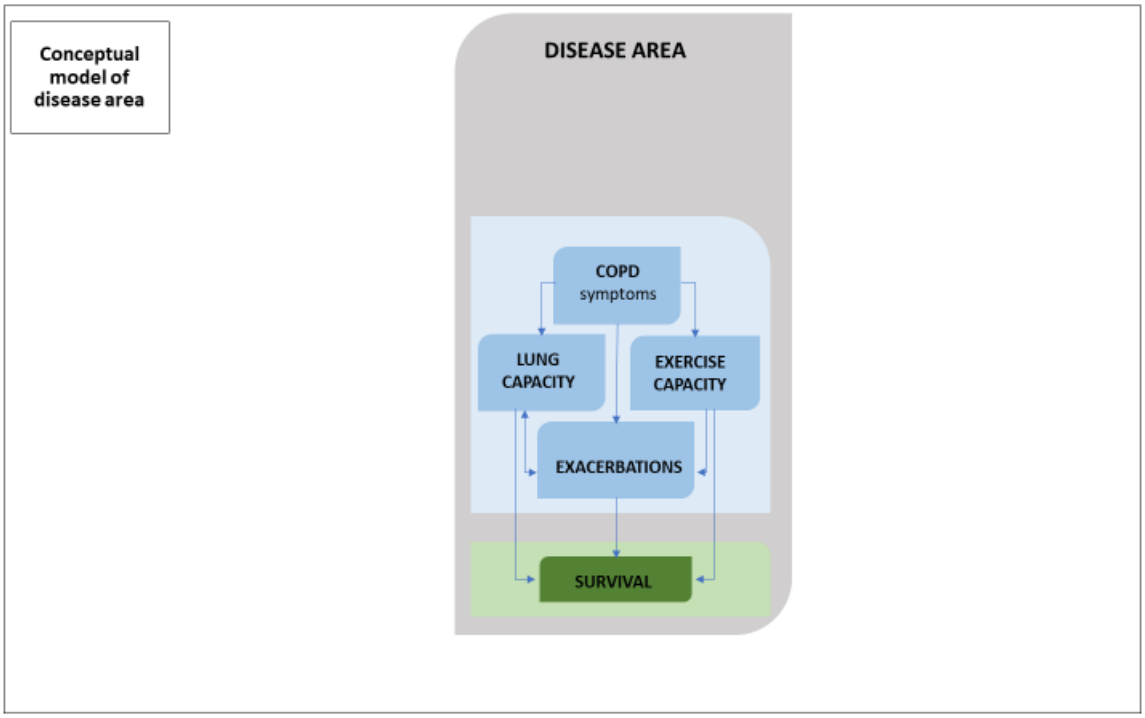


Figure 34: COPD conceptual model building - disease area

The next step was to include healthcare costs related to COPD (Figure 35), two categories of healthcare costs are added; costs associated directly to COPD that result from disease severity and COPD related health events, and costs associated with improved survival.

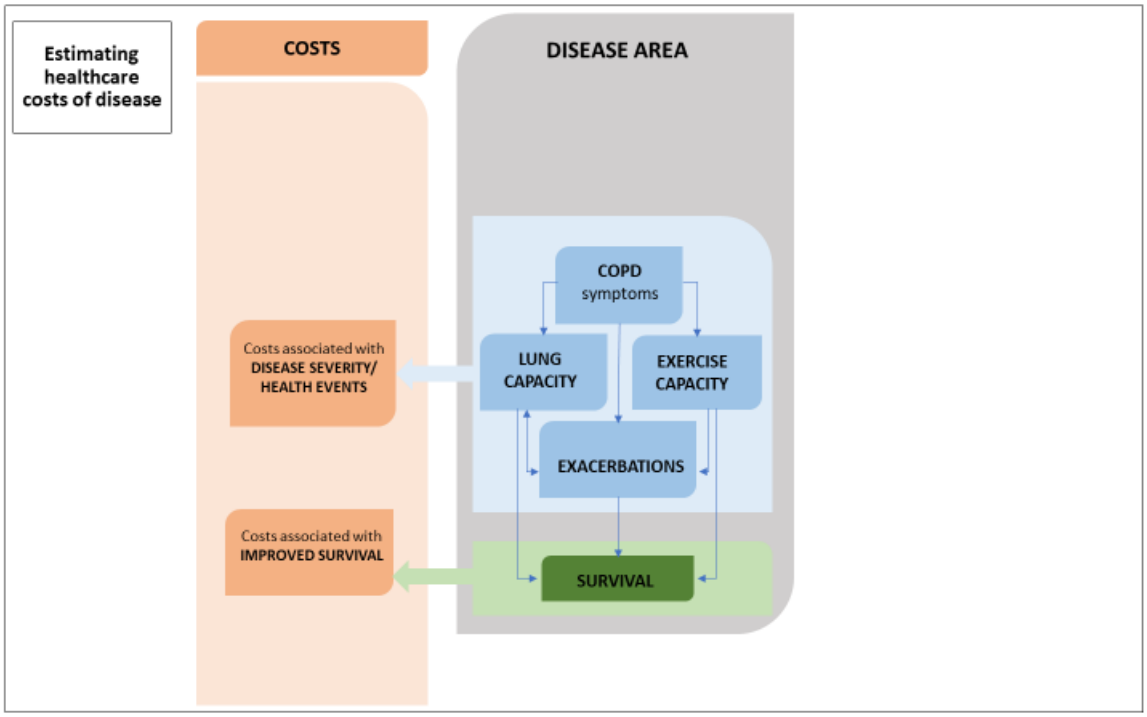


Figure 35: COPD conceptual model building - healthcare costs

Then health benefits in terms of health-related quality of life relating to COPD were added (Figure 36). The grey health-related quality of life utility panel represents the step in calculating QALYs where clinical events are measured using utilities. Clinical events are expected to result from the two elements of the disease area: symptoms, lung capacity, exercise capacity and exacerbations, and survival. The yellow 'Health benefits' panel represents utilities translated into QALYs.

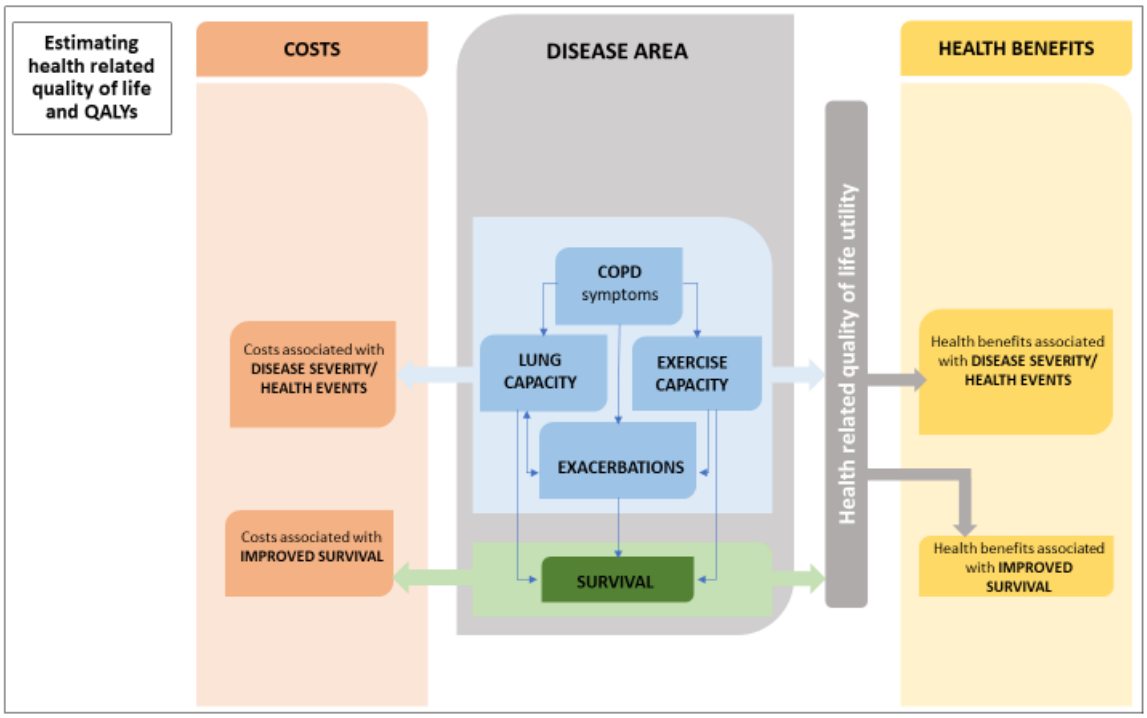


Figure 36: COPD conceptual model building - health related quality of life and quality adjusted life-years

The next step is to layer in a generic treatment to include the treatment effects consistent with the disaggregated cost-effectiveness template (Figure 37). Layering in the generic treatment (and comparator) illustrates the incremental analysis of costs and health benefits; all cost and health benefit components are now incremental. In the 'Disease Area' panel a new red section is added with two components: the generic treatment itself and side effects resulting from the treatment. These components are associated with costs and health benefits: potential costs are those associated with treatment side effects and those directly linked to treatment, and health benefits are the quality of life lost as a result of treatment side effects.

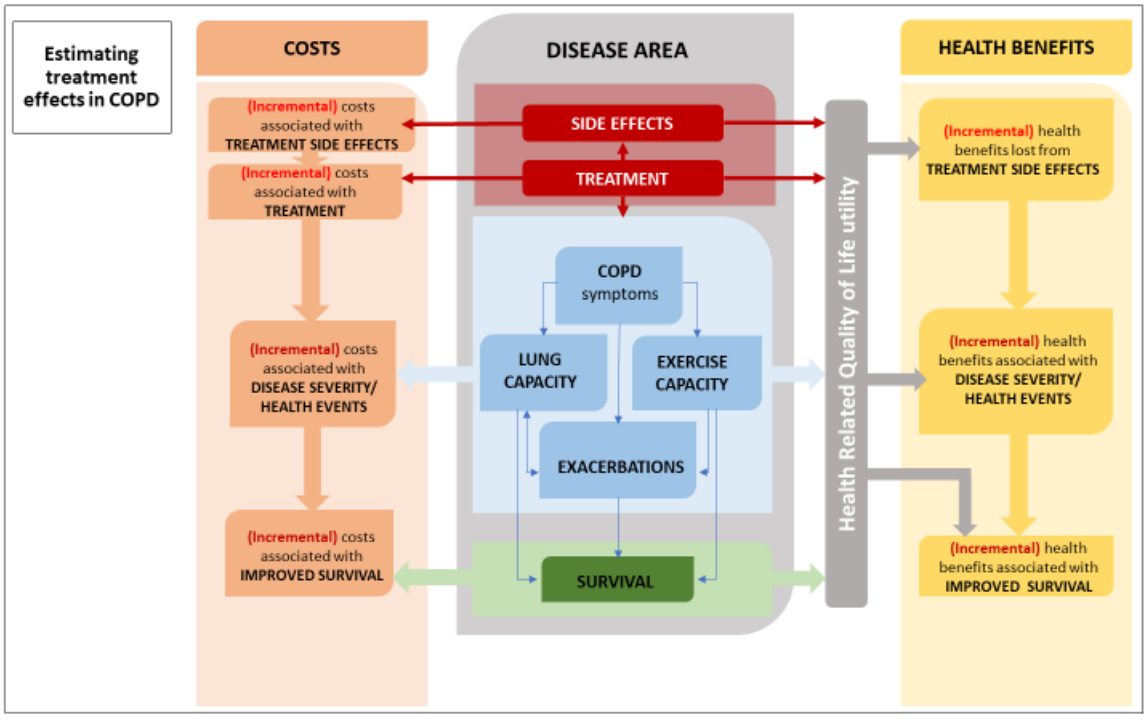


Figure 37: COPD conceptual model building - generic treatment

Once generic treatment effects are included in the conceptual model the cost-effectiveness can be illustrated (Figure 38). This is depicted in the bottom row of the diagram where all cost components are summarised in 'Net incremental costs' and all health benefits are summarised in 'Net incremental quality adjusted life-years'. These summary components feed into the economic evaluation summary measure of cost per quality adjusted life-year (the ICER). This diagram in the conceptual model building process depicts a full economic evaluation. The next step is to adapt it to TWICS, replacing the generic treatment with theophylline and taking out unnecessary elements that are not expected to effect or be relevant to theophylline and TWICS.

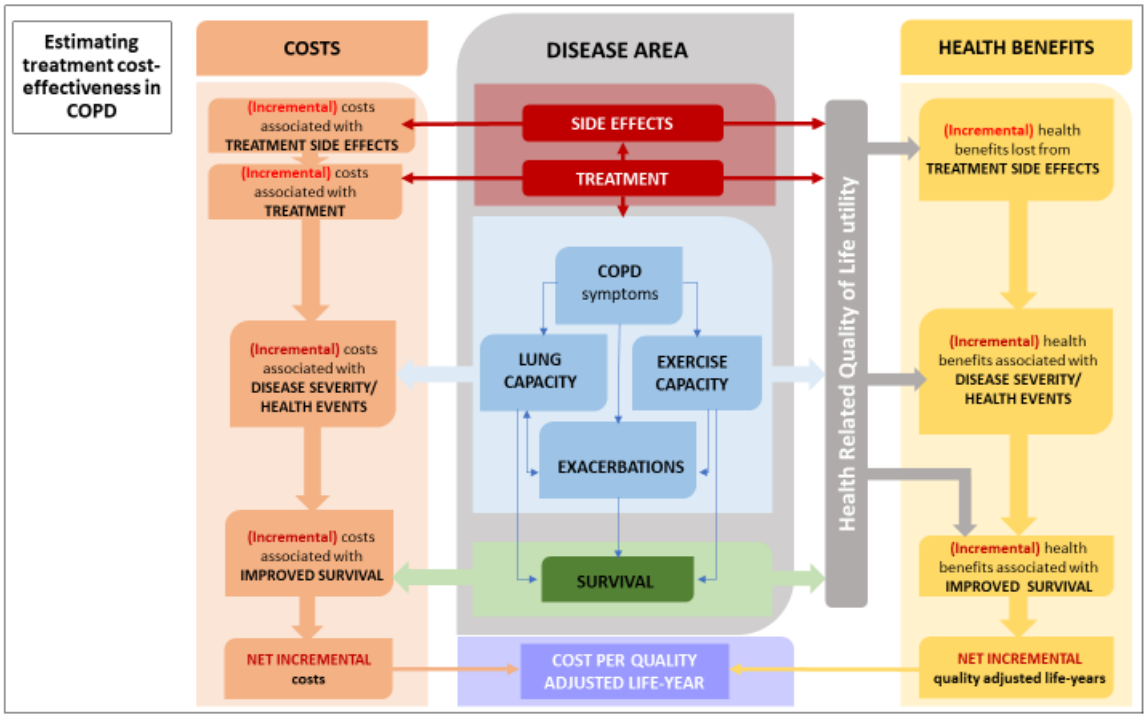


Figure 38: COPD conceptual model building - cost-effectiveness

The next step of applying TWICS to the COPD generic treatment conceptual model is depicted in Figure 39. This diagram is a copy of the COPD cost-effectiveness conceptual model with components 'crossed out' that are not relevant to the treatment effect of theophylline. The assumptions made in developing this step were not made with clinical input and are an illustrative example. First, treatment side effects were not expected to feature in TWICS as low dose theophylline has not shown the same side effects of high doses of theophylline. Removing treatment side effects from the conceptual model central disease area also led to removal of potential costs associated with treatment side effects and losses of health benefits associated with treatment side effects. Disease severity was not measured in TWICS, nor included in the expectation of effects of theophylline on exacerbations, therefore 'disease severity' was removed from both costs and health benefits. The TWICS clinical trial did not measure survival as an outcome and theophylline was not expected to directly affect survival, furthermore, survival was not expected to be a factor bearing in mind the one-year follow-up period, therefore survival was also removed from the COPD generic treatment conceptual model. Removing survival from the conceptual model central disease area also led to removal of potential costs associated with improved survival and potential health benefits associated with improved survival. Finally, COPD symptoms and lung and exercise capacity were removed from the conceptual model. These 'Disease Area' components were removed because the expected trial mechanism was that theophylline changes biological processes in the participant allowing inhaled corticosteroids to reduce numbers of exacerbations, it was not anticipated that theophylline would directly affect COPD symptoms, and lung and exercise capacity.

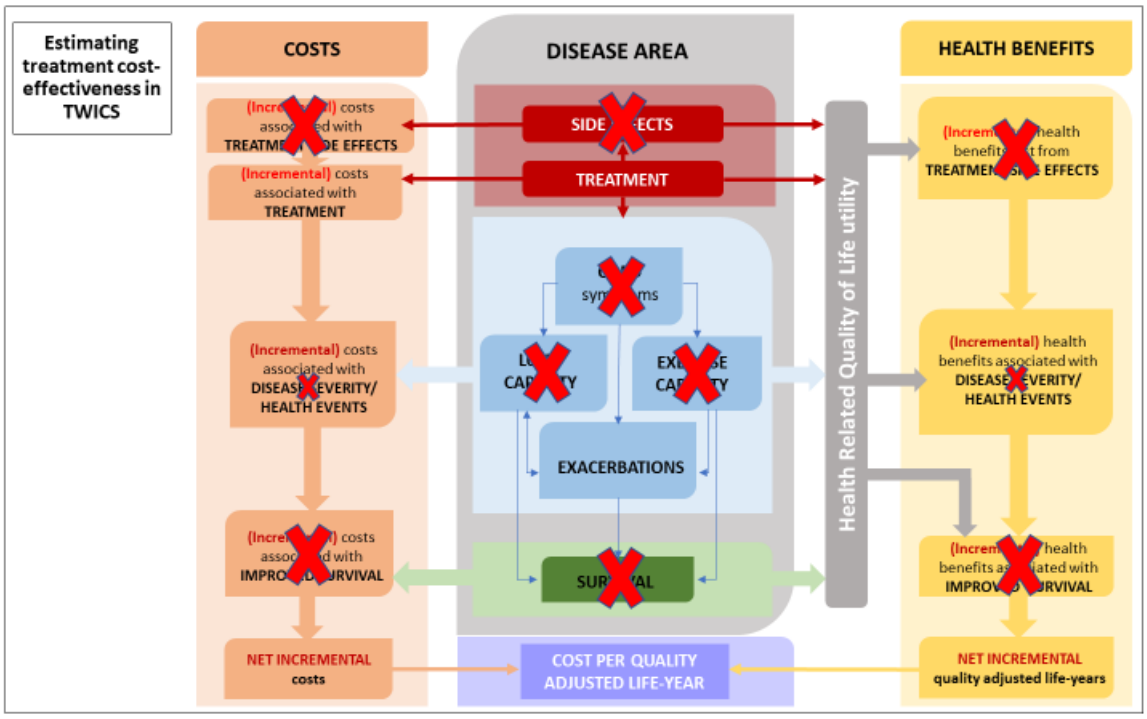


Figure 39: COPD conceptual model building - applying theophylline

Once the unnecessary components were removed from the COPD conceptual model the key components left were those expected to be important in the TWICS economic evaluation (Figure 40). This conceptual model was a simplified illustration of the key components of the TWICS economic evaluation, the last step was to remove the conventional cost-effectiveness components of net incremental costs, net incremental quality adjusted life-years and cost per quality adjusted life-years necessary to conduct an economic evaluation, leaving the trial mechanism and key cost and health benefit components needed to conclude the development of the conceptual model and enable its application in a new role to conduct a conceptual model driven analysis. A final conceptual model was developed from this diagram and is presented in sub-section 6.2.2.

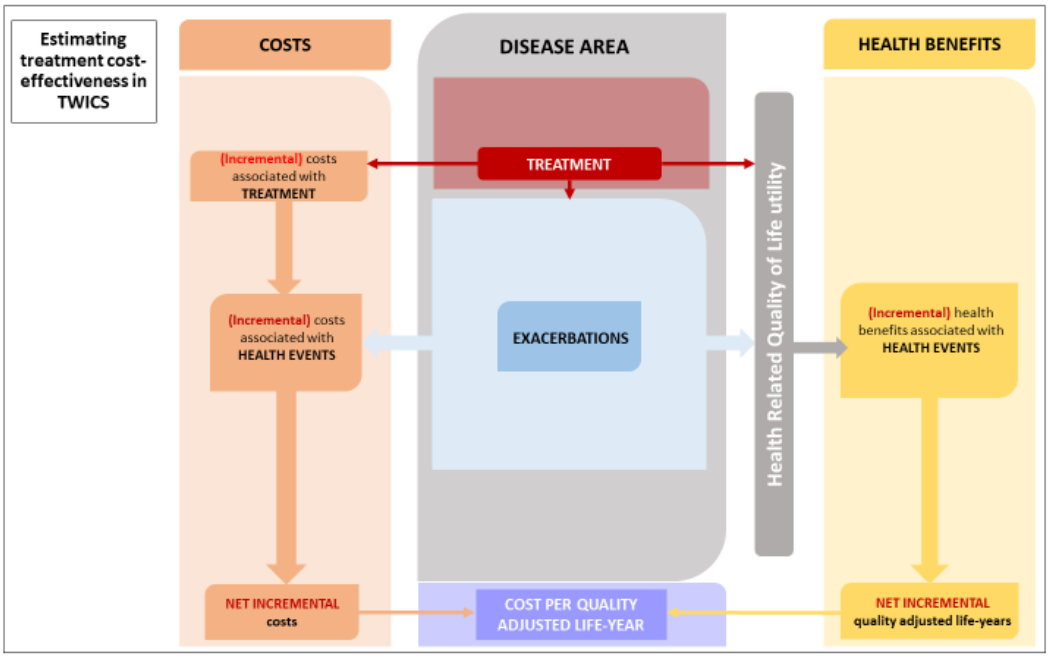


Figure 40: TWICS conceptual model building - economic evaluation

Stage 6 Diagram and document

The final TWICS conceptual model diagram and document are presented in sub-section 6.2.2, the process of applying the methodological framework to the TWICS case study has been described above and further iterations of the conceptual model are included in Appendix 11: Iterative development of TWICS case study conceptual model diagram (*Chapter 6)*.

Stage 7 Review, revise and refine

The earliest iterations of the conceptual model presented in Appendix 11: Iterative development of TWICS case study conceptual model diagram (*Chapter 6)* were developed with input from the TWICS Chief Investigator and Senior Health Economist, Figure A11.94 was presented to TWICS investigators to get feedback on its accuracy, understandability and usefulness. This feedback was positive, with comments that it was useful and understandable. Whilst presenting the conceptual model it was apparent that removing the lines from patient to data (quality of life and health resource use) and 'treatment' to 'health resource use' did not correctly represent the links between patients and data, it made the conceptual model too simplified, these lines were added back in. The description of 'Data' in the third column was changed back to 'Measures' as this is a better description of the components in this column (Figure A11.93).

6.2.2 Final TWICS conceptual model

This sub-section presents the final TWICS conceptual model and accompanying document.

Diagram

The final TWICS conceptual model diagram is a simplified representation of the TWICS trial mechanism and how it links the key components of the cost and health benefits (Figure 41).

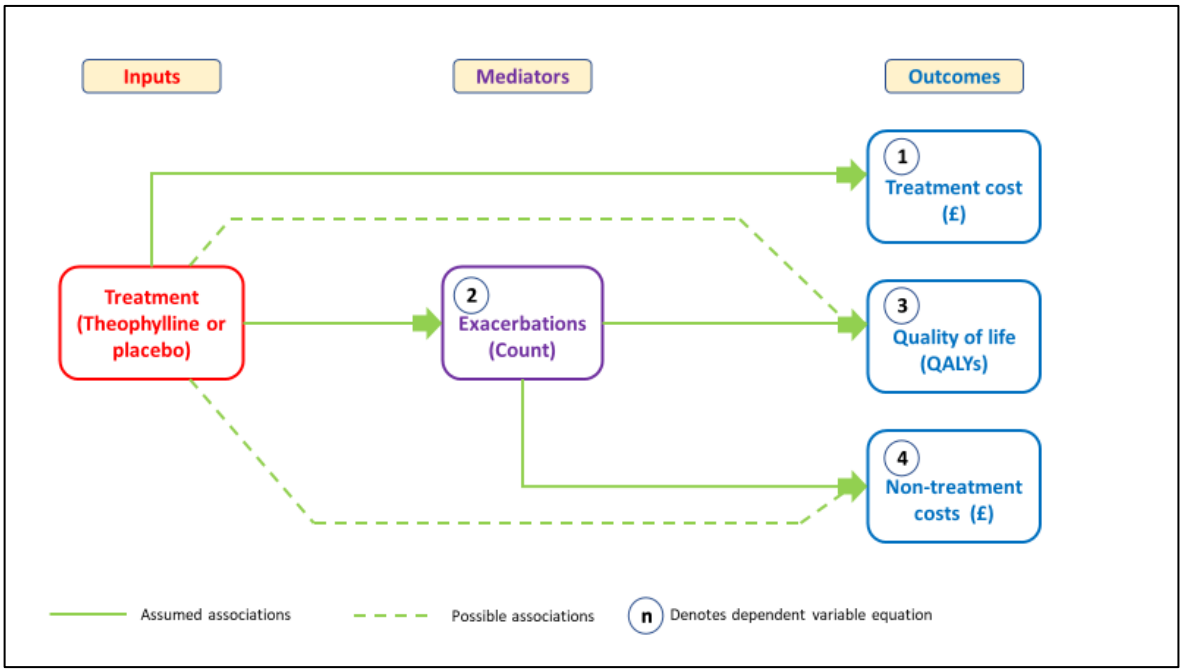


Figure 41: Final TWICS conceptual model

Document

Background

Chronic obstructive pulmonary disease (COPD) is a progressive lung disease associated with breathlessness on exertion, lung function, disability, absence from work, morbidity, early retirement, and premature death. The main cause of COPD is smoking and most cases are diagnosed from the age of 50 onwards. Exacerbations are a key feature of COPD; symptoms include breathlessness, coughing, and expelling mucus. Annually around 15% of COPD patients

experience an exacerbation needing hospitalisation, and 12% of those COPD patients who are hospitalised will die within a year of hospitalisation (116). Exacerbations are linked with increased decline in lung function, reduced physical activity, lower quality of life and increased mortality (115). Mortality can also be predicted using the BODE index which includes **B**ody-mass index, airflow **O**bstruction, **D**yspnea (difficult or laboured breathing) and **E**xercise capacity (188). Patients are classified for treatment using a combination of symptom severity and exacerbation based on the Global initiative for chronic Obstructive Lung Disease (GOLD) assessment tool (189).

It is predicted that in 2030 COPD will be the third leading cause of death globally, compared to the fifth leading cause of death in 2002 (112). The burden in the UK is high; in 2012 there were almost 1 million diagnosed cases of COPD, and COPD accounted for 5-6% of all deaths, financially COPD costs the NHS about £1 billion per year (111) and exacerbations account for 60% of the direct costs of COPD to the NHS (190).

The effects of the exacerbations are: increased mortality, decreased lung function, decrease in quality of life, an increased risk of hospitalisation, accelerated lung function decline and an increase in symptoms related to COPD.

Current COPD treatment includes inhaled corticosteroids, combined with inhaled long-acting β_2 agonists, however patients still have exacerbations despite treatment (111). Oral theophylline has also been used to reduce exacerbations, but the high doses involved cause unpleasant side effects which has led to treatment with other, better tolerated bronchodilators. However, recent preclinical trials indicate that a low dose of oral theophylline is beneficial in treating COPD.

A novel rich picture depicting the factors predicting a person's severity of COPD and risk of death is presented in Figure 42.

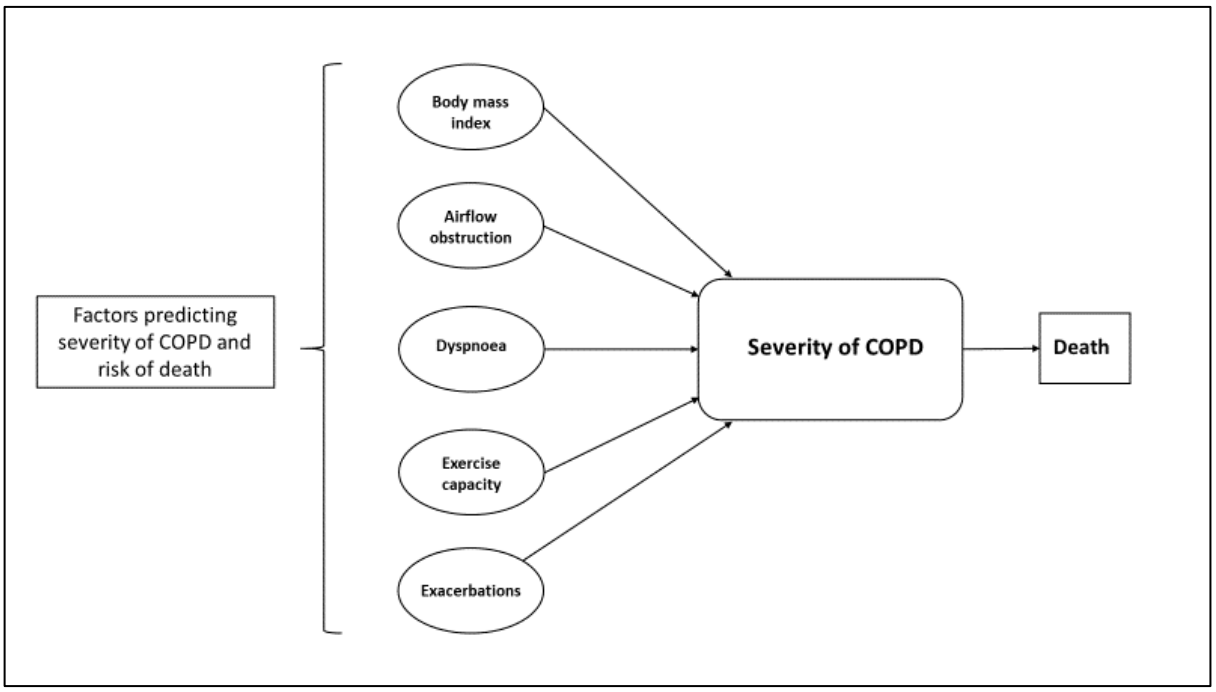


Figure 42: COPD rich picture

The target population were adults aged 40 or above with a diagnosis of COPD and who are likely to exacerbate during the 12 months follow-up of the trial, as evidenced by two or more exacerbations in the year prior to recruitment into the trial. Other inclusion criteria include; smoking history of at least 10 pack years and no exacerbations in the 4 weeks prior to recruitment. Full details on inclusion and exclusion criteria are published elsewhere (111).

The intervention was usual care plus 200mg or 400mg (dose was dependant on ideal body weight and smoking status) theophylline for 52-weeks. The comparator was usual care plus placebo.

The primary outcome of TWICS was the number of exacerbations in the 52-week follow-up period of the clinical trial requiring treatment with antibiotics or oral corticosteroids. The health economic outcome was the quality adjusted life-year, measured using the EQ-5D-3L questionnaire (74). The study type was a two-arm clinical trial.

The perspective taken in the economic evaluation was NHS direct healthcare costs and personal social services.

The research question was 'what are the trial mechanism and key economic components relevant to the TWICS economic evaluation?'

The conceptual modelling team comprised three people: a junior and a senior health economist and a medical consultant with expertise on COPD.

Objectives

The modelling objective was to simplify and represent the key components driving the economic evaluation of TWICS. The conceptual model would be used as a communication tool, to test the logic of the associations represented in it, and to conduct an analysis based on these associations.

Scope

The scope was restricted to the key economic components of TWICS: treatment arm, treatment cost, exacerbations (as theophylline was expected to decrease the number of exacerbations experienced), non-treatment cost and quality of life. The treatment and non-treatment costs were presented separately as the original economic evaluation showed that treatment costs were precise, whereas non-treatment costs were uncertain.

Detail

The detail of the model was at aggregate level for treatment, exacerbations, and quality adjusted life-years. Costs were disaggregated into treatment and non-treatment costs.

Model contents

The conceptual model is split into three sections: inputs (red); mediators (purple), and outcomes (blue). Components in each section are represented by rectangles in the same colour. Expected associations between components are represented by arrows; solid lines represent assumed associations, and dashed

lines represent possible associations. Components with lines leaving from the right are explanatory variables; the component to which their arrow enters is predicted by them, for example treatment arm predicts treatment costs and exacerbations. All components in the 'inputs' section are explanatory. Components with arrow heads entering on the left are dependant variables; they are predicted by the components from which the arrow originates. For example, treatment costs are predicted by the treatment arm. All components in the 'outcomes' section are dependent variables. A component can be both explanatory and dependent, mediating components are both explanatory and dependant; they have arrows entering and leaving. For example, exacerbations are predicted by the treatment arm and also predict quality of life and non-treatment costs.

Assumed associations (solid lines) - Participants are randomised to one of two treatment arms: either theophylline or placebo. The hypothesis of the clinical trial was that there would be less exacerbations needing treatment in the theophylline arm (association 2) due to chemical factors specific to theophylline. Exacerbations are expected to predict quality of life; participants experiencing lower numbers of exacerbations would report higher levels of quality of life (association 3), conversely participants experiencing higher numbers of exacerbations were expected to report lower levels of quality of life. Treatment costs are expected to be higher in the theophylline arm (association 1) due to zero costs for the placebo treatment. The number of exacerbations experienced are expected to predict non-treatment costs (association 4), most likely in the costs of treating exacerbations but also, if the severity of a participant's COPD progresses, for non-exacerbation related healthcare costs.

Possible associations (dashed line) – It is possible that the treatment arm the participant is randomised to can directly influence quality of life not mediated through exacerbations (association 3). There is a possibility that non-treatment costs (association 4) may be directly affected by the allocated treatment arm. These possible associations are equivalent to the outcomes-driven 'black box' evaluation, where there is no consideration of the process driving the results.

Assumptions and simplifications

The overarching simplification was to base the conceptual model on the disaggregated cost-effectiveness template using disaggregated key economic evaluation components, and that there would not be any key components identified outside of this template.

The disease area initially included four key components: symptoms, lung and exercise capacity, and exacerbations (Figure 34), this was further reduced to only include exacerbations as the assumption was that theophylline affects a patient's biological mechanism, which in turn decreases exacerbations. There was no expectation that theophylline would directly affect symptoms, lung and exercise capacity.

It was assumed that there would be no treatment side effects as preclinical trials did not witness any side effects of using low dose theophylline. It was also assumed that survival would not be an aspect of the trial mechanism, due to the short (12-months) follow-up period, also, survival was not an outcome of the clinical trial.

The key components of costs were assumed to split into treatment and non-treatment, it was not expected that further sub-division of costs would be useful.

Finally, it was assumed that all clinical events and factors of COPD would be picked up by the EQ-5D-3L questionnaire to give a measure of health benefits. Whilst this may be a strong assumption based on the generic nature of the EQ-5D questionnaire, research has found that the EQ-5D measure is a reliable and responsive measure of health-related quality of life in people with COPD (191, 192). Furthermore, this is an illustrative example to show how a chosen economic outcome measure from the trial data would feed into the conceptual model, not a formal conceptual model driven analysis.

Chapter 6

6.3 Case study: application of the methodological framework - BeatIt case study

This section of the chapter describes applying the methodological framework to the BeatIt case study which was presented earlier in Chapter 3 Section 3.3. In sub-section 6.3.1 the conceptual model development process is described and the final iterations are presented, the initial iterations of the conceptual model are presented in Appendix 12: Iterative development of BeatIt case study conceptual model diagram (*Chapter 6)*, and the output of diagram and documentation is presented in sub-section 6.3.2.

6.3.1 Application of methodological framework

This sub-section describes the development process, presenting each stage separately.

Stage 1: Understand the problem

The recommendation in this stage is to describe the relevant components of the decision problem, these include understanding the context of the research system and the problem. Other aspects of this stage include developing a rich picture, determining the project team and defining the research question.

To gain an understanding of the research problem the questions in Table 14 were considered, and answers were included in the conceptual modelling document. Information and evidence was taken from the BeatIt protocol (128) and published literature.

Developing a rich picture

The next step was to develop a rich picture of the disease area, an informal diagram of the problem area. To do this I took my understanding of depression from the questions above, and from existing diagrams depicting the nature of depression.

Depression is characterised by an initial episode of depression during which symptoms worsen in severity, followed by complete remission and recovery, or periods of relapse and recurrence, this is illustrated well in the following figures (Figure 43 and Figure 44).

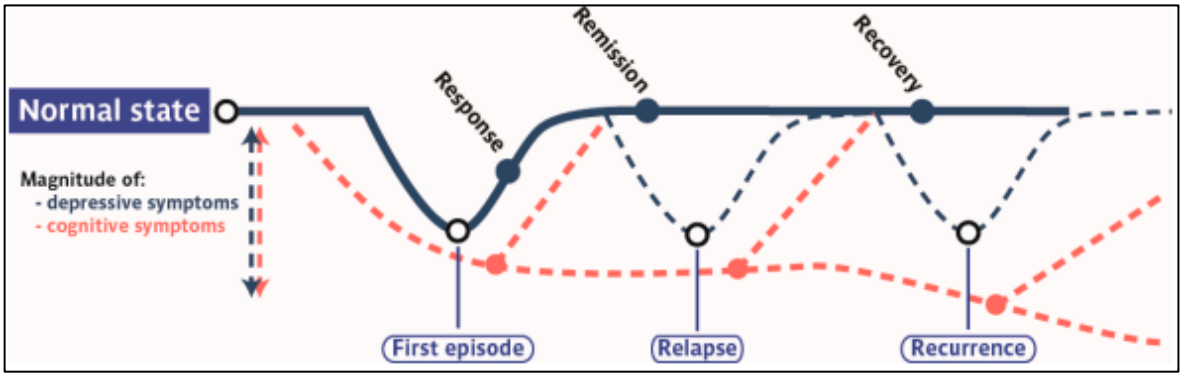


Figure 43: Recurrent depression from Darcet et al. 2016 (193)

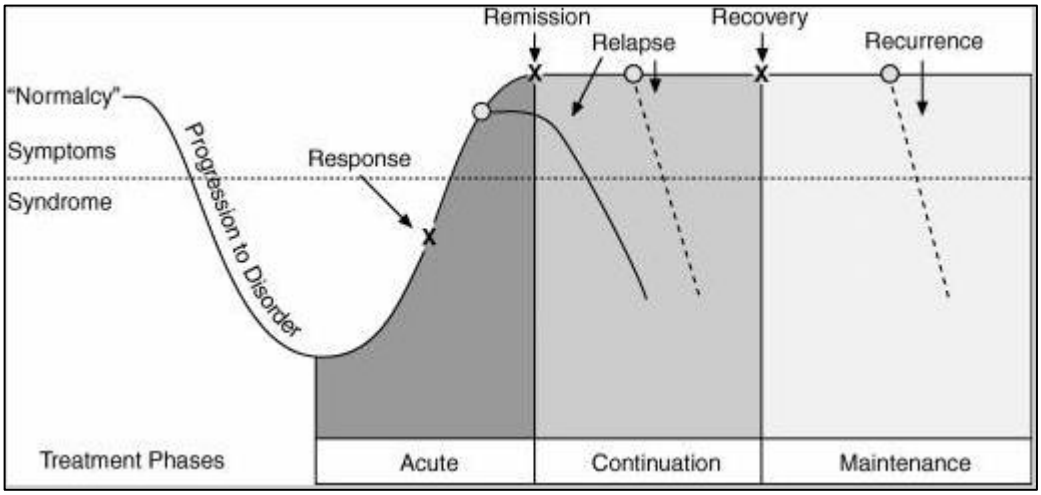


Figure 44: Response, remission, recovery, relapse, and recurrence of depression during treatment stages from "Preventing recurrent depression: long-term treatment for major depressive disorder." (194)

The rich picture is presented in the conceptual model document in 6.3.2.

Finally, in this stage the research question for the conceptual model was defined and the team determined.

Stage 2: Setting objectives

The recommendation in this stage is to describe the objective of the conceptual model, this was linked to the research question.

Stage 3 Determine scope of the conceptual model

The recommendation in this stage is to describe the scope of the conceptual model. In this stage I narrowed down the elements included in the rich picture relevant to the objectives and research question. The trial mechanism did not include severity of depression so that was omitted. Activity was not included in the rich picture but was expected to be an effect of the BeatIt therapy so was included in the conceptual model diagram.

Stage 4 Determine the detail of the conceptual model

The recommendation in this stage is to determine the level of detail of the conceptual model.

Stage 5 Determine the content of the conceptual model

The recommendation of this stage is to describe the trial mechanism and how the key components of the mechanism interact.

The format and iterations of the BeatIt conceptual models follow the format and iterations of the final TWICS conceptual models, using the disaggregated cost-effectiveness template (Figure 3) then following the steps already presented in 6.2.1 for the TWICS conceptual model.

The first step was to develop a conceptual model of the disease area, this was a simplification of the disease. For depression to be diagnosed in a patient a number of symptoms have to be experienced over a specific period of time, symptoms vary between people and are complex. To simplify the disease area symptoms were represented as social, physical and psychological (Figure 45).

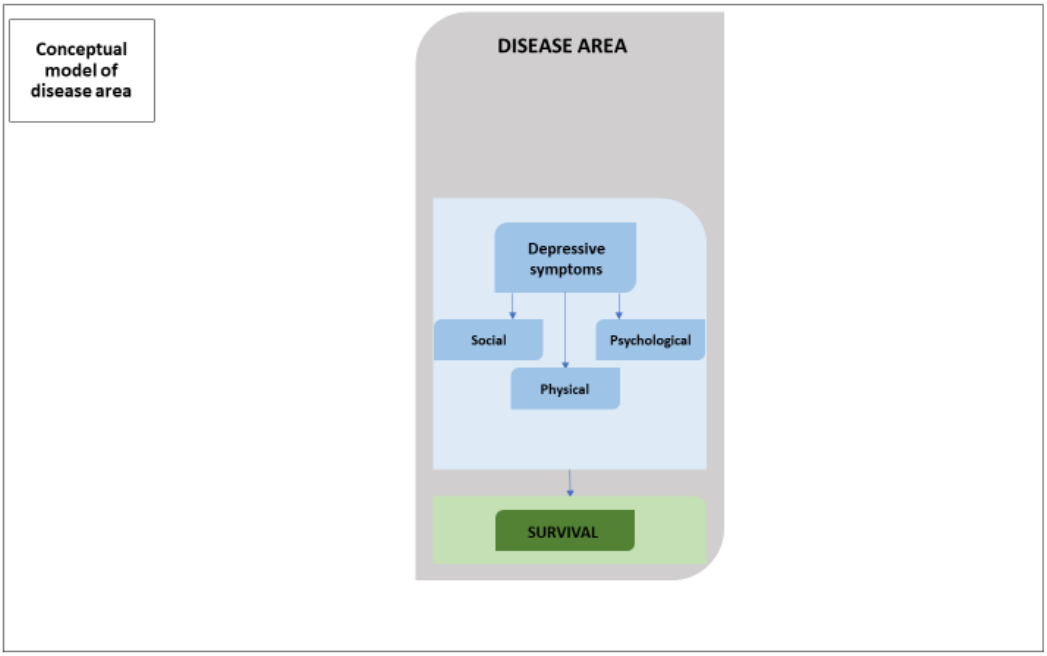


Figure 45: BeatIt conceptual model building - disease area

The next step was to include healthcare costs, two categories of healthcare costs are added; costs associated directly to depression that result from disease severity and related health events, and costs associated with improved survival (Figure 46).

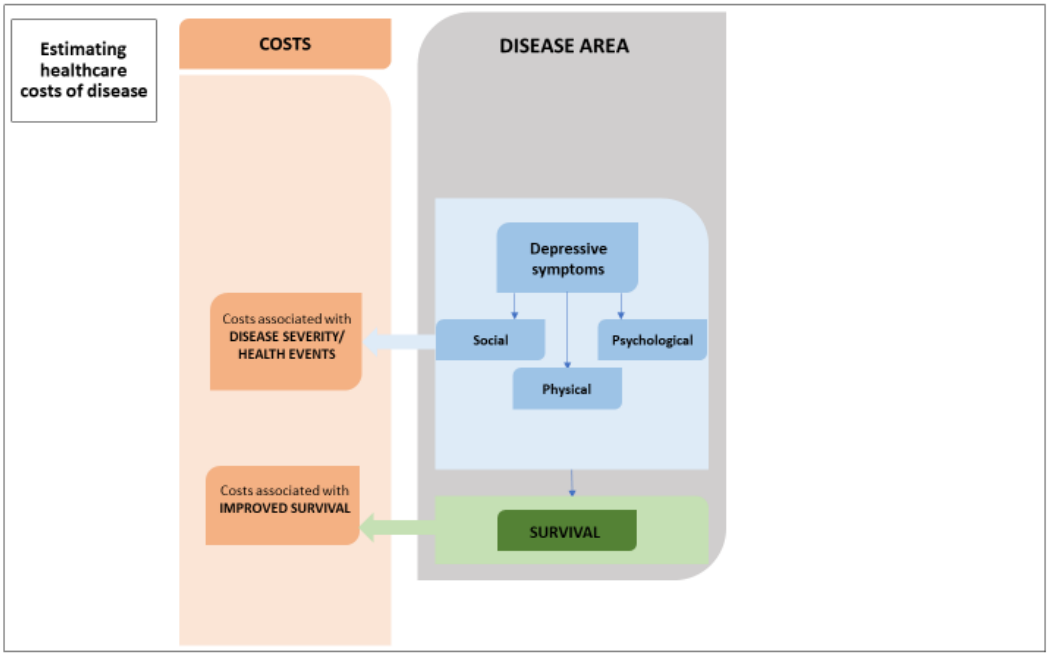


Figure 46: BeatIt conceptual model building - healthcare costs

Then health benefits in terms of health-related quality of life relating to depression were added (Figure 47). The grey health-related quality of life utility panel is the step in calculating QALYs where clinical characteristics are measured using utilities. The yellow 'Health benefits' panel represents utilities translated into QALYs.

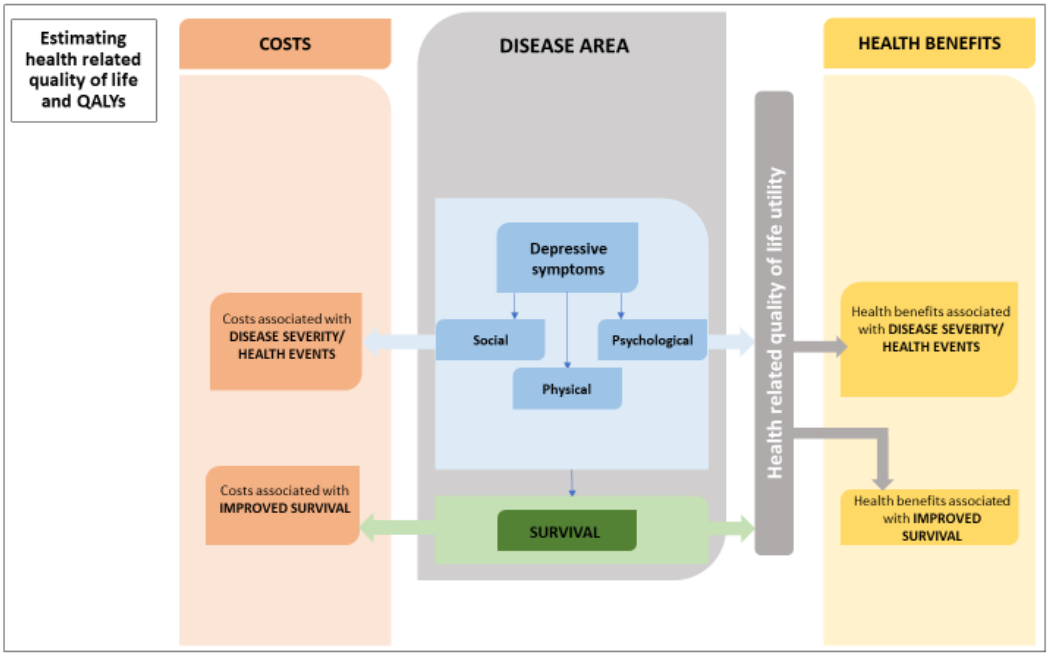


Figure 47: BeatIt conceptual model building - health related quality of life and quality adjusted life-years

The next step is to layer in a generic treatment to include the treatment effects consistent with the disaggregated cost-effectiveness template (Figure 48). Layering in the generic treatment (and comparator) illustrates the incremental analysis of costs and health benefits; all cost and health benefit components are now incremental. In the 'Disease Area' panel a new red section is added with two components: the generic treatment itself and side effects resulting from the treatment. These components are associated with costs and health benefits: potential costs are those associated with treatment side effects and those directly linked to treatment, and health benefits are the quality of life lost as a result of treatment side effects.

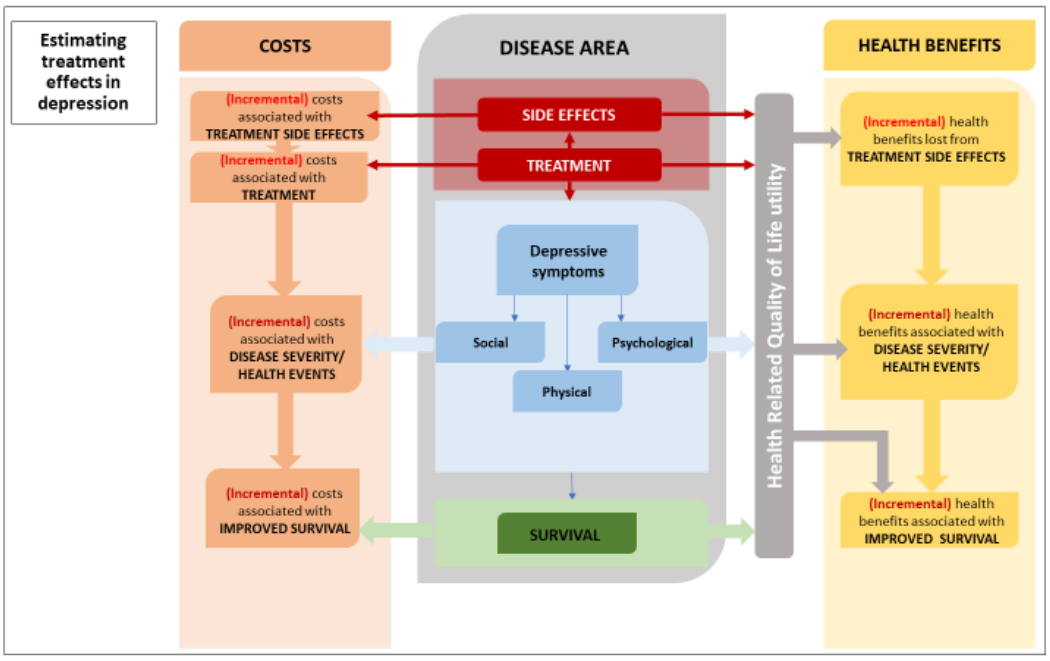


Figure 48: BeatIt conceptual model building - generic treatment

Once generic treatment effects are included in the conceptual model the cost-effectiveness can be illustrated (Figure 49). This is depicted in the bottom row of the diagram where all cost components are summarised in 'Net incremental costs' and all health benefits are summarised in 'Net incremental quality adjusted life-years'. These summary components feed into the economic evaluation summary measure of cost per quality adjusted life-year (ICER). This diagram in the conceptual model building process depicts a full economic evaluation. The next step is to adapt the conceptual model to BeatIt, replacing the generic treatment with BeatIt and taking out unnecessary elements that are not expected to effect or be relevant to BeatIt.

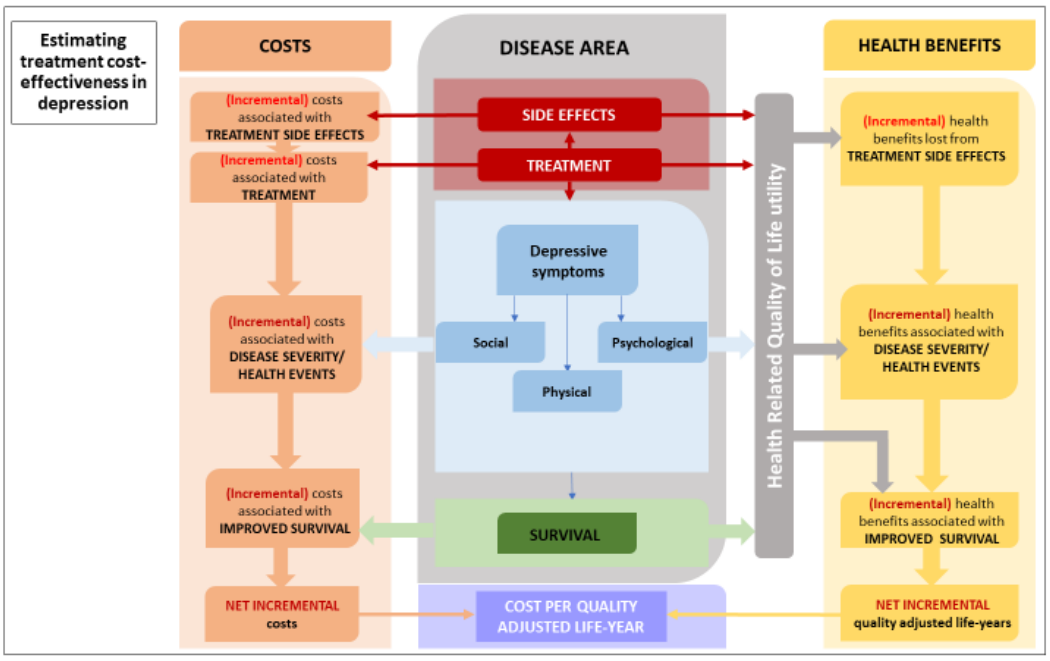


Figure 49: BeatIt conceptual model building - cost-effectiveness

The next step of applying the BeatIt therapy to the depression generic treatment conceptual model is depicted in two diagrams (Figure 50 and Figure 51). The assumptions made in developing this step were not made with clinical input and are an illustrative example. The first diagram is a copy of the depression cost-effectiveness conceptual model with components 'crossed out' that are not relevant to the treatment effect of the BeatIt therapy. First, treatment side effects were not expected to feature in BeatIt. Removing treatment side effects from the conceptual model central disease area removes potential costs associated with treatment side effects and losses of health benefits associated with treatment side effects. The BeatIt clinical trial did not expect survival to be affected by the therapy, only depressive symptoms, therefore survival was also removed from the depression generic treatment conceptual model. Removing survival from the conceptual model central disease area also led to removal of potential costs associated with improved survival and potential health benefits associated with improved survival. The severity of depression is not pertinent to the BeatIt clinical trial so the 'disease severity' aspects of costs and health benefits were removed. Finally, the separate social, physical and psychological components of depressive symptoms were removed as BeatIt measured depression as a whole as a primary outcome.

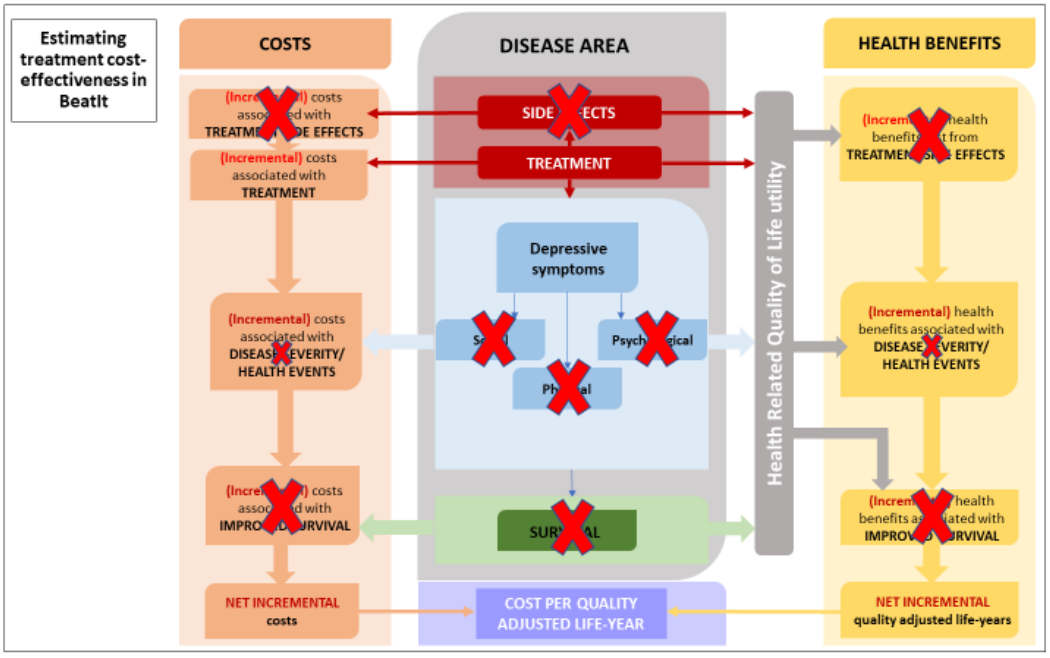


Figure 50: BeatIt conceptual model building - applying BeatIt

Once the unnecessary components were removed from the BeatIt conceptual model the key components left were those expected to be important in the economic evaluation (Figure 51). This conceptual model was a simplified illustration of the key components of the BeatIt economic evaluation, the last step was to remove the conventional cost-effectiveness components of net incremental costs, net incremental QALYs and cost per quality adjusted life-years necessary to conduct an economic evaluation, leaving the trial mechanism and key cost and health benefit components needed to conclude the development of the conceptual model and enable its application in a new role to conduct a conceptual model driven analysis. As the BeatIt therapy is based on behaviour change a component to represent the expected increase in activity was added to the final conceptual model, this was not relevant to add to the conceptual model as it developed as it did not relate to treatment or disease area. This final conceptual model is presented in sub-section 6.3.2.

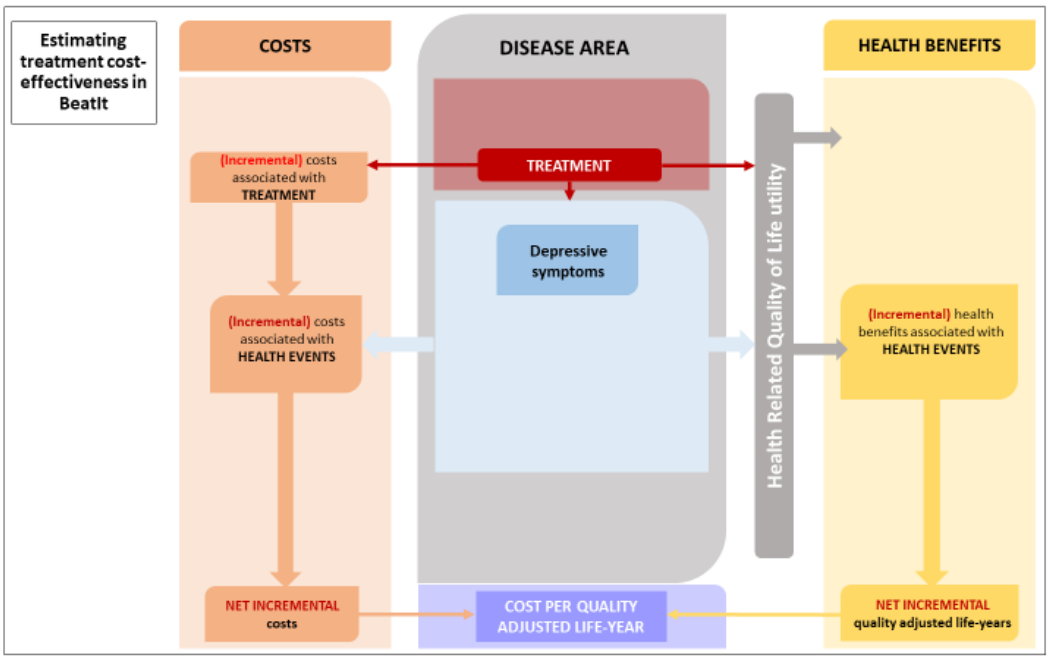


Figure 51: BeatIt conceptual model building - economic evaluation

Stage 6 Diagram and document

The final BeatIt conceptual model diagram and document are presented in sub-section 06.3.2, the process of applying the methodological framework to the BeatIt case study has been described above and further iterations of the

conceptual model are included in Appendix 12: Iterative development of BeatIt case study conceptual model diagram (*Chapter 6)*.

Stage 7 Review, revise and refine

The earliest iterations of the conceptual model presented in Appendix 12: Iterative development of BeatIt case study conceptual model diagram (*Chapter 6)* were developed with input from the BeatIt Chief Investigator and Senior Health Economist. However, due to time constraints the final conceptual model was not seen by the Chief Investigator.

6.3.2 Final BeatIt conceptual model

This sub-section presents the final BeatIt conceptual model and accompanying document.

Diagram

The final BeatIt conceptual model diagram is a simplified representation of the BeatIt trial mechanism and how it links the key components of the cost and health benefits (Figure 52).

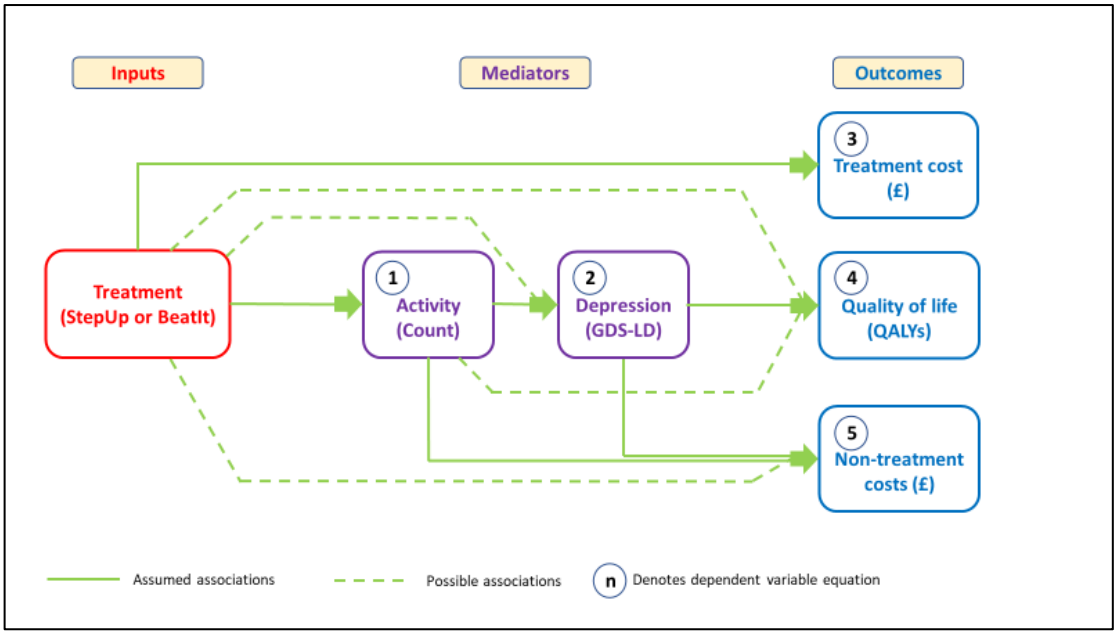


Figure 52: Final BeatIt conceptual model

Chapter 6
Document

Background

Depression is a common mental health issue, an episode may be mild, moderate or severe. During a mild episode the person will not cease to function but may struggle to work and take part in normal social activities, during a severe episode most activities will be limited (195). Symptoms includes low mood most of the day nearly every day, taking less interest or pleasure in usual activities, insomnia, fatigue, worthlessness, and lack of concentration (196). Depression is the leading cause of suicide, and suicide and suicide attempts are 10 times as high in people with psychiatric diseases compared to the general population (197).

Depression is highly prevalent in adults with intellectual disabilities, with a point prevalence of 5% (132). Evidence suggests that depression is five times more common in adults with intellectual disabilities compared with the general population (133).

Depression affects around 300 million people globally (129). The burden of depression includes costs of prescription and treatments, and time off work. In 2007 1.24 million people in England were estimated to have depression, with healthcare costs of £1.7 billion and £5.8 billion in lost earnings (130). In 2017/18 over 900,000 patients were prescribed at least on antidepressant drug in Scotland, and the cost of antidepressants was £44.8 million (131).

Psychological therapies for treating depression in the general population are well established. However, approaches like cognitive behavioural therapy (CBT) are essentially talking therapies and their reliance on verbal communication makes them less accessible for many adults with intellectual disabilities (126). While research concerning the use of CBT for adults with intellectual disabilities and depression has been encouraging (134), behavioural activation therapy which is less reliant on verbal and cognitive skills and maybe more suited to adults with intellectual disabilities.

Behavioural activation focuses on the link between mood and activity. It aims to increase participation in purposeful and motivating activities, bringing the individual into contact with positive experiences and helping to lift their mood (135). Research in the general population with severe depression has shown that behavioural activation is as effective as antidepressant medication and more effective than CBT (136, 137), with the positive treatment effects shown to last as long as those for CBT (138). Behavioural activation is also recommended in the NICE (2009) guidelines for the treatment of depression in the general population (139).

The target population were adults with mild to moderate intellectual disabilities, a diagnosis of depression, and a carer willing to accompany the participant to therapy visits and follow-up visits. Full inclusion and exclusion criteria are published elsewhere (128).

The intervention was a behavioural activation therapy (BeatIt) delivered over 12 sessions. The comparator was a guided self-help therapy (StepUp) delivered over eight weeks. Both treatments were delivered one-to-one with the participant's supporter present.

The primary outcome of BeatIt was depressive symptoms measured by the Glasgow Depression Scale for Learning Disabilities (GDS-LD) at the 12-month follow-up period (141). The economic outcome was the quality adjusted life-year measured by the EQ-5D-Y questionnaire, this questionnaire is specifically designed aimed at children and adolescents using simplified language, making it suitable for adults with intellectual disabilities (77).

The perspective taken was NHS and personal social services, personal costs and time off work were not included.

The rich picture shows the key symptoms of depression, the potentially repetitive circular nature of the illness, plus an indication that there are different severities of depression (Figure 53).

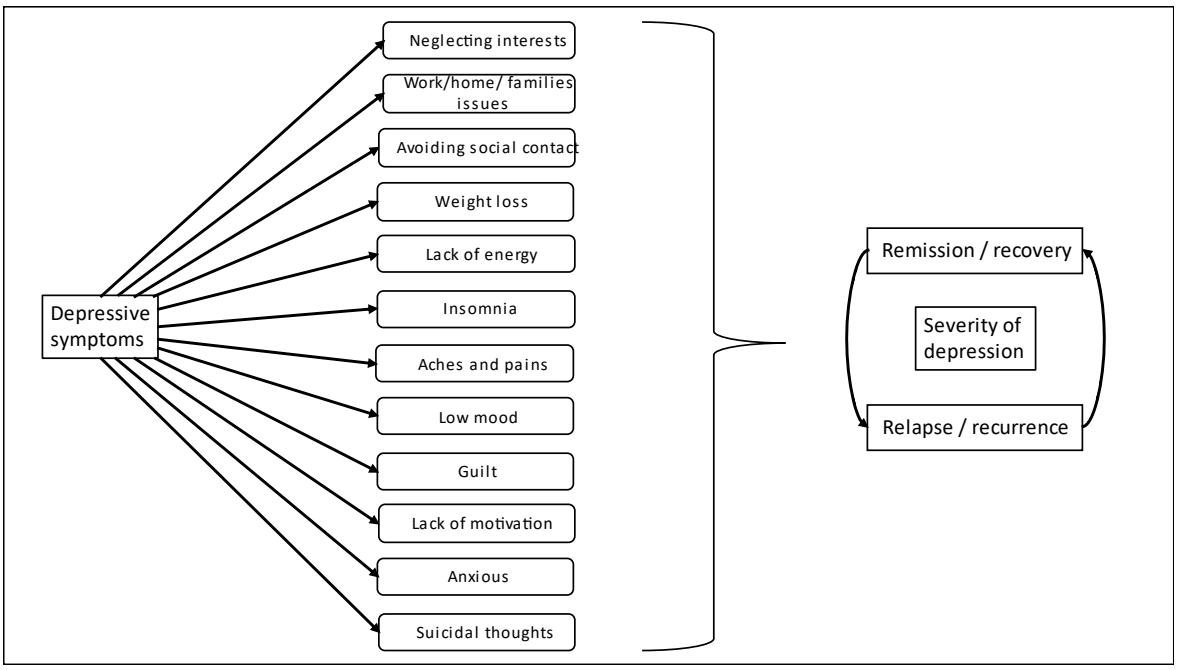


Figure 53: Depression rich picture

The research question for the conceptual model driven analysis was: 'what are the key components in the BeatIt economic evaluation and how are they related the trial mechanism?'

The conceptual modelling team comprised three people: me, a senior health economist who was also PI on the BeatIt study and the Chief Investigator on the BeatIt study.

Objectives

The objective was to develop a conceptual model to depict the trial mechanism of BeatIt, linking the key components of the economic evaluation in terms of inputs and outputs; this could be used as a communication tool, to test the logic of the conceptual model and to conduct further analysis based on the conceptual model.

Scope

The primary outcome of BeatIt was depressive symptoms so these were included along with the key economic components of treatment arm, treatment cost, non-treatment cost and quality of life. Activity was also included as the expectation of the behavioural activation was that it would increase participation in activities, which in turn would decrease depressive symptoms. The treatment and non-treatment costs were kept separate as we knew from the original economic evaluation that treatment costs were precise, whereas non-treatment costs were uncertain.

Detail

Treatment allocation, the quality of life measure taken from EQ-5D-Y questionnaire, depressive symptoms taken from the GDS-LD and levels of activity were all at an aggregate level. Costs were disaggregated into costs of treatment and non-treatment costs.

Model contents

The BeatIt conceptual model is split into three sections; inputs (red), mediators (purple) and outcomes (blue); components in each of these sections are represented by rectangles in the same colour. Expected associations between components are represented by arrows; solid lines represent assumed associations, and dashed lines represent possible associations. Components with lines leaving from the right are explanatory variables; the component to which their arrow enters is predicted by them, for example treatment arm predicts treatment costs and exacerbations. All components in the 'input' section are explanatory. Components with arrow heads entering on the left are dependant variables; they are predicted by the components from which the arrow originates, for example treatment costs are predicted by the treatment arm. All components in the 'outcome' section are dependent variable. A component can be both explanatory and dependent, mediating components are both explanatory and dependant; they have arrows entering and leaving, for example exacerbations are predicted by the treatment arm and also predict quality of life and non-treatment costs.

Once the expected links and associations are defined and the dynamics of the mechanism represented, each dependant variable is allocated a number, this makes describing the conceptual model clearer and easier for the user to follow, as well as identifying regressions for the next stage of the new approach.

Assumed associations (solid lines) - Participants were randomised to one of two treatment arms: BeatIt or StepUp. In the clinical trial the hypothesis was that the BeatIt therapy increased the level of activities in the participants in the BeatIt arm (association 1). Levels of activity would affect symptoms of depression (association 2) with participants reporting higher levels of activity expected to report fewer depressive symptoms (association 2), conversely participants reporting lower levels of activity were expected to also report more depressive symptoms. The level of depressive symptoms reported were expected to directly impact quality of life (association 4), with participants reporting fewer depressive symptoms expected to also report better quality of life and vice versa. Treatment costs were expected to be higher in the BeatIt arm (association 3) due to more sessions included in the BeatIt therapy compared to the StepUp therapy. Levels of activity were expected to affect the non-treatment costs (association 5), with higher reported levels of activity increasing non-treatment costs, and healthcare use relating to depression was expected to be associated with non-treatment costs (association 5).

Possible associations (dashed lines) – It is possible that the treatment arm that participants are randomised to will directly affect the levels of depressive symptoms, for example through the therapy itself improving depressive symptoms, not mediated through activity (association 2). There is also the possibility that treatment arm will affect quality of life directly, (association 4), and that treatment arm will also affect non-treatment costs directly (association 5), neither of these possible associations would be mediated through activity and depression, similar to a 'black box' evaluation. There is a possibility that levels of activity are directly associated with quality of life without being mediated through depression (association 4), for example through the 'usual activities' or 'mobility' domains of the EQ-5D-Y questionnaire.

Whilst depression may be more enduring in people with intellectual disabilities it was assumed that the characteristics of the disease would be the same in this population as in the general population when developing the rich picture and disease area in the conceptual model.

Symptoms are amalgamated as it was not possible to extract specific symptoms which predicted others.

It was assumed that the link between the BeatIt therapy and depression is mediated by activity only.

It was assumed that no treatment side effects would be experienced so costs and benefits relating to side effects were removed. Severity of depression was not included, this was not an inclusion factor or measured as an outcome; the disease severity was removed from costs and health benefits. Survival was not an outcome in the BeatIt clinical trial, or expected as an effect of the BeatIt therapy; it was removed from costs and health benefits.

Finally, it was assumed that quality of life measured by the EQ-5D-Y would pick up the effects of depression measured in the GDS-LD.

6.4 Limitations

These conceptual models were illustrative in nature, and whilst this highlighted the key attributes of development, it was potentially at the cost of omitting some details. Due to time pressures, it was not possible to confirm the appropriateness of the conceptual models with the clinical trial team. However, early iterations of the conceptual models did have input from the trial teams, these are included in 0Appendix 11: Iterative development of TWICS case study conceptual model diagram (Chapter 6) and Appendix 12: Iterative development of BeatIt case study conceptual model diagram (Chapter 6). Although there was input into their initial development, there was not any expert input for their refinement. Furthermore, it is realistic to assume that, had the conceptual

model been developed in a pilot or feasibility study compared to later in a definitive trial, the conceptual model may have included different components.

6.5 Summary

This chapter has demonstrated the application of the methodological framework presented in Chapter 5 to the two thesis case studies. For each case study a step-by-step description of the application the methodological framework to the case study was included, and the final version of the conceptual model presented.

The next paragraphs describe lessons learnt whilst applying the methodological framework. The methodological framework was largely a useful guide to developing the conceptual models. The questions in Table 14 acted as a checklist to ensure that all aspects of understanding the problem were considered and to systematically guide the gathering of information for understanding the problem. The explanations in Stage 5 were very useful, guiding the consideration and identification of components and associations for the content of the conceptual model. This guidance helped give a clear understanding of what should be included in the conceptual model and how the components of the model are related. Identifying the conceptual model key components based on the disaggregated cost-effectiveness template was very useful, helping to focus on what is important in an economic evaluation; first a conceptual model of the disease or behaviour expected in the trial mechanism was developed, this formed the centre panel of the conceptual model, a generic treatment was then layered into this trial mechanism which includes components related to treatment including costs and the expected effect of treatment (side effects and mortality), to replicate the disaggregated cost-effectiveness template; layering in the treatment allowed the conceptual model to depict the incremental costs and effects expected in a conventional economic evaluation. This gave a generic conceptual model for the trial mechanism. Once this was developed the generic treatment was replaced with the specific treatments included in the clinical trial, then the modeller asks ‘what is relevant to the clinical trial?’ as not all components in this generic conceptual model will be relevant to individual trials, any components that are not relevant are

removed. Another issue is 'what is practical?' although Roberts et al. guidance (26) recommends not considering data when developing a conceptual model, it is impractical to include components that cannot be represented by data, so once the generic treatment conceptual model was developed with all key components it was pragmatic to exclude components that are not possible to measure due to lack of data from the clinical trial.

The iterative nature of developing a conceptual model was evident, particularly in TWICS as several iterations were developed. There were also iterations developed in the 'Review, revise and refine' stage and the final three stages were closely linked.

Finally, it would be useful to include the conceptual model diagram and document in the health economics plan for reference, for continuity when there are staff changes, and for communicating and agreeing understanding with the trial team.

The next chapter uses the conceptual models developed in this chapter to demonstrate the new role using illustrative case studies; a conceptual model driven analysis.

Chapter 7 Case studies to illustrate the new role for conceptual models

7.1 Chapter overview

The aim of this chapter is to demonstrate the new role for conceptual models using the two case studies introduced in Chapter 3, this fulfils the first objective of the thesis. In demonstrating the new role, it can be evaluated and any strengths and limitations highlighted. First a brief recap of the background and original results of the case studies from Chapter 3 is given, then the three aspects of the proposed novel approach are applied to each case study, these three aspects are: 1) the conceptual model for each case study is presented visually and described; 2) the regressions capturing the relationships identified in each case study are tested and results reported, 3) bootstrapped samples from these regression results are used to calculate point estimates, present cost-effectiveness planes and cost-effectiveness acceptability curves, and estimate value of information figures, these results are then compared to the original treatment arm-based results.

Both of the case studies are for illustrative purposes only, to demonstrate a potential method for implementing the conceptual model analysis. The focus of the case studies was to introduce the concept of conceptual models in this role, not to rigidly recommend a predefined method. The method used in these illustrative case studies should not be assumed to be the most statistically rigorous suitable for the new role.

The layout of the chapter is as follows: Section 7.2 refreshes the reader's memory of the background to the TWICS case study; Section 7.3 describes the application of the novel approach to the TWICS case study; Section 7.4 reintroduces the BeatIt case study; Section 7.5 describes the application of the novel approach to the BeatIt case study; Section 7.6 presents a discussion about the new role for conceptual models, and Section 7.7 summarises the chapter. Stata 16 (StataCorp. 2019. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC.) was used to operationalise the conceptual model analysis.

(To avoid unnecessary repetition in the BeatIt case study, there is more detail on the reasons for including each aspect of the novel approach presented in the TWICS case study.)

7.2 Case study #1 TWICS

7.2.1 Background of TWICS

The TWICS clinical trial investigated the effectiveness and cost-effectiveness of adding low-dose theophylline as an adjunctive therapy to inhaled corticosteroids in patients with COPD and a history of frequent exacerbations, compared to placebo. The primary outcome was the number of moderate or severe exacerbations requiring treatment over the 12-month follow-up period for each participant. The population was adults with COPD who had a history of at least 2 exacerbations (requiring treatment) in the previous 12 months and who were using inhaled corticosteroids as a maintenance treatment. 1,578 participants were randomised, 11 were excluded after randomisation, leaving 1,567 participants: 788 in the theophylline arm and 779 in the placebo arm.

The aim of the health economic evaluation was to evaluate the cost-effectiveness of adding low-dose theophylline as an adjunctive therapy to inhaled corticosteroids, compared to a placebo, for reducing exacerbations in the trial population; cost-effectiveness was assessed using the incremental cost per QALY summary ICER measure. The categories of health economic resource use collected included treatment costs, exacerbation costs, COPD costs, emergency non-COPD hospital admissions and health service use not related to exacerbations.

The main trial analysis found that there was no statistically significant difference in the number of exacerbations between arms; there was no treatment effect observed for theophylline. The economic evaluation complete case cost results (Table 17) found that the total costs were higher (statistically significant) in the placebo arm. Further investigation found that this was due to higher exacerbation costs in the placebo arm, which was driven by more exacerbations in the placebo arm requiring hospital treatment; exacerbations treated in hospital are more costly than exacerbations treated in other

locations. (This is explained in more detail in Chapter 3 sub-section 3.2.4). Treatment costs showed a statistically significant difference between arms, higher in the theophylline arm; £22 compared to no cost for placebo. No other cost categories showed a difference between arms. QALYs were higher in the placebo arm compared to the theophylline arm, although this was not a statistically significant result.

Table 17: TWICS health economic evaluation results

	Theophylline	**Placebo**	**Difference between arms (95% CI)**
Exacerbations (Mean)	2.24	2.23	Unadjusted IRR 1.00 (0.92 to 1.09)
Complete case			
Treatment costs	£22	£0	£22 (£22 to £22)
Exacerbation costs	£585	£1,033	-£447 (-£709 to -£186)
Non-treatment, non-exacerbation costs	£2,075	£2,101	-£26 (-£234 to £181)
Total costs	£2,684	£3,136	-£452 (-£771 to -£133)
Total QALYs	0.626	0.637	-0.011 (-0.040 to 0.018
Multiple imputation (unadjusted)			
Total costs	£2,702	£3,141	-£439 (-£846 to -£32)
Total QALYs	0.617	0.621	-0.004 (-0.031 to 0.024)
Multiple imputation (adjusted)			
Total costs	£2,784	£3,006	-£222 (-£472 to £27)
Total QALYs	0.621	0.616	0.005 (-0.015 to 0.025)

CI – confidence interval, QALY – quality adjusted life-year

When multiple imputation was used to compensate for missing data, without adjustment for baseline characteristics, the direction of the differences in total cost and total QALYs mirrored complete case results. When multiple imputed total costs and QALYs were adjusted for baseline characteristics total costs were still higher in the placebo arm, but this was no longer statistically significant. The direction of the difference in QALY results swapped to being higher in the theophylline arm compared to placebo, however this difference remained a statistically insignificant result. In summary, after using recommended methods

to conduct the economic evaluation and characterise uncertainty the results were still uncertain: the only thing that was known for certain was that the treatment cost was higher in the intervention arm, and that exacerbation costs were higher in the placebo arm as a result of more participants requiring hospital treatment for exacerbations in that arm compared to the theophylline arm, however the trial team concluded that this was not a result of any possible treatment effect.

7.3 Applying the novel approach to the TWICS case study

7.3.1 Conceptual model

The first step in applying the novel approach to the TWICS data was to develop a conceptual model to illustrate the expected links and associations (direct and mediated) between the key components of the mechanism driving the TWICS trial, in relation to the health economic evaluation. The conceptual model development is described in detail in Chapter 6 and the TWICS conceptual model for analysis is presented again here in Figure 54. The main purpose for developing the conceptual model is to provide a framework on which to base the conceptual model driven analysis, it also provides a tool for communicating with the wider trial team to explain the economic evaluation process and why economic data collection is important, and to ask for feedback on the accuracy of the links and associations in the conceptual model.

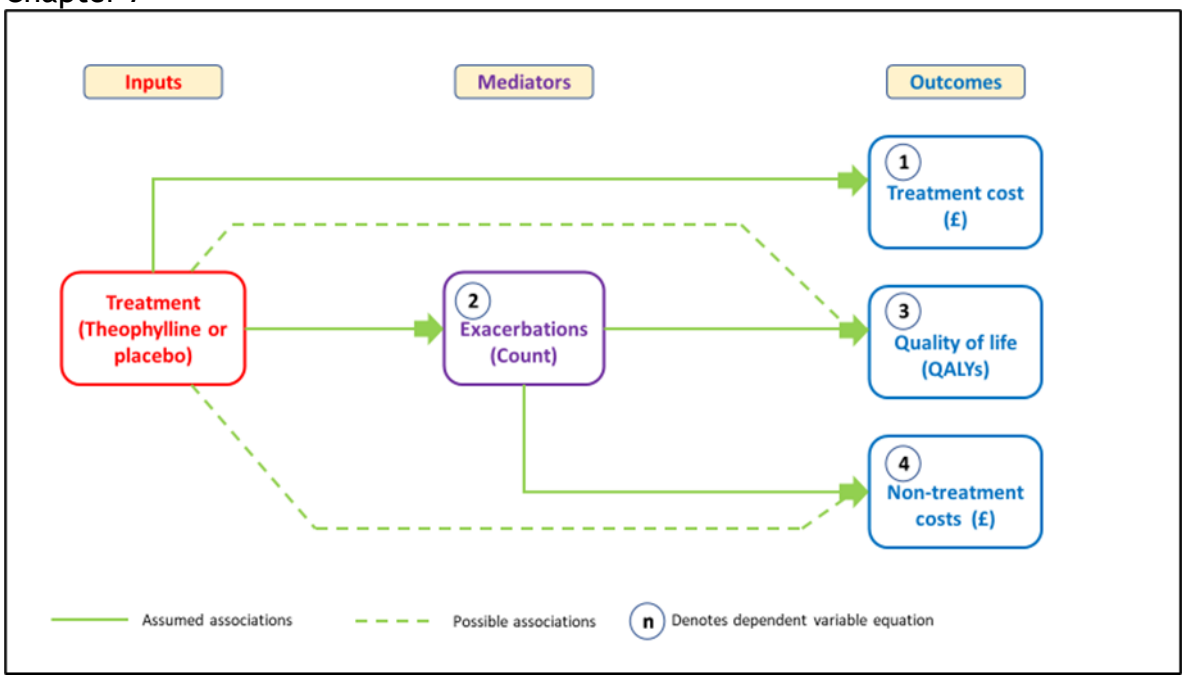


Figure 54: Final TWICS conceptual model

The conceptual model is described in full in the previous chapter, briefly it is split into three sections: 'Inputs' which is the treatment arm a participant is randomised to; 'Mediators', there is one mediator which is number of exacerbations, and 'Outcomes' which are treatment and non-treatment costs, and quality of life measured in QALYs. Each dependant variable has been given a number to identify a regression, this is to allow ease of description of the conceptual model and assumed and possible associations.

The assumed associations in the conceptual model (solid lines) are that treatment arm predicts treatment costs (regression 1); it was known a priori that the intervention medication had a small cost, and that placebo has no cost. Regression 2 assumes that treatment arm affects the number of exacerbations experienced (this is based on the overarching trial hypothesis that theophylline may reduce the number of exacerbations a participant experiences). Other assumed associations are that the number of exacerbations a participant experiences predicts quality of life (regression 3); it was expected that the number of exacerbations a participant experienced would be reflected in the quality of life reported (more exacerbations experienced would lead to lower reported quality of life). The exacerbations could be considered a surrogate endpoint; where the effect of the treatment on QALYs is completely captured by the exacerbations, this effect would need to be validated by quantifying the

relationship between exacerbations and QALYs. The final assumed associations are that the number of exacerbations also predicts non-treatment costs (regression 4); exacerbations are a cost event themselves, plus if a participant experiences a deteriorating health status due to exacerbations they may face an increase their healthcare use.

There are two possible associations in the conceptual model, both are predicted by the treatment arm; 1) quality of life (regression 3), and 2) non-treatment costs (regression 4). These possible associations are equivalent to an intervention focussed evaluation, where there is no consideration of the process driving the results, only considering the treatment and outcomes. Possible drivers for these associations include treatment side effects (although these were not expected in TWICS.)

Overall, the main assumptions in the conceptual model were that the intervention would lower the number of exacerbations experienced by participants in the intervention arm, conversely the participants in the placebo arm were expected to experience higher numbers of exacerbations. The number of exacerbations a participant experienced would drive the quality of life of that participant, it was also possible that treatment arm would directly affect quality of life without being mediated through the number of exacerbations. It was assumed that the treatment arm would drive the level of treatment costs. Non-treatment costs were assumed to be driven by the number of exacerbations experienced, however it was also possible that non-treatment costs were driven by treatment arm too. The possible associations replicated the 'black box' evaluation where treatment arm is regarded as being directly responsible for quality of life and non-treatment costs.

7.3.2 Testing the accuracy of the conceptual model

The next step in the novel approach is to test the accuracy of the conceptual model; once the conceptual model is finalised the accuracy of the predicted associations should be tested, this is done using regression techniques based on the numbered dependent variables in the conceptual model. Regression techniques were chosen for this as they estimate the relationships (or associations) between the independent variable and one of more predicted

variables; this identifies which predicted variables are impacted by the independent variable. Testing the accuracy of the associations confirms the accuracy of the structure of the conceptual model and whether the assumptions made about associations and links are correct. The following equations represent the regressions identified and presented in the conceptual model diagram:

1. Treatment cost = constant + (beta1 * treatment arm)

2. Exacerbations = constant + (beta2 * treatment arm)

3. Quality of life = constant + (beta3 * exacerbations) + (beta4 * treatment arm)

4. Non-treatment costs = constant + (beta5 * exacerbations) + (beta6 * treatment arm)

Details of the variables in the TWICS clinical trial data used to represent the components in the regressions are presented in Table 18. As there was a small amount of missing data in the trial dataset, for this demonstration missing cost and QALY data were replaced with a treatment arm specific mean, this produced a full dataset on which to base this demonstration.

Table 18: Variables used in TWICS case study regressions

Entity/variable	Description
Treatment arm	Binary indicator of the treatment arm each participant was randomised to; theophylline or placebo
Exacerbations	Number of exacerbations reported by each participant (requiring treatment) over the follow-up period
Treatment cost	Cost of theophylline treatment to participants in the theophylline arm (placebo was zero cost)
Non-treatment costs	Cost per participant of all non-treatment resources used during the follow-up period
Quality of life	Quality adjusted life-years accumulated over the follow-up period, reported in EQ-5D-3L questionnaire

The results from the equations defined in the conceptual model are presented in Table 19, with strong associations between independent and predicted variables highlighted in red and summarised below.

Equation 1) confirms that treatment costs only apply to the theophylline arm and are strongly associated with the intervention arm; £22 per participant annually.

Equation 2) shows a trend towards 0.008 more exacerbations in the theophylline arm, a weak association, confirming the main clinical trial results; that theophylline does not affect the number of exacerbations needing treatment.

Equation 3) demonstrates that treatment arm does not have a direct effect on quality of life, predicting 0.011 less QALYs in the theophylline arm, a weak association. However, regression results demonstrate that the number of exacerbations does have an effect on quality of life; each additional exacerbation results in a 0.022 reduction in QALYs, this is a strong association.

Equation 4) shows that treatment arm does predict non-treatment costs; costs are £447 lower in the intervention arm (a strong association). Exacerbations have a direct effect on non-treatment costs; each additional exacerbation contributes an additional £529 to non-treatment costs, this is a strong association.

Table 19: Regression results from the conceptual model equations

Equation	Dependant variable	Predictor variable(s)	Constant	Co-efficient	95% CI
1	Treatment cost	Treatment arm	£0	£22.0	(21.7 to 22.3)
2	Exacerbations	Treatment arm	2.23	0.008	-0.190 to 0.206
3	Quality of life	Treatment arm	0.687	-0.011	-0.034 to 0.012
		Exacerbations		-0.022	-0.028 to -0.017
4	Non-treatment costs	Treatment arm	£1,956	-£447	-853 to -103
		Exacerbations		£529	434 to 624

CI – confidence interval

In equation 4) it was known that the strong association between treatment arm and non-treatment costs was likely driven by a larger number of exacerbations requiring hospital treatment in the placebo arm, which are more costly to treat

than exacerbations treated in the primary care sector or at home. A sensitivity analysis was conducted on equation 4) results to assess whether the treatment arm effect on non-treatment costs remained independently of the location of treatment for exacerbations (Table **20**). Exacerbation costs were split into the costs of exacerbations treated in hospital and costs of exacerbations not treated in hospital, plus the additional costs of medication and oxygen required to manage the exacerbation were assessed separately. As expected, treatment arm was a strong predictor of the costs of exacerbations treated in hospital but not of non-hospital treated exacerbations, suggesting that theophylline reduces the number of exacerbations requiring hospitalisation, (further investigation in the original trial clinical effectiveness results concluded that this result was driven by a small number of participants in the placebo arm and not by any treatment effect; theophylline did not reduce the number of exacerbations treated in hospital). Treatment arm was not a strong predictor of the cost medications and oxygen used to treat exacerbations. The same regression was run with non-treatment, non-exacerbation costs. The results confirmed that when removing treatment and exacerbation costs from total costs, treatment arm does not have any effect on the remaining costs, and furthermore, the number of exacerbations is a strong predictor of costs; each exacerbation predicted an increase of £193 (95% CI £137 to £249) in non-treatment, non-exacerbation costs.

The results of this sensitivity analysis showed that the difference between arms for non-treatment costs was driven by the cost of hospital treated exacerbations.

Table 20: Additional analysis on equation 4

Dependant variable	Predictor variable(s)	Constant	Co-efficient	95% CI
Exacerbation costs	Treatment arm	£989	-£409	-£672 to -£147
Exacerbation costs - treated in hospital	Treatment arm	£812	-£381	-£624 to -£138
Exacerbation costs - not treated in hospital	Treatment arm	£103	-£5	-£21 to £12
Non-location exacerbation treatment costs	Treatment arm	£74	-£24	-£47 to £0.04
Non-treatment, non-exacerbation costs	Treatment arm	£1,674	-£28	-£251 to £196
	Exacerbations		£193	£137 to £249

CI – confidence interval

To explore this further the mean cost of an exacerbation treated in hospital in the placebo arm was calculated, £3,613, compared to £2,671 in the theophylline arm, a difference of £941 (95% CI £140 to £1,743). This difference was driven by longer lengths of stay in the placebo arm compared to the theophylline arm; overall the distribution of length of stay was similar in both arms, however there were 10 participants in the placebo arm with lengths of stay greater than 40 days (and therefore high exacerbation costs), two of these had a length of stay of over 100 days. This compares to no participants in the theophylline arm with a length of stay greater than 40 days. This can be seen in Figure 55 which illustrates that the majority of stays in both arms were below 40 days. Furthermore, when these 10 participants (with longer lengths of stay) are omitted from data in equation 4, treatment arm is no longer a strong predictor of non-treatment costs.

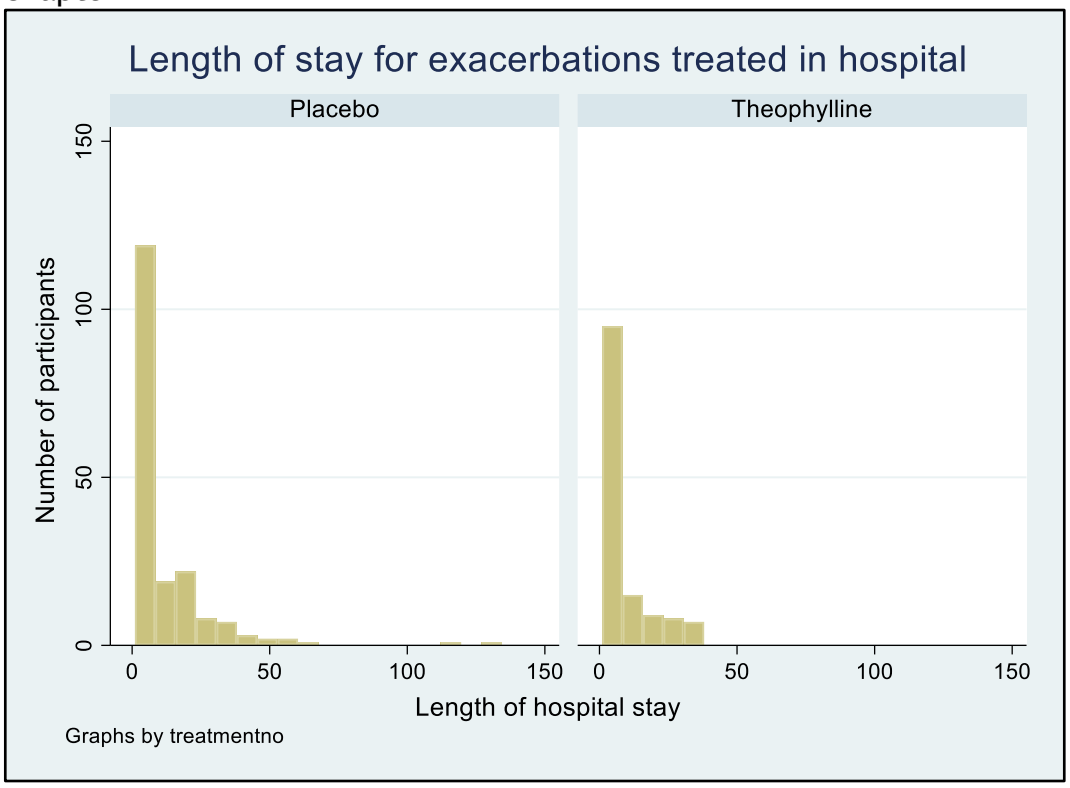


Figure 55:TWICS length of stay - hospital treated exacerbations (all length of stays)

This step of performing regressions in the novel approach tested the accuracy of the conceptual model. The regressions confirmed that there were two predictor components in the TWICS conceptual model: treatment arm and number of exacerbations. Treatment arm was strongly associated with treatment cost and non-treatment costs. However, the association between treatment arm and non-treatment costs resulted from a small number of participants in the placebo arm requiring hospital treatment for exacerbations, and treatment arm was not associated with other resource categories included in non-treatment costs. Clinicians in the TWICS trial did not believe there was any plausible biological mechanism linking theophylline to a reduction in exacerbations needing hospital treatment (109). The sensitivity analysis presented in this section showed that treatment arm was only strongly associated with the costs of exacerbations treated in hospital, not the costs of exacerbations treated elsewhere, which was driving the regression results of the non-treatment costs. The number of exacerbations were strongly associated with quality of life and non-treatment costs; these were both assumed associations.

In conclusion the assumed associations in the conceptual model were accurate, except treatment arm predicting number of exacerbations. Furthermore, the

regression results confirmed that the possible association of treatment arm predicting quality of life was not strong, but that treatment arm predicted non-treatment costs directly, this goes some way to confirm that the trial assumption of no treatment side effects from theophylline, and confirms the original treatment arm-based analysis results of no difference between arms for QALYs, the only difference was treatment and non-treatment costs. The regression results also validate the inclusion of exacerbations as a mediator in the conceptual model as they predict both quality of life and non-treatment costs. The results from this step in the novel approach could be used to inform a future trial; the number of exacerbations affect quality of life when measured using the EQ-5D-3L, and if a future intervention was expected to affect the severity of exacerbations the economic evaluation should look at the location of treatment for different severity of exacerbations.

7.3.3 Final analysis results

The final step in the new approach is to use the conceptual model as a framework to further analyse trial data, in this step the results from the regressions (testing the accuracy of the conceptual model) are used to inform similar summary measures and measures of uncertainty to the original treatment arm-based analysis. The aim of this final analysis is to apply the trial mechanism driving the economic evaluation in the conceptual model to the trial data, to give a more detailed understanding of the trial than the treatment arm-based analysis. The results of this final analysis are compared to the results of the original analysis; this comparison is important as it provides the additional interpretation and understanding of the economic evaluation results based on the expected causal mechanism in the clinical trial; presenting the final analysis results on their own would not provide this additional understanding.

This final analysis is operationalised by applying the results of the above regressions in sub-section 7.3.2 and taking forward regressions with strong predictors to inform bootstrapped samples using the TWICS trial data. Where there are mediating relationships, the output of one regression should feed into another. The regressions included in the TWICS case study were:

1. Treatment costs were predicted by treatment arm

2. Exacerbations were predicted by treatment arm

3. Quality of life was predicted by exacerbations

4. Non-treatment costs were predicted by exacerbations

Regression 1 is used to predict treatment costs, regression 2 is used to predict exacerbations. The exacerbations predicted in regression 2 are used to predict quality of life in regression 3 and non-treatment costs in regression 4. Treatment and non-treatment costs were summed to create a new variable for total costs and the difference in total costs and QALYs was calculated. Treatment costs, non-treatment costs, total costs and total QALYs are included in bootstrapping syntax to replicate 1,000 iterations. Annotated Stata syntax is included in Appendix 13: Stata syntax for TWICS and BeatIt conceptual model analysis case studies. The resulting bootstrap dataset is then used to calculate mean cost and QALY differences and produce cost-effectiveness planes and a cost-effectiveness acceptability curve. The purpose of producing cost-effectiveness planes and cost-effectiveness acceptability curves is to graphically represent uncertainty in the conceptual model driven analysis.

Bootstrapping techniques have been chosen as they do not make assumptions about distributions and estimate a large sample size treating the study sample as a population and allowing the presentation of confidence intervals, uncertainty using cost-effectiveness planes and cost-effectiveness acceptability curves; the larger the sample size the more confidence can be placed on the summary statistic. In this analysis the association between treatment arm and non-treatment costs was removed despite it being strong, this decision was made based on the clinical assumption that theophylline does not reduce the number of exacerbations requiring hospitalisation. This may lead to bias in the final TWICS conceptual model analysis results, potentially creating bias away from theophylline; if this association had been left in the conceptual model results, they may have favoured theophylline as being cost saving. More generally, if associations are removed from conceptual models, as they have been in this case study, there is a potential for the conceptual model analysis to produce misleading results if the association stands. In this illustrative example the clinician’s opinion that this was a chance finding was taken at face value when

developing the conceptual model, however the additional post-hoc analysis undertaken in this thesis suggests a different model, where exacerbation costs are separate from the remaining non-treatment costs. In this different conceptual model, and contra to clinical opinion, the association between treatment arm and the cost of exacerbations remains, but the link between treatment arm and remaining non-treatment costs is removed. However, without a prior hypothesis that theophylline reduces exacerbations requiring hospitalisation, justification of post-hoc findings would be problematic, particularly if these go against clinical opinion and notions of implausibility from clinicians who designed the trial. Finally, future clinical trials in this field should consider the potential for this finding to reoccur, possibly reassessing current evidence about the link.

Please note that although treatment arm was not a strong predictor of number of exacerbations this regression is included to operationalise the bootstrapped samples. This approach is discussed in sub-section7.6.3.

Figure 56 presents the cost-effectiveness plane using all regressions in the conceptual model bootstrap samples. This is similar to the treatment arm-based complete case cost-effectiveness plane (Figure 4) where the majority of the bootstrapped samples fall into the southwest quadrant; decision makers would need to decide what level of cost saving per loss in health benefit they would be willing to accept. However, the confidence ellipse crosses into the northwest, southwest and southeast quadrants illustrating the uncertainty in the samples. This cost-effectiveness plane is included to show the results of including all associations in the conceptual model regardless of their accuracy.

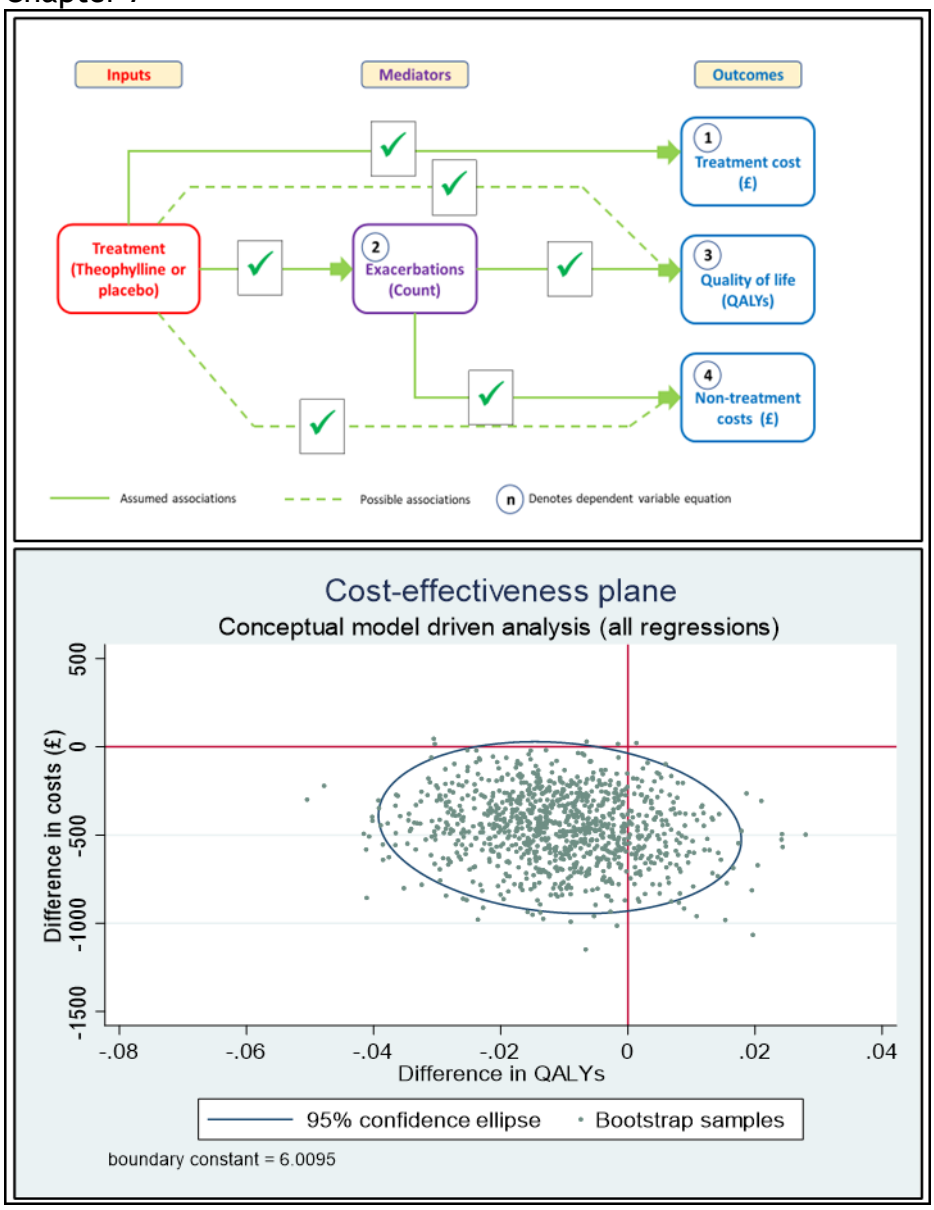


Figure 56: TWICS cost-effectiveness plane – conceptual model driven analysis (including all regressions in the conceptual model)

Although in this case study each component was predicted by only one other component, the annotated syntax in Appendix 13: Stata syntax for TWICS and BeatIt conceptual model analysis case studies includes syntax to produce Figure 56 where all equations were included, this gives guidance to the reader to operationalise a conceptual model which has more than one predictor for one or more equations.

Figure 57 presents the conceptual model driven analysis incorporating the regression results: in equation 1) treatment costs are predicted by treatment arm, in equation 2) the number of exacerbations is predicted by treatment arm,

in equation 3) QALYs are predicted by the number of exacerbations and in equation 4) non-treatment costs are predicted by the number of exacerbations. The cost-effectiveness plane shows that, compared to the cost-effectiveness plane with all associations, the plots are more concentrated, covering a smaller area, they also cross the axis close to the origin, however the ellipse still crosses into three quadrants. The shape of the plot clearly shows that as incremental QALYs decrease the incremental costs increase, this is driven by the exacerbations included in the analysis; as the number of exacerbations experienced by a participant increases so do the costs for treating those exacerbations, and a higher number of exacerbations also leads to a decrease in quality of life. By removing the association between treatment arm and non-treatment costs I have mimicked using an average cost for exacerbations, which removes the noise of the more expensive hospitalised exacerbations in the placebo arm. Removing treatment arm from predicting quality of life also removes the noise of this regression, leaving a less uncertain prediction of quality of life from exacerbations. To present this cost-effectiveness plane in more detail it has been replicated using different scales on the axes (Figure 58).

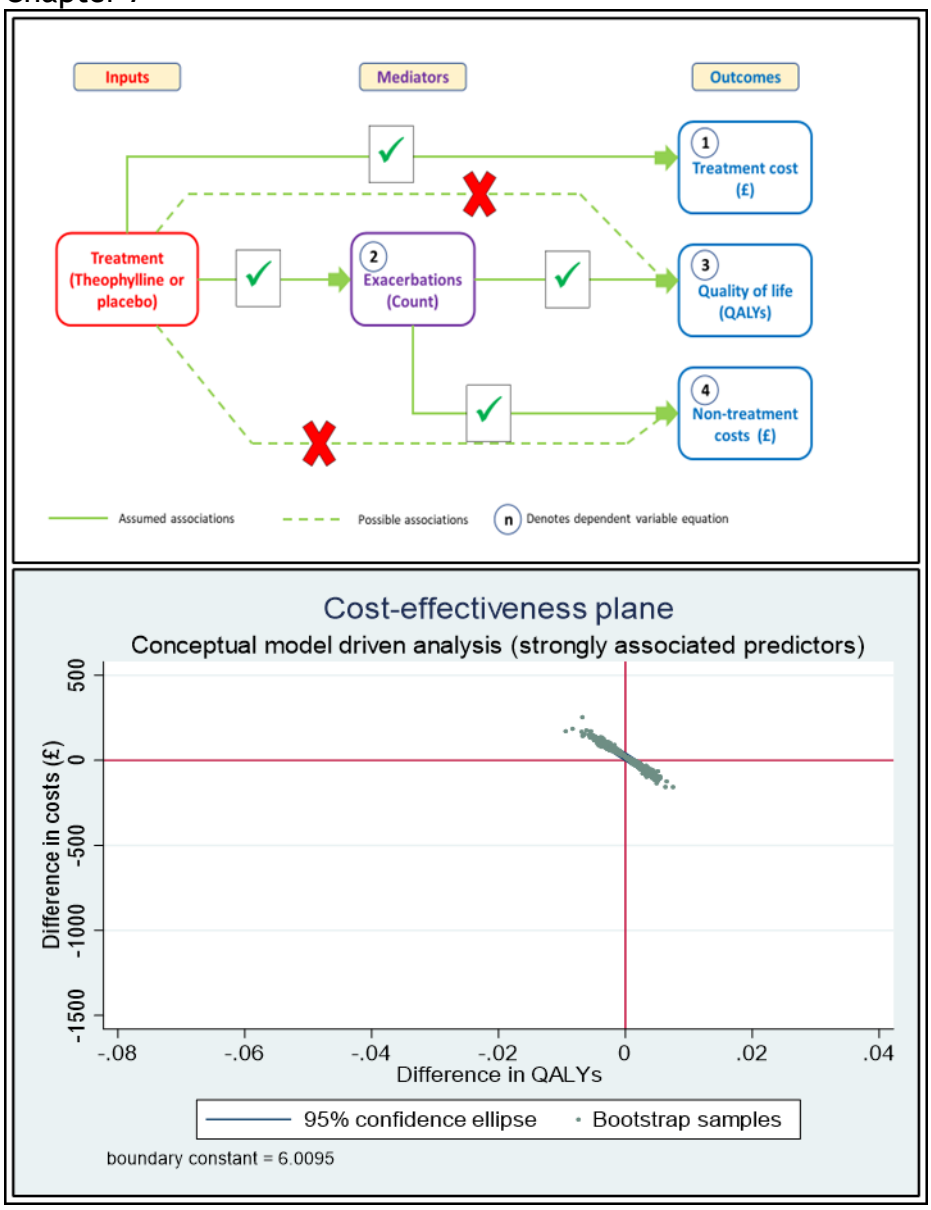


Figure 57: TWICS cost-effectiveness plane – conceptual model driven analysis (with strongly associated predictors)

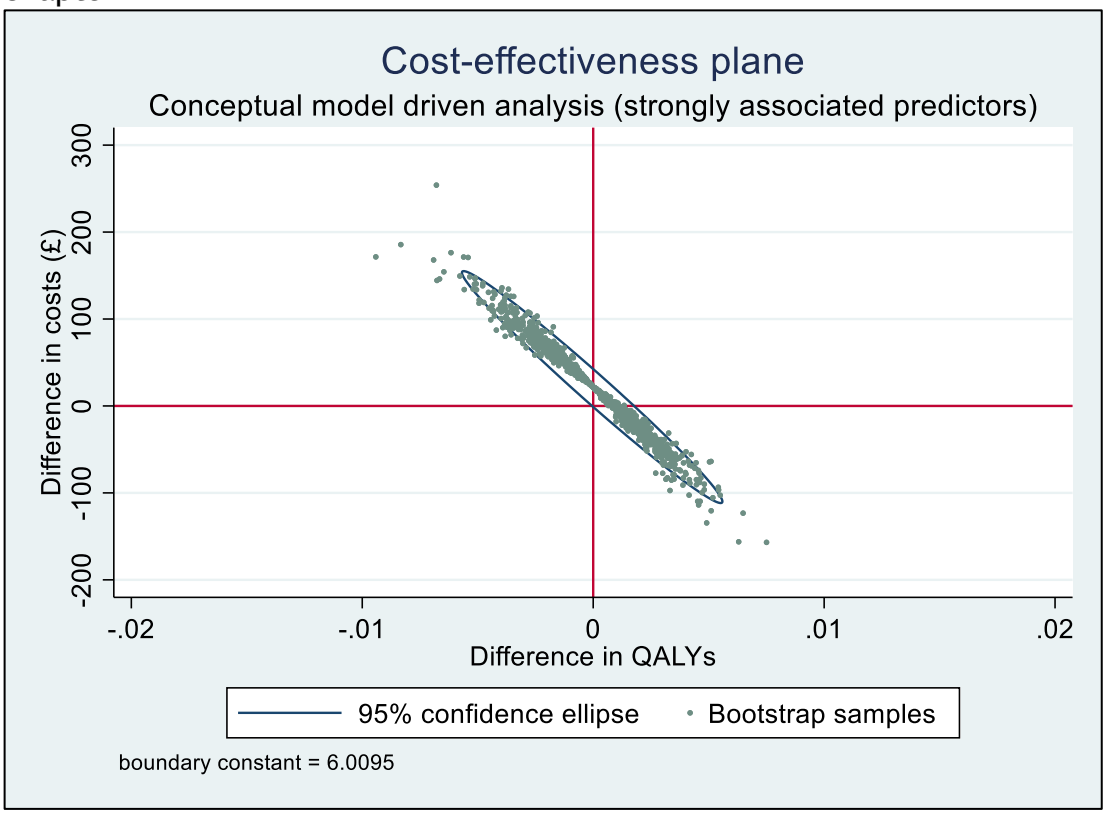


Figure 58: TWICS cost-effectiveness plane – conceptual model driven analysis (with strongly associated predictors) - large scale axis

Next bootstrapped samples should be used to produce a cost-effectiveness acceptability curve, presented in Figure 59, this shows that at a willingness-to-pay threshold of £20,000 the theophylline arm has a 41.3% chance of being cost-effective.

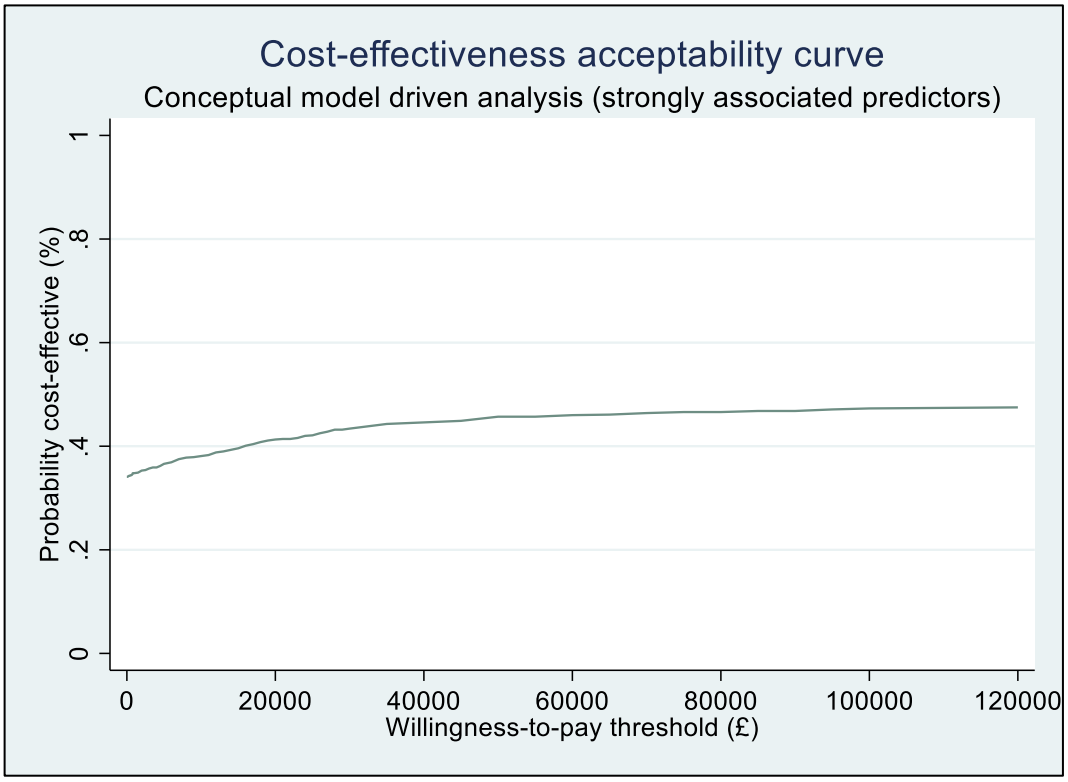


Figure 59: TWICS cost-effectiveness acceptability curve - conceptual model driven analysis

The last step in this final analysis in the novel approach is to use the bootstrapped samples to calculate point estimates, summary measures (for example ICER and net monetary benefit), and value of information estimates, and to compare these to the original treatment arm-based analysis results. The crucial aspect of this last step is the comparison of these conceptual model driven analysis results (including the cost-effectiveness planes and cost-effectiveness acceptability curves presented above) to the original treatment arm-based analysis results. The aim of this comparison is to provide additional insight and more detailed understanding of the original treatment arm-based results, by interpreting the differences in these measures and estimates.

Point estimates and 95% confidence intervals for cost and QALY differences and incremental net monetary benefits, ICERs and measures of value of information are presented in Table 21 for the treatment arm-based and conceptual model driven analyses. Measures included are taken from bootstrapped results for both the treatment arm-based analysis and conceptual model driven analysis to allow direct comparison between the results for both approaches. Measures to characterise uncertainty around the ICERs are not included as they are ratios and there are statistical complexities attached to calculating uncertainties around ratios particularly when the bootstrapped samples fall into all quadrants of the cost-effectiveness plane as it is not possible to identify which quadrant the positive and negative ICERs fall into, however, uncertainty was presented in the cost-effectiveness planes using a confidence ellipse which characterises uncertainty around the ICER clearly, with none of the pitfalls just described.

Table 21: TWICS results from both analyses

	Treatment arm-based analysis Mean (95% CI)	**Conceptual model driven analysis Mean (95% CI)**
Cost differences	-£458 (-£863 to -£86)	£22 (-£85 to £131)
QALY differences	-0.011 (-0.034 to 0.012)	-0.00004 (-0.005 to 0.004)
ICER	£48,186	-£18,869
Incremental NMB	£244 (-£370 to £894)	-£23 (-£217 to £168)
EVPI	£41	£29
EVPPI – QALYs	£19	£9
EVPPI - Costs	£8	£12
EVPPI – Treatment costs		£0
EVPPI – Non-treatment costs		£12

CI - confidence interval, EVPI - expected value of perfect information, EVPPI - expected value of perfect parameter information, ICER - incremental cost-effectiveness ratio, NMB - net monetary benefit, QALY - quality adjusted life-year

In the treatment arm-based analysis the theophylline arm is less costly and less effective than the placebo arm; the decision maker would need to decide what level of compensation (or cost saving) is acceptable for each QALY lost. In the conceptual model driven analysis the theophylline arm is more costly and less effective than the placebo arm, putting the ICER estimate in the northwest quadrant of the cost-effectiveness plane where theophylline would be considered to be dominated, however the 95% confidence intervals for cost and

QALY differences cross zero so there is some uncertainty as to the exact value of these differences. This comparison suggests that theophylline is not cost-effective and should not be adopted. The 95% confidence interval decreases more than three times for cost differences and five times for QALY difference in the conceptual model driven analysis compared to the treatment arm-based analysis.

The less uncertain results from the conceptual model driven analysis are clearly illustrated in Figure 60; a comparison of cost-effectiveness planes for both analyses. In the treatment arm-based analysis cost-effectiveness plane (top) most of the bootstrapped samples are in the southwest quadrant where the decision maker should consider their willingness-to-accept a reduction in health benefits. In the conceptual model driven analysis cost-effectiveness plane (bottom) most samples fall into the northwest quadrant where theophylline is considered dominated and should be rejected. However, a significant number of samples fall into the southeast quadrant where theophylline is considered dominant and would be adopted, this inconclusive result is driven by the number of exacerbations a participant experiences and should be interpreted with caution due to the uncertainty. Despite uncertain results in both analyses it is clear that the conceptual model driven analysis results in less uncertain estimations, shown by the smaller spread of bootstrapped samples.

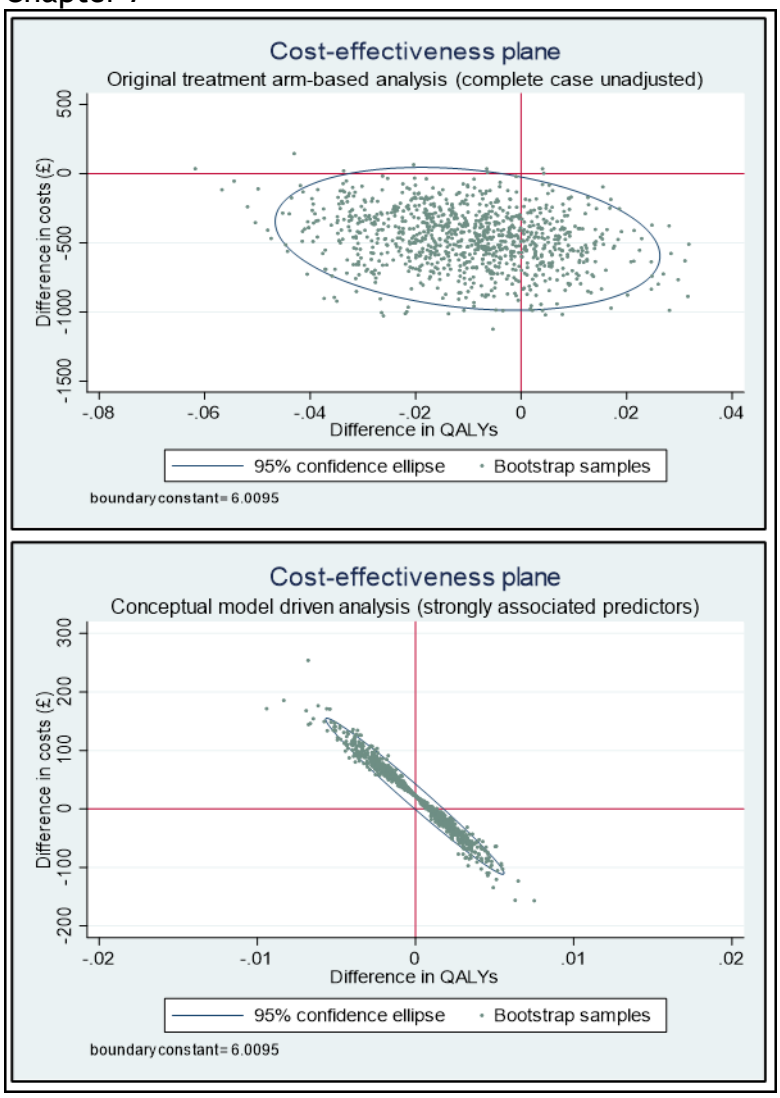


Figure 60: BeatIt cost-effectiveness planes - comparison of treatment arm-based and conceptual model driven analyses

The cost-effectiveness acceptability curves from both analyses are presented in Figure 61 for comparison. The treatment arm-based analysis curve (top) shows that at a willingness-to-pay threshold of £20,000 the theophylline arm has a 75% chance of being cost-effective compared to placebo, however in the conceptual model driven analysis the chance of being cost-effective compared to placebo is 41%. The treatment arm-based analysis cost-effectiveness acceptability curve shows that at a low willingness-to pay threshold theophylline has a high probability of being cost-effective compared to placebo, but that this probability decreases as the threshold increases. In the conceptual model driven analysis theophylline has a probability of less than 40% of being cost-effective compared to placebo, this increases slightly as the threshold increases but remains constant beyond a threshold of £40,000. This result confirms the results from the cost-effectiveness planes; in the treatment arm-based analysis the decision

maker would need to access their willingness-to-accept, and in the conceptual model driven analysis theophylline is not cost-effective at any level of willingness-to-pay threshold.

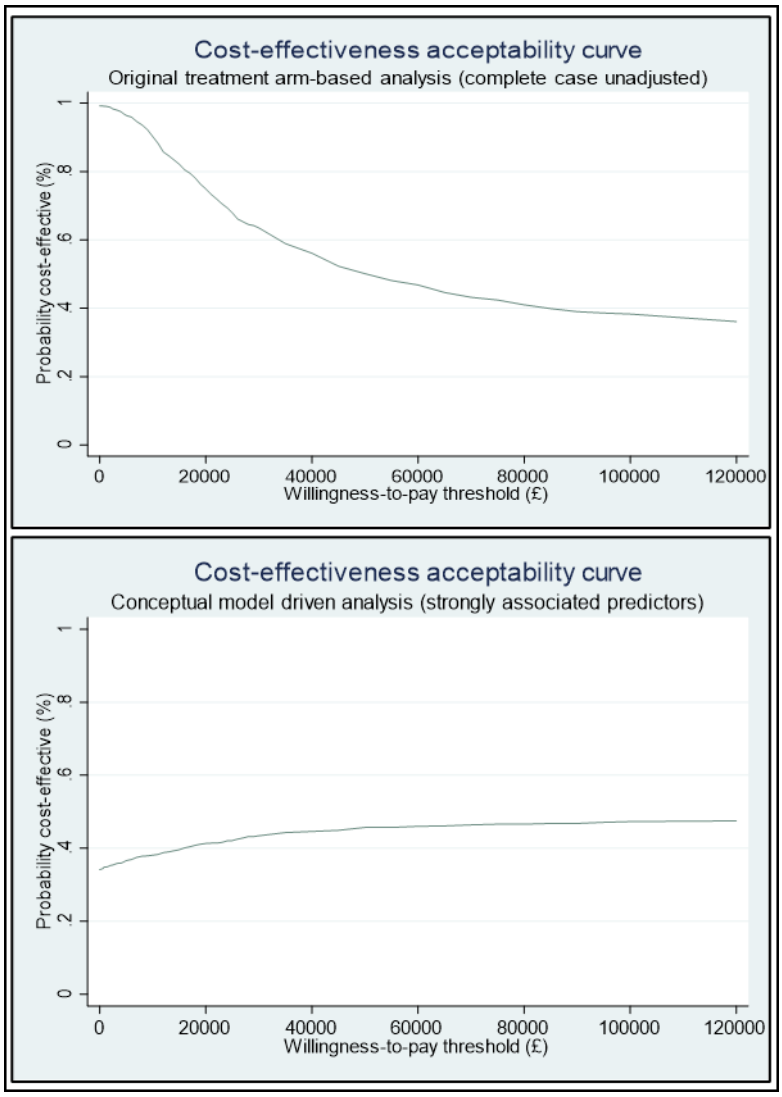


Figure 61: TWICS cost-effectiveness acceptability curves – comparison of treatment arm-based and conceptual model driven analyses

The incremental net monetary benefit is positive in the treatment arm-based analysis, indicating that the theophylline intervention should be accepted, although the 95% confidence interval around this measure crosses zero so the results are uncertain. In the conceptual model driven analysis, the incremental net monetary benefit is negative, indicating that the theophylline intervention should be rejected, again the 95% confidence interval around the incremental net monetary benefit crosses zero. However, the 95% confidence interval is three times narrower in the conceptual model driven analysis compared to the

treatment arm-based analysis, suggesting the summary measure is less uncertain in the conceptual model driven analysis. The net monetary benefit results from both approaches are plotted in a histogram to illustrate the variation of the results in each approach in Figure 62. Again, these histograms demonstrate the narrower spread of results from the conceptual model approach compared to the treatment-arm based results and the movement of the point estimate closer to zero. These results mirror the interpretations above; at some levels of willingness-to-pay decision makers would accept theophylline looking at the treatment arm-based analysis results, whereas there is no evidence in the conceptual model driven analysis results that theophylline is cost-effective.

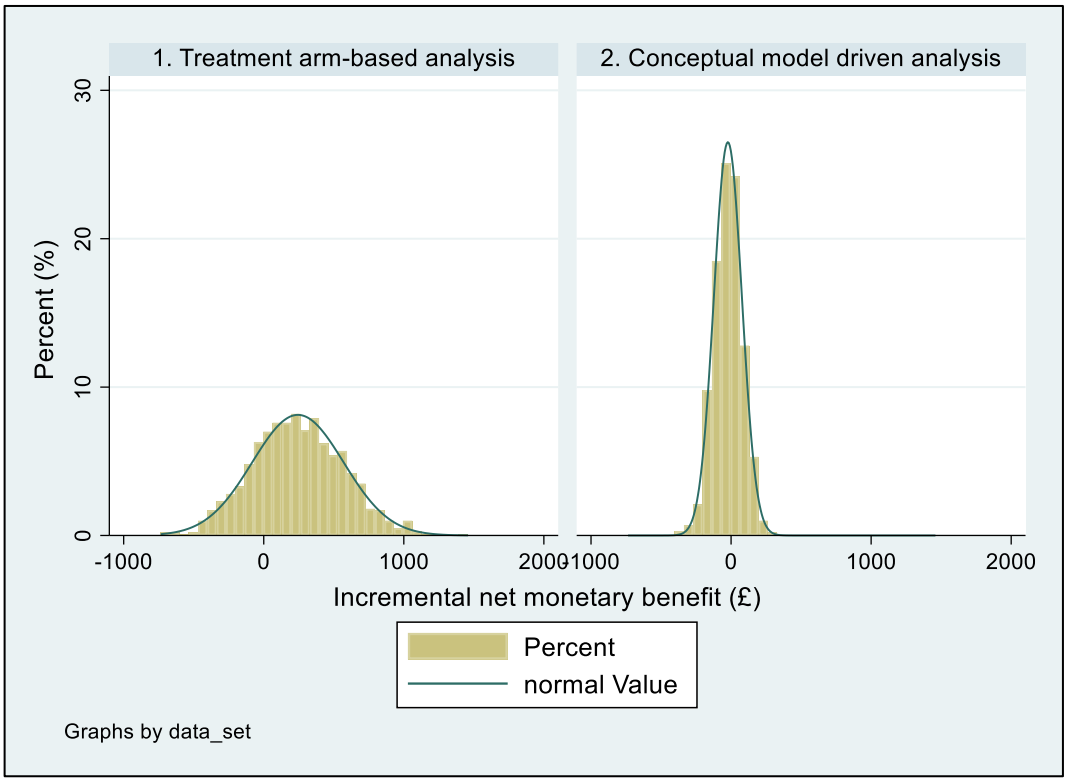


Figure 62: TWICS - histogram of net incremental monetary benefit results showing the distribution of the results from both approaches

Value of information results per person are also included in Table 21 to explore the value of eliminating sampling uncertainty; what can be gained if there was no uncertainty. These value of information results show that the expected value of perfect information is higher in the treatment-arm based approach compared to the conceptual model driven approach, £41 compared to £29; there is potentially more value in conducting further research into the treatment arm-based analysis compared to the conceptual model driven analysis, however both

EVPIs are low and it is debatable whether either would convince decision makers to commission further research. The reduced EVPI estimate is a result of the additional structure in the conceptual model driven analysis providing a more detailed understanding. Results of EVPPI for total costs are £8 in the treatment arm-based analysis and £12 in the conceptual model driven analysis, for total QALYs the results are £19 in the treatment arm-based analysis and £9 in the conceptual model driven analysis. These estimates indicate that there is more decision uncertainty linked to QALYs than to costs, and therefore potentially more to be gained by removing uncertainty in QALYs in the treatment arm-based analysis and total costs in the conceptual model driven analysis. To further investigate these results the EVPPI for treatment and non-treatment costs was estimated; this showed that there was more decision uncertainty in non-treatment costs compared to treatment costs (£12 v £0); this result confirms that treatment costs are precise are known, but the accuracy of the non-treatment cost estimate could be improved.

7.3.4 TWICS summary

In the original treatment arm-based analysis the complete case and multiple imputed unadjusted costs and QALYs were higher in the placebo arm compared to the theophylline arm, the difference in costs was statistically significant, but not for the difference in QALYs. In the multiple imputed adjusted results the costs were still higher in the placebo arm compared to the theophylline arm but the results were no longer statistically significant. The direction of the QALY results changed in the multiple imputed adjusted results with more QALYs reported in the theophylline arm, this was not a statistically significant result. Overall, although there was a trend for higher costs in the placebo arm this result was uncertain, and there was a lot of uncertainty in the QALY results. In the conceptual model driven analysis costs were marginally higher in the theophylline arm and QALYs were barely higher in the placebo arm, with the 95% confidence crossing zero for both differences.

The conceptual model driven analysis found that the treatment arm predicts treatment costs and non-treatment costs but does not predict exacerbations. A sensitivity analysis found that the effect of treatment on non-treatment costs

only held true for the costs of treating exacerbations requiring a stay in hospital. The treatment arm does not predict the costs of exacerbations requiring treatment at home or in a community setting, nor non-exacerbation costs. This sensitivity analysis illustrated that unexpected results can be analysed further. The removal of this equation (treatment arm predicting exacerbations) was based on clinical opinion, with further analysis described above showing that this strong prediction was driven by the costs of exacerbations requiring hospitalisation, however it is important to keep in mind that by excluding this regression, bias may have resulted in the final conceptual model analysis, a limitation of this case study. The conceptual model driven analysis also found that exacerbations predict quality of life and non-treatment costs, including non-treatment non-exacerbation costs. The assumed associations in the conceptual model gave an accurate representation of the trial mechanism except for the treatment arm predicting exacerbations, whereas none of the possible associations were strong and were removed from the final analysis.

As neither of the possible associations depicted in the conceptual model were strong it could be argued that this confirms: 1) the trial assumption of no treatment side effects from theophylline, and 2) the original treatment arm-based analysis results of only treatment costs showing a difference between arms. The original trial results found no treatment effect on the number of exacerbations, the conceptual model driven analysis confirmed this but also found that exacerbations do influence quality of life and non-treatment, non-exacerbation costs; validating the inclusion of exacerbations as a mediator in the conceptual model as they predict both quality of life and non-treatment costs. The results from this aspect of the novel approach could be used to inform a future trial design.

The incremental net monetary benefit favoured TWICS in the treatment arm-based analysis but not for the conceptual model driven analysis, suggesting that despite the uncertainty in the treatment arm-based results decision makers might adopt the theophylline intervention, however the net monetary benefit summary measure is calculated using complete case bootstrap results, and may overestimate the difference in costs between arms. EVPPI for the treatment arm-based analysis suggests that there is potentially more value in further

research into QALYs compared to costs resulting from this comparison of treatments, due to more decision uncertainty in QALYs. However, in the conceptual model driven analysis, there is less value in further research into QALYs compared to costs.

Overall, the conceptual model driven analysis concluded that theophylline would not be a cost-effective intervention; it does not reduce the number of exacerbations experienced or result in cost savings.

7.4 Case study #2 BeatIt

7.4.1 Background of BeatIt

The BeatIt trial compared a behavioural activation therapy (BeatIt) with a guided self-help therapy (StepUp) in a cohort of adults with an intellectual disability and a diagnosis of depression, evaluating improvements in depressive symptoms. The primary outcome was a change in measure of depressive symptoms using the GDS-LD score at 12 months follow-up. The economic outcome was the quality adjusted life-year measure using the EQ-5D-Y questionnaire. In total 161 participants were randomised with 141 providing data at the primary endpoint of 12 months (68 BeatIt and 73 StepUp).

The aim of the health economic evaluation was to evaluate the cost-effectiveness of BeatIt compared to StepUp, measured by an incremental cost per quality adjusted life-year ICER. Resource use and outcome measures were collected at baseline, 4 and 12 months with an additional collection of resource use data at 8 months reported by carers. The analysis was conducted using the intention-to-treat population from an NHS and social services perspective.

A summary of the trial-arm based analysis results is presented in Table 22. Briefly, main trial results at 4 and 12-month follow-ups demonstrated no statistical difference between arms in the GDS-LD score, however, there was a statistically significant improvement in GDS-LD scores over the follow-up period in both BeatIt and StepUp arms, suggesting that both therapies were successful in reducing depressive symptoms in this population. The economic evaluation results showed a statistically significant difference at baseline for health utilities

between treatment arms; but for 4, 8 and 12-months follow-up there were no statistically significant differences after adjusting for baseline health utilities. Similar to the primary outcome there was a statistically significant improvement in health utility scores in both arms suggesting that both therapies were successful in improving quality of life, however there was a trend for higher QALYs in the StepUp arm (0.691 compared to 0.628, a difference of 0.063 (5% CI -0.052 to 0.178). The treatment cost was higher in the BeatIt arm, a statistically significant difference between arms, however there was no statistically significant difference between arms in non-treatment costs or total costs. There was a trend for higher total costs in the BeatIt arm (£27,158 compared to £26,786), this was driven by treatment costs. Bootstrapped plots on the cost-effectiveness plane showed no difference in costs nor effects between arms. When missing data was replaced using multiple imputation, results mirrored the complete case results, however when these multiple imputation results were adjusted for baseline characteristics the direction of the previous results reversed; total costs were higher in the StepUp arm and QALYs were marginally higher in the BeatIt arm. In summary, after applying recommended methods in conducting the original treatment arm-based economic evaluation, the results of the economic evaluation were uncertain and gave little evidence for decision makers to base a decision on. The treatment costs were certain, they were higher in the intervention arm, however there was no statistically significant difference between arms for total costs or total QALYs despite using multiple imputation and adjusting for baseline characteristics. It was also clear that both treatments improved depressive symptoms but the mechanism of how this happened was not known.

Table 22: Results from BeatIt trial-arm based analysis at 12 months follow-up (complete case)

	BeatIt	StepUp	Difference between arms (95% CI)
GDS-LD score	12.43	12.03	0.40 (-2.26 to 2.70)
Treatment costs	£1,788	£1,050	£738 (£586 to £890
Non-treatment costs	£25,370	£25,736	-£367 (-£13,418 to £12,684)
Total costs	£27,158	£26,786	£371 (-£12,689 to £13,432)
Total QALYs	0.628	0.691	-0.063 (-0.178 to 0.052)
Multiple imputation (unadjusted)			
Total costs	£27,223	£26,021	£1,201 (-£11,299 to £13,702)
Total QALYs	0.617	0.693	-0.076 (-0.185 to 0.033)
Multiple imputation (adjusted)			
Total costs	£26,369	£27,962	-£1,593 (-£5,194 to £2,008)
Total QALYs	0.657	0.655	0.002 (-0.082 to 0.085)

7.5 Applying the novel approach to the BeatIt case study

7.5.1 Conceptual model

The first step in the novel approach is to develop a conceptual model illustrating the links between the inputs and outputs of the economic evaluation, mediated through the trial mechanism, the development of the conceptual model was described in detail in Chapter 6 and the final BeatIt conceptual model is presented again in Figure 63.

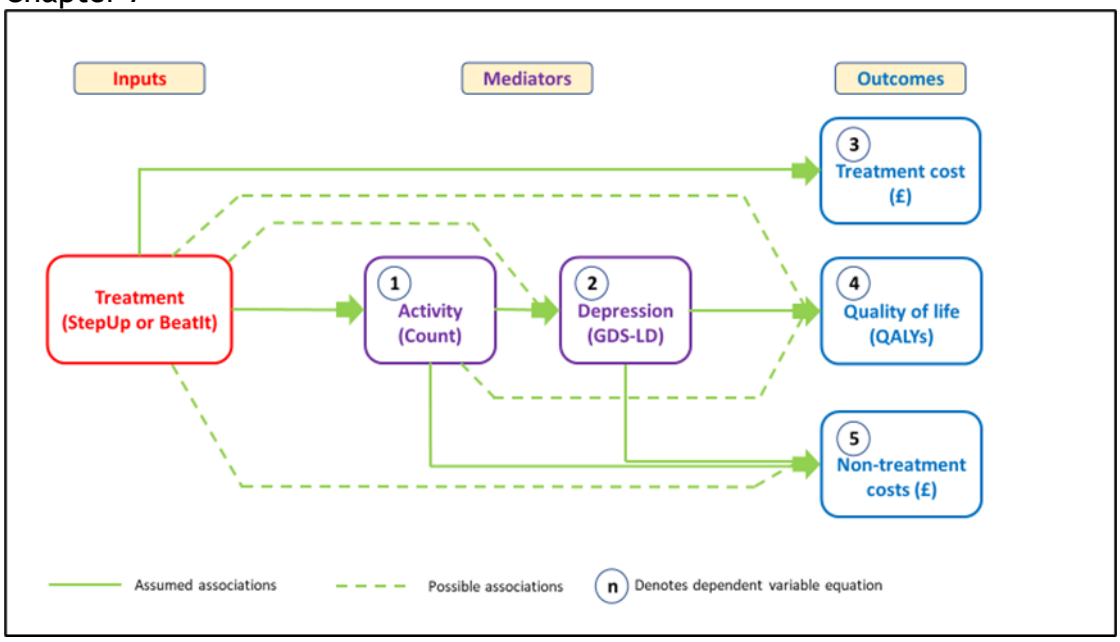


Figure 63: Final BeatIt conceptual model

As the conceptual model is described in full in the previous chapter it will only be briefly described here. The conceptual model is split into three sections: 'Inputs' which is the treatment arm; 'Mediators' which are activity and depression, and outcomes which are treatment and non-treatment costs, and quality of life. Each dependant variable has been allocated an identifying number, this enables a clear description of the conceptual model and the associations included in it.

Assumed associations include the treatment arm predicting treatment costs (regression 3) and activity levels (regression 1). The latter assumption is based on the hypothesis of the trial; the BeatIt therapy increases engagement with activities. It was also assumed that levels of activity would predict symptoms of depression (regression 2), with participants reporting higher levels of activity expected to report fewer depressive symptoms and vice versa. The level of depressive symptoms reported were assumed to predict quality of life (regression 4), with participants reporting fewer depressive symptoms expected to also report better quality of life and vice versa. Levels of reported depression could be considered a surrogate endpoint for QALYs if effects of the treatment on QALYs is completely captured by levels of depression with no other possible links to QALYs. Finally, it was assumed that activity levels and depressive symptoms would predict non-treatment costs (regression 5); higher reported levels of activity would increase non-treatment costs, and more reported

depressive symptoms were expected to lead to higher healthcare use and therefore higher non-treatment costs.

Possible associations included the treatment arm predicting levels of depressive symptoms, without being mediated through activity levels (regression 2), and treatment arm predicting quality of life and non-treatment costs, not being mediated through activity and depression (regressions 4 and 5). There is also a possibility that levels of activity are directly associated with quality of life without being mediated through depression (regression 4), for example through the 'usual activities' or 'mobility' domains of the EQ-5D-Y questionnaire.

Overall, the main assumptions in the conceptual model were that the BeatIt therapy would increase the levels of activity of participants in the intervention arm, which would in turn reduce depressive symptoms. The reduction in depressive symptoms and increase in activities were predicted to improve quality of life in these participants. The StepUp therapy was predicted to reduce depressive symptoms, but not mediated through a change in activity levels. There was a possibility that quality of life would be directly affected by treatment arm, not mediated through depression (and activity possibly). It was expected that treatment arm would directly predict treatment costs, and could possibly affect non-treatment costs, although it was assumed that non-treatment costs would be mediated through activity (and depression). The direct links between treatment arm and outcomes represents the 'black box' evaluation where no trial mechanism is considered.

7.5.2 Testing the accuracy of the conceptual model

The next step was to test the accuracy of the structure of conceptual model; to do this the associations presented in the conceptual model are tested using the numbered dependant variables in the conceptual model to represent regressions. The suitability of regression in this purpose is described in the TWICS case study. The following equations represent the regressions identified and presented in the conceptual model diagram (Figure 63), attached to each dependent variable:

1. Activity = constant + (beta1 * treatment arm)

2. Depression = constant + (beta2 * treatment arm) + (beta3 * activity)

3. Treatment cost = constant + (beta4 * treatment arm)

4. Quality of life = constant + (beta5 * treatment arm) + (beta6 * activity) + (beta7 * depression)

5. Non-treatment cost = constant + (beta8 * treatment arm) + (beta9 * activity) + (beta10 * depression)

The details of the clinical trial variables used to represent the components included in the conceptual model are presented in Table 23. To allow for full use of the data in the trial dataset missing cost and QALY data was replaced with a treatment arm specific mean.

Table 23: Variables used in the regressions

Entity/variable	Description
Treatment arm	Binary indicator of whether the treatment arm represents StepUp or BeatIt
Activity	Total reported activity over the 12 months follow-up
Depression	GDS-LD scores at 12 months follow-up. A high score indicates more depressive symptoms (more severe depression) than a low score
Quality of life	Quality adjusted life-years accumulated over the follow-up period collected using the EQ-5D-Y
Treatment cost	The cost of the therapy each participant received
Non-treatment cost	All costs relating to each participant excluding the treatment cost

Results from the equations defined in the conceptual model are reported in Table 24, with strong associations highlighted in red, and summarised below.

Equation 1) demonstrates a small trend towards higher reported levels of activity over the 12 months of the trial in the BeatIt arm, however this is not a strong association.

Equation 2) demonstrated that higher reported activity levels are associated with lower reported levels of depressive symptoms after 12 months (0.003 for each 1-point decrease in depressive symptoms), this is a small but strong association. The BeatIt treatment arm is associated with lower reported depressive symptoms; however, this is not a strong association.

Equation 3) shows that, as expected, BeatIt is associated with higher treatment costs of £738, a strong association.

Equation 4) demonstrates that lower reported depressive symptoms are associated with higher reported quality of life, this is a small but strong association of 0.017 QALYs. Higher reported levels of activity are associated with lower reported quality of life; this is a small trend and not a strong association. BeatIt is associated with a small reduction in quality of life, this is not a strong association of 0.101 QALYs.

Equation 5) shows that an improvement in reported depressive symptoms is associated with higher non-treatment costs, however this is not a strong association. Higher reported levels of activity are associated with higher non-treatment costs of £19, a small but strong association. BeatIt is associated with a £3,756 reduction in non-treatment costs, however this is a weak association.

In summary, the regressions established the relationships in the trial mechanism and found that participants with higher levels of activity also reported lower levels of depressive symptoms, treatment costs were higher in the BeatIt arm compared to the StepUp arm, participants reporting lower levels of depressive symptoms also reported higher quality of life, and finally, higher reported levels of activity resulted in higher non-treatment costs.

Table 24: Regressions results

Equation	Dependant variable	Predictor variable(s)	Constant	Co-efficient	95% CI
1	Activity	Treatment arm	495	96.5	-131 to 324
2	Depression	Activity	13.4	-0.003	-0.006 to -0.0003
		Treatment arm		-1.79	-4.74 to 1.17
3	Treatment cost	Treatment arm	£1,050	£738	£608 to £868
4	Quality of life	Depression	0.952	-0.017	-0.025 to -0.009
		Activity		-0.00006	-0.0002 to 0.00005
		Treatment arm		-0.101	-0.216 to 0.013
5	Non-treatment costs	Depression	£25,367	-£435	-£1,496 to £625
		Activity		£19.2	£5.13 to £33.2
		Treatment arm		-£3,756	-£18,581 to £11,069

CI – confidence interval

7.5.3 Final analysis results

The final step in the new approach is to use the confirmed conceptual model as a framework to further analyse trial data. In this step the results from the regressions above are used to inform estimates of similar summary measures and measures of uncertainty to the original treatment arm-based analysis. The aim of this final analysis is to apply the trial mechanism driving the economic evaluation to the trial data, the results of this final analysis are compared to the results of the original treatment arm-based analysis. This comparison is important as it provides additional interpretation and understanding of the economic evaluation results based on the expected causal mechanism in the clinical trial, compared to the traditional 'black box' evaluation; presenting the final analysis results on their own would not provide this additional understanding.

The final analysis is operationalised by applying the results of the regressions in sub-section 7.5.2, taking forward the regressions with strong predictors to

inform bootstrapped samples using the BeatIt trial data. The following regressions included in the BeatIt case study were:

1. Activity was predicted by treatment arm
2. Depressive symptoms were predicted by activity levels
3. Treatment costs were predicted by treatment arm
4. Quality of life was predicted by depressive symptoms
5. Non-treatment costs were predicted by activity levels

In regression 1 treatment arm is used to predict activity, the output from this regression is used to predict depression (equation 2) and non-treatment costs (equation 5). The output of equation 2 (depression) is used to predict quality of life in equation 4. Treatment costs are predicted by treatment arm in equation 3. Treatment and non-treatment costs were summed to produce total costs and the difference in total costs and total QALYs were used to inform the bootstrap syntax. 1,000 iterations were run and these bootstrap samples were used to calculate mean cost and QALY differences and to produce cost-effectiveness planes and cost-effectiveness acceptability curves. The aim of including these figures is to provide graphical representations of uncertainty.

Please note that although treatment arm did not strongly predict levels of activity it has been included to operationalise the bootstraps.

Figure 64 presents the cost-effectiveness plane when all associations depicted in the conceptual model are included in the bootstrapped samples. The conceptual model driven cost-effectiveness plane shows that costs are higher in the BeatIt arm compared to the StepUp arm, and QALYs are also higher in the BeatIt arm. Whilst the majority of the bootstrapped samples fall into the northeast quadrant, where BeatIt would be considered cost-effective if the ICER is below the willingness-to-pay threshold of £20,000, there is uncertainty in this result as the confidence ellipse crosses into all four quadrants of the cost-effectiveness plane. This is similar to the cost-effectiveness plane presented for the original

treatment arm-based analysis (Figure 11) where the confidence ellipse crosses into all four quadrants, but the samples in the original cost-effectiveness plane show more uncertainty in the difference between arms for costs and QALYs with samples more evenly spread in all four quadrants. The aim of including this cost-effectiveness plane is to replicate the full conceptual model regardless of the associations' accuracy.

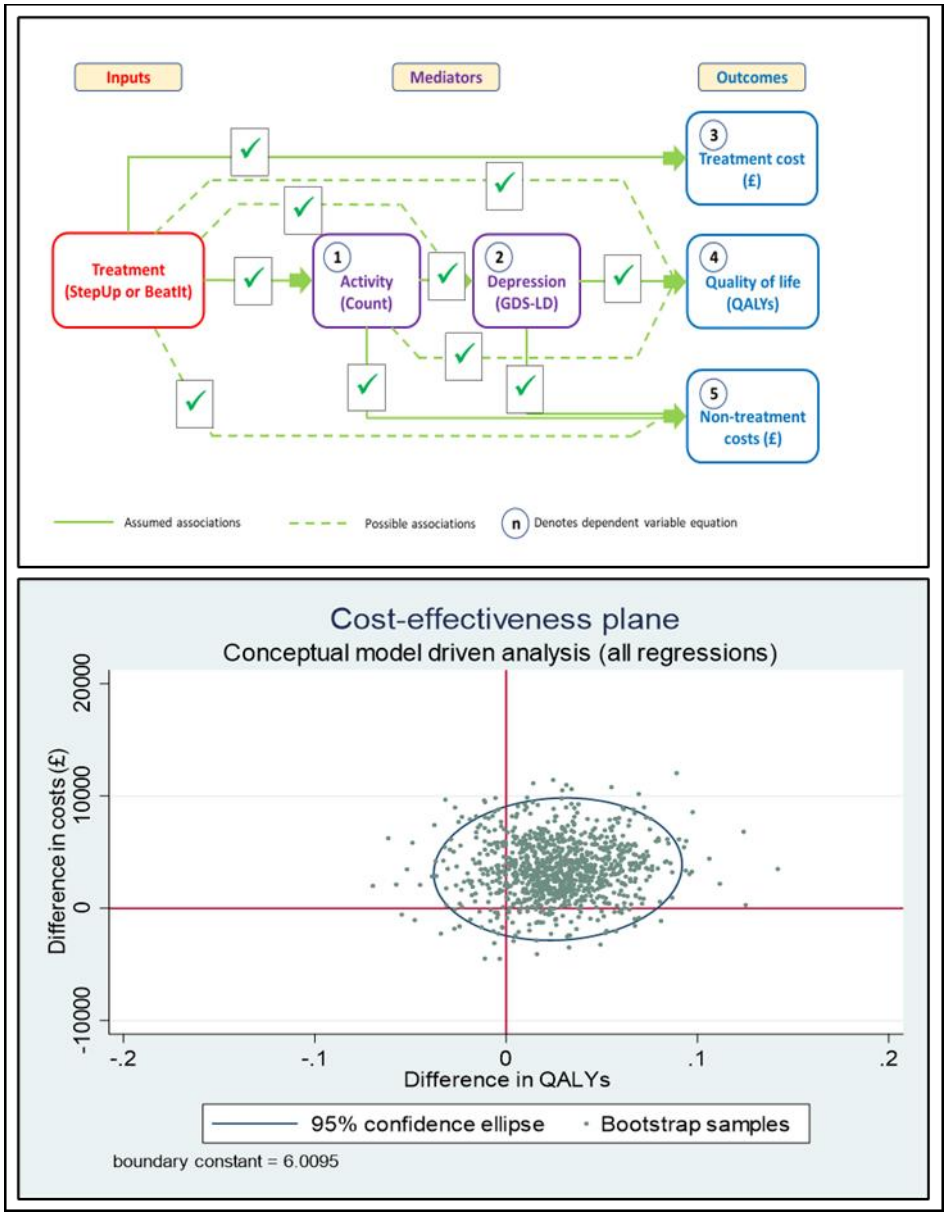


Figure 64: BeatIt cost-effectiveness plane – conceptual model driven analysis (including all regressions in the conceptual model)

Although in this case study, as in the TWICS case study, each component was predicted by only one other component, the annotated syntax in Appendix 13: Stata syntax for TWICS and BeatIt conceptual model analysis case studies

includes syntax to produce Figure 64 where all equations were included, this enables to the reader to understand how to operationalise a conceptual model which has more than one predictor for one or more equations.

Figure 65 presents the conceptual model driven analysis informed by the results from the regressions, this analysis includes the following associations: in equation 1) the treatment arm predicts levels of activity, in equation 2) levels of activity predict depressive symptoms, in equation 3) the treatment arm predicts treatment costs, in equation 4) depressive symptoms predict QALYs, and in equation 5) levels of activity predict non-treatment costs. Compared to the cost-effectiveness plane presented above with all associations (Figure 64), the plots in this cost-effectiveness plane are concentrated around the origin and cover a smaller area than previously, however the ellipse still crosses into all four quadrants. The shape of the bootstrap samples show that as incremental costs increase so do incremental QALYs, however the incremental QALYs remain close to the vertical axis. QALYs are predicted by depression, which is predicted by activity, and the regression results show that higher levels of activity led to high non-treatment costs.

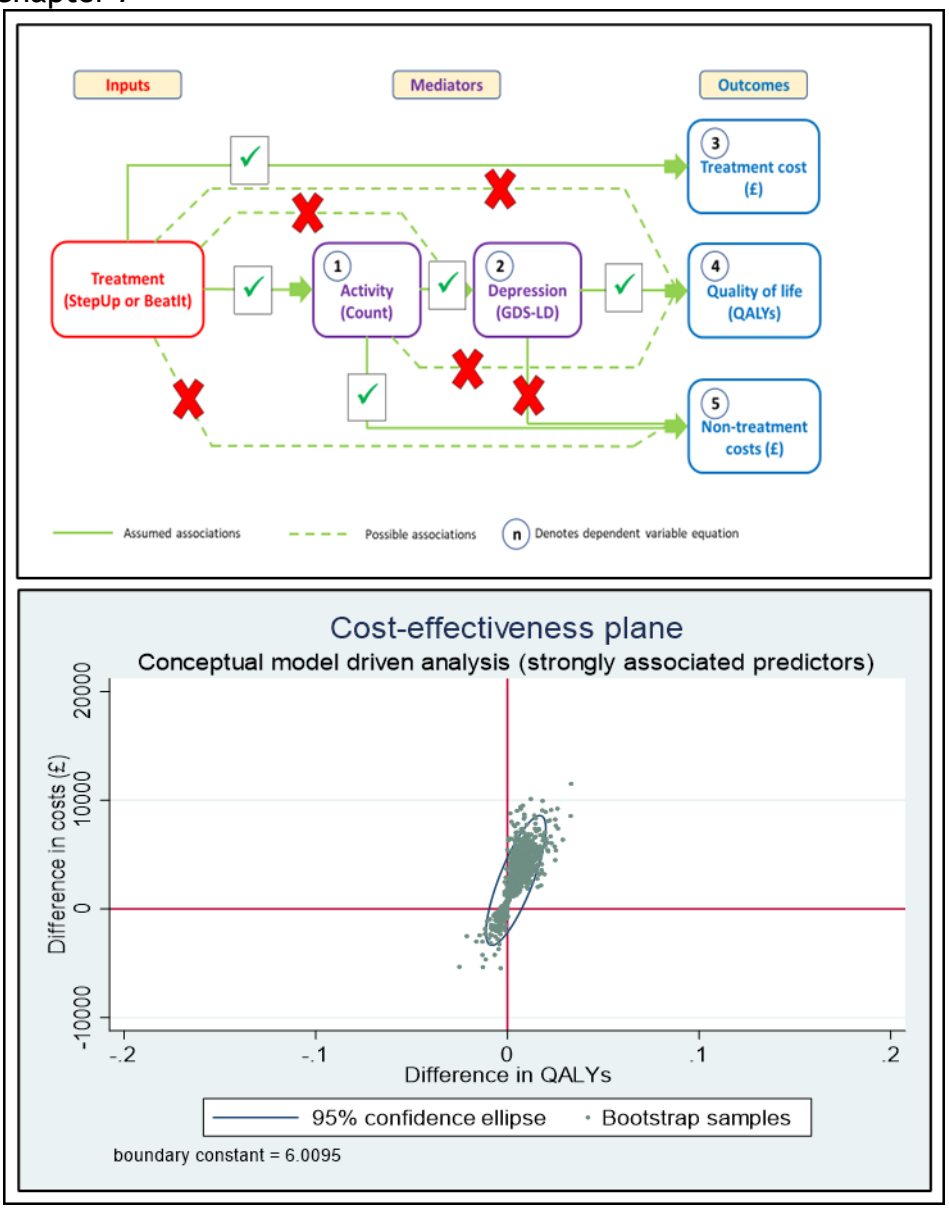


Figure 65: BeatIt cost-effectiveness plane – conceptual model driven analysis (with strongly associated predictors)

Next the bootstrapped samples should be used to produce a cost-effectiveness acceptability curve, presented in Figure 66, this shows that at a willingness-to-pay threshold of £20,000 the BeatIt arm has a 12.1% chance of being cost-effective.

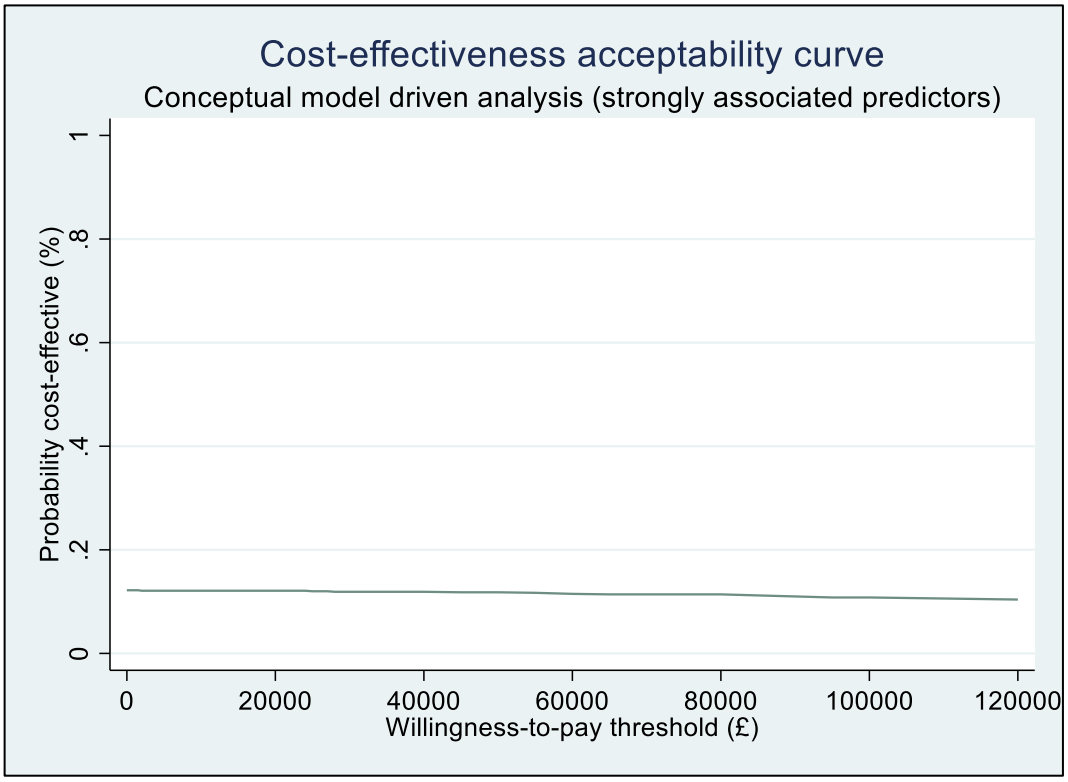


Figure 66: BeatIt cost-effectiveness acceptability curve - conceptual model driven analysis

The last step in this final analysis is to use the bootstrapped samples to calculate point estimates, summary measures (for example ICER and net monetary benefit), and value of information estimates, and compared to the original treatment arm-based analysis results. The comparison of these conceptual model driven analysis results (including the cost-effectiveness planes and cost-effectiveness acceptability curves presented above) to the original treatment arm-based analysis results is a crucial aspect of the final analysis; the aim of this comparison is to provide additional insight and understanding to the original treatment arm-based results, by interpreting the differences in these measures and estimates.

The differences in costs and QALYs and incremental net monetary benefits, with 95% confidence intervals, and ICERs and value of information estimates are presented in Table 25 for both the treatment arm-based and conceptual model driven analyses. The measures included are derived from bootstrapped results for both analyses to allow a direct comparison between the results for both approaches. Measures to characterise uncertainty around the ICERs are not included as they are ratios and there are statistical complexities attached to calculating uncertainties around ratios.

Table 25: BeatIt results from both analyses

	Treatment arm-based analysis Mean (95% CI)	**Conceptual model driven analysis Mean (95% CI)**
Cost differences	£444 (-£10,911 to £11,876)	£2,621 (-£1,965 to £7,758)
QALY differences	-0.064 (-0.168 to 0.037)	0.004 (-0.007 to 0.019)
ICER	-£105,823	£496,160
Incremental net monetary benefit	-£1,729 (-£13,495 to £10,241)	-£2,533 (-£7,558 to £1,892)
EVPI	£1,679	£148
EVPPI - QALYs	£21	£0
EVPPI – Total costs	£1,570	£171
EVPPI – Treatment costs		£0
EVPPI – Non-treatment costs		£168

CI - confidence interval, EVP - expected value of perfect information, EVPPI - expected value of perfect parameter information, ICER - incremental cost-effectiveness ratio, QALY - quality adjusted life-year

In the treatment arm-based analysis the BeatIt is more expensive and less effective than the StepUp arm, in this scenario the decision makers should reject BeatIt. In the conceptual model driven analysis BeatIt is more expensive and more effective than StepUp, in this scenario decision makers should adopt BeatIt if the ICER is below £20,000, however the ICER is much larger than £20,000; it is £496,160. The 95% confidence interval for the difference in costs and QALYs in both analyses crosses zero, indicating uncertainty in both of the results. The difference in costs is larger in the conceptual model driven analysis but the 95% confidence interval is narrower by a factor of nearly two, illustrating that uncertainty in the point estimate has decreased despite BeatIt being more costly than StepUp. The difference in QALYs between arms is smaller in the conceptual model driven analysis compared to the treatment arm-based analysis

and the 95% CI is narrower by a factor of almost eight, again confirming that uncertainty in the point estimate has decreased.

The less uncertain results from the conceptual model driven analysis are clearly illustrated in Figure 67, where the cost-effectiveness planes for both approaches are compared. The spread of bootstrapped samples is wider in the treatment arm-based analysis (top) and they are relatively evenly spread in all four quadrants of the cost-effectiveness plane, the confidence ellipse is also crossing into all four quadrants confirming the uncertainty in these results. In the conceptual model driven analysis (bottom) the samples are more focussed around the origin. In this analysis the costs show a broader spread north of the horizontal axis and less spread south of the horizontal axis. There are more bootstrapped samples in the northeast quadrant compared to the southwest quadrant, showing a stronger trend towards BeatIt being more expensive and more effective than StepUp. The costs and QALYs have a narrower spread in the conceptual model driven analysis compared to the treatment arm-based analysis indicating less uncertainty in the difference, however the 95% confidence interval for differences still crosses zero in the conceptual model driven analysis. Furthermore, despite a narrower spread of samples the confidence ellipse still crosses into all four quadrants.

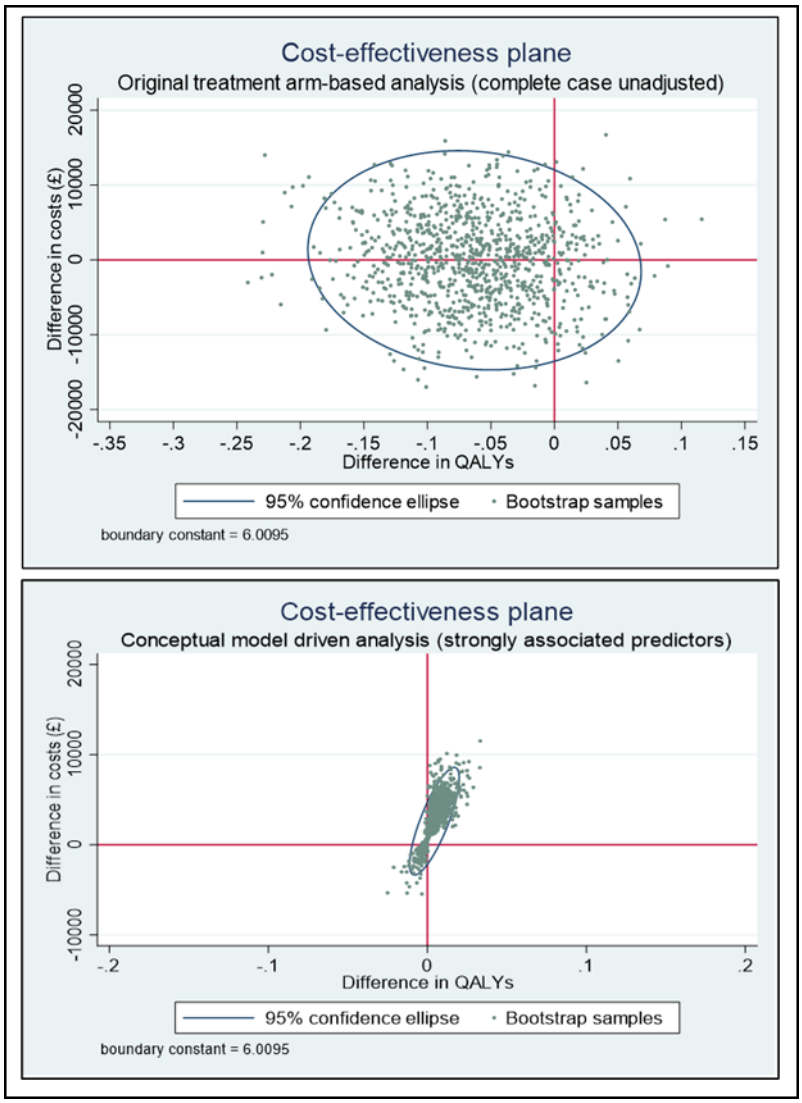


Figure 67: Cost-effectiveness planes - comparison of treatment arm-based and conceptual model driven analyses

The cost-effectiveness acceptability curves from both analyses are presented in Figure 68 for comparison. The treatment arm-based analysis curve (top) shows that at a willingness-to-pay threshold of £20,000 the BeatIt therapy has less than a 50% chance of being cost-effective compared to the StepUp therapy, however in the conceptual model driven analysis (bottom) the chance of BeatIt being cost-effective compared to StepUp is approximately 12%. The treatment arm-based analysis cost-effectiveness curve shows that the probability of BeatIt being cost-effective compared to StepUp is fairly constant at all willingness-to-pay thresholds; there is little chance of BeatIt being cost-effective. In the conceptual model driven analysis there is also an almost constant probability of BeatIt being cost-effective compared to StepUp, this is lower than the probability in the treatment arm-based analysis. This result confirms the

interpretation of the cost-effectiveness planes, there is little chance of BeatIt being cost-effective in the treatment arm-based analysis, this is confirmed in the conceptual model driven analysis with less chance of BeatIt being cost-effective.

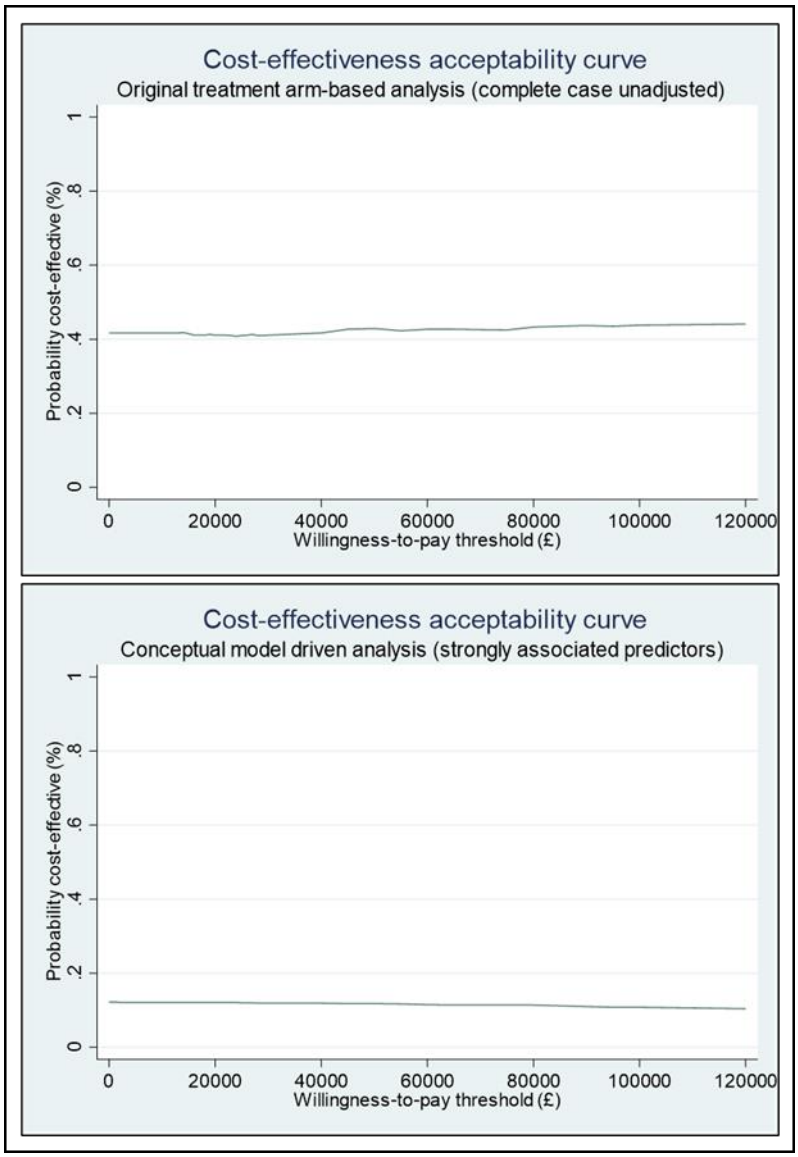


Figure 68: BeatIt cost-effectiveness acceptability curves - comparison of treatment arm-based and conceptual model driven analyses

The incremental net monetary benefit measure is particularly helpful for interpreting results when bootstraps fall into more than one quadrant of the cost-effectiveness plane, as in this situation. The incremental net monetary benefit for both analyses is negative, indicating that decision makers should not adopt the BeatIt intervention in either analyses. The negative incremental net monetary benefit from the conceptual model driven analysis is greater than the

net monetary benefit resulting from the treatment arm-based analysis indicating that this approach less cost-effective. The uncertainty (95% CI) around the net monetary benefit measures is over two times bigger in the treatment arm-based analysis, again showing that the conceptual model driven analysis results has less uncertainty around the results, but these results still confirm that BeatIt is not cost-effective. The incremental net monetary benefit results from both approaches are plotted in a histogram to visually illustrate the variation of the results in each approach in Figure 69. This demonstrates the narrower spread of results from the conceptual model approach compared to the trial-arm based results and the lower incremental net monetary benefit point estimate, confirming the previous results that BeatIt would be considered less cost-effective in the conceptual model driven analysis compared to the treatment arm-based analysis.

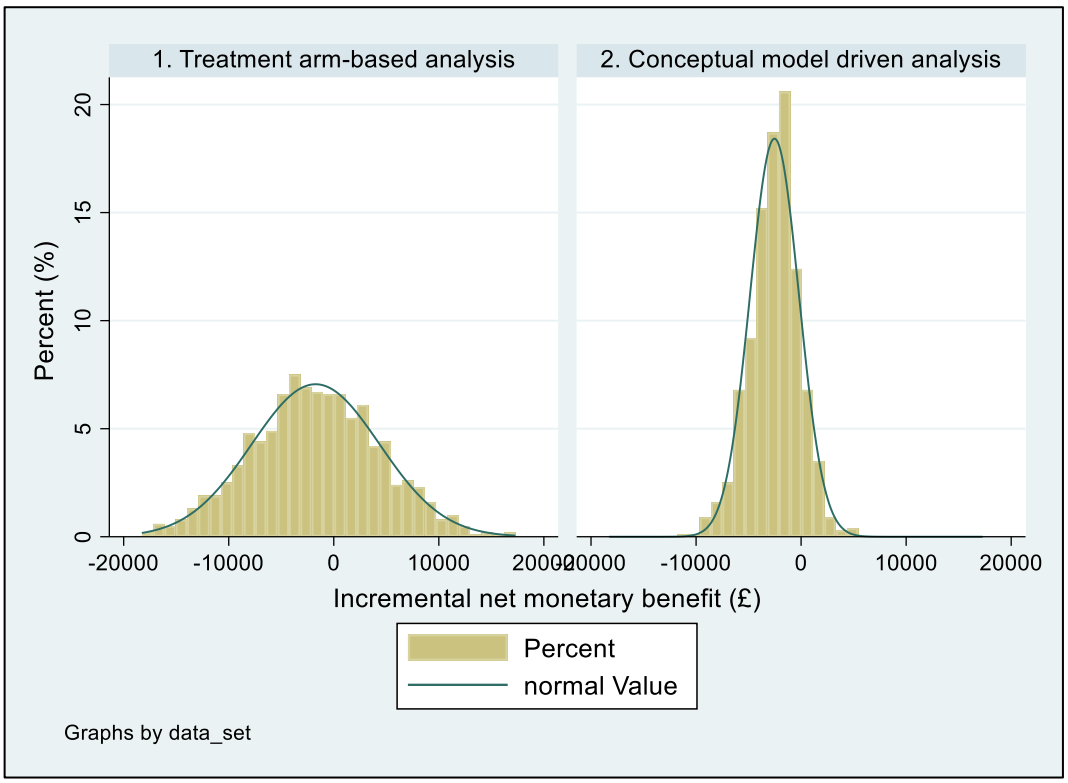


Figure 69: BeatIt - histogram of net incremental monetary benefit results showing the distribution of the results from both approaches

Value of information results per person are also included in Table 25 to explore the value of eliminating sampling uncertainty and having perfect information. These results show that there is more value to be gained by doing more research

in the treatment arm-based analysis compared to the conceptual model driven analysis (£1,679 v £148), the lower EVPI estimate for the conceptual model driven analysis suggests that this analysis provides a more detailed understanding than the treatment arm-based analysis as a result of the additional structure in the conceptual model. EVPPI results indicate there is more decision uncertainty in total costs, and potential value in further research on total costs in both analyses compared to further research on QALYs. In the treatment arm-based analysis the value of more research per person is £21 for QALYs and £1,570 for total costs. In the conceptual model driven analysis, the value of more research per person is £0 for QALYs and £171 for total costs. These results confirm that there is more decision uncertainty in the treatment arm-based analysis. When the EVPPI for total costs is broken down into treatment and non-treatment costs in the conceptual model driven analysis, the results show that there is no value to be gained by reducing the uncertainty in treatment costs (EVPPI results are £0); this result is expected as treatment costs are known and precise, there is no decision uncertainty. For non-treatment costs the EVPPI is £168 indicating a level of decision uncertainty, and that the value of further research into these costs is greater than further research into treatment costs.

7.5.4 BeatIt summary

In the original economic evaluation, the treatment arm-based results for the complete case analysis reported higher costs and lower QALYs for the BeatIt arm compared to the StepUp arm, however there was uncertainty in these results as bootstrapped samples fell into all four quadrants, as did the confidence ellipse. These treatment arm-based results were confirmed in the multiple imputed unadjusted results; however, the multiple imputed adjusted results are reversed with lower costs and higher QALYs in the BeatIt arm compared to the StepUp arm. These results showed there was uncertainty in the economic evaluation, the only certain element was that treatment costs were greater in the BeatIt arm compared to the StepUp arm. In the conceptual model driven analysis costs and QALYs were higher in the BeatIt arm compared to the StepUp arm, these results were also uncertain with the confidence ellipse crossing all four quadrants of the cost-effectiveness plane, however the 95% confidence interval was narrower in this analysis compared to the original treatment arm-based

analysis, showing the greater precision of the conceptual model driven analysis results.

The conceptual model driven approach demonstrated that the treatment arm predicted treatment costs, but not depression, activity nor non-treatment costs. Levels of activity predicted non-treatment costs and depression but not quality of life, and depressive symptoms predicted quality of life but not non-treatment costs. In the original treatment arm-based analysis the quality of life (in terms of QALYs) improved in both arms, the conceptual model driven analysis can give a more detailed understanding of the mechanism driving this; it showed that there was no difference in activity levels between arms, but that there was a link between activity and depression, and depression and quality of life demonstrating that both therapies affected depressive symptoms and quality of life. This result validated including the link between activities, depressive symptoms and quality of life in the conceptual model.

None of the possible associations were found to be strong and they were all omitted from the conceptual model final analysis. One assumed association was omitted in the final analysis: depressive symptoms predicting non-treatment costs, this may be a result of the increase in activity costs linked to increased activity levels masking this association. The conceptual model driven analysis confirmed the original treatment arm-based analysis findings that treatment arm only predicted treatment cost, not depressive symptoms nor quality of life. Additionally, it found that the BeatIt intervention did not alter activity levels significantly differently compared to the StepUp comparator. These findings could feed into future trials with knowledge that the GDS-LD is linked to the EQ-5D-Y outputs and that activity levels predict depressive symptoms. However, it should be borne in mind that the sample in this trial was small and that any conclusions reached should be treated with caution.

The incremental net monetary benefit showed that neither analysis considered BeatIt to be cost-effective. Value of information results showed that additional research was valued higher in the treatment arm-based analysis compared to the conceptual model driven analysis, and that the less uncertain estimates in the latter analysis indicated that there would be no value to be gained by further research into treatment costs and QALYs.

Overall, the conceptual model driven analysis demonstrated that BeatIt is not a cost-effective therapy compared to StepUp.

7.6 Discussion of new role for conceptual models

7.6.1 Overview

The purpose of the new role for conceptual models is to provide decision makers with additional insight into the results of conventional treatment arm-based analyses, going beyond an intervention focussed, outcomes-driven analysis to provide understanding of the trial mechanism driving the economic evaluation results.

This new role involves developing a conceptual model to represent the trial mechanism linking inputs to economic outputs via expected associations. The approach comprises three aspects: 1) develop a conceptual model which can be used as a **communication tool** to explain the economic evaluation to the clinical trial team, and confirm assumptions made and the clinical integrity of the conceptual model: 2) **regressions** are performed to confirm or otherwise the accuracy of the mechanism in the conceptual model, this provides additional understanding to the treatment arm-based results, and 3) provides a framework on which to carry out further **analysis** to assess point estimates and produce cost-effectiveness planes, cost-effectiveness acceptability curves, net monetary benefit, and value of information estimates. These outputs are compared to the original treatment arm-based results to provide a more nuanced understanding of the treatment arm-based results.

The case studies were used to evaluate the new role and establish whether it achieved its purpose of providing additional insight and understanding.

In both case studies the results from confirming the accuracy of the conceptual model provided additional understanding of the treatment arm-based results, for example the TWICS case study confirmed the validity of including exacerbations as a mediator as the number of exacerbations was found to predict quality of life and non-treatment costs. In BeatIt the results confirmed that levels of activity predicted depressive symptoms, which in turn predicted quality of life.

Neither case study showed that the treatment arm directly predicted quality of life or non-treatment costs, this illustrates the value of this new role when most economic evaluations are conducted based solely on the difference in economic outcomes between arms. The results from understanding the trial mechanism could inform the design of future trials in terms of resource use categories collected and evidence of associations in trial mechanisms.

Applying the case study did not reverse or significantly change the original trial results but did provide a more less uncertain estimate of cost-effectiveness, witnessed by narrower variances in uncertainty measured by 95% confidence intervals and presented on the cost-effectiveness planes with confidence ellipses. Both case study cost-effectiveness acceptability curves showed a lower probability of the intervention being cost-effective, and overall, the conceptual model driven analyses reported lower EVPI estimates than the original treatment arm-based analyses suggesting less value to be gained by future research based on the conceptual model driven results, it could be argued that the additional structure of the conceptual model driven analysis has added value without the need for future research.

Despite the similarity in results between the two analyses in the case studies there may be a times where the treatment arm-based analysis produces conflicting results to the conceptual model driven analysis, in this scenario the new role will not help interpret the original results but will add confusion. These conflicting results may result from an inaccurate conceptual model, in which case the accuracy of the conceptual model should be checked, however if the conflicting results are genuine then this should be presented in a transparent way.

The regressions in the case studies were kept deliberately straightforward as their purpose was to illustrate the new concept. Ordinary least squares regressions were chosen as they are simple to interpret, easy to understand, are unbiased, and have less assumptions than other regression techniques. Whilst I acknowledge that cost and count data is often skewed and other regression techniques are better suited to account for this skewedness, the focus of the case studies was to illustrate the new conceptual model driven analysis approach without the distraction of statistical model selection. For this same reason

covariates were not included in the regressions, nor multiple imputation in the dataset (except a straightforward replacement of missing data with treatment arm specific means).

To prevent bias in the conceptual model it should be developed prior to receiving the clinical trial data, if the clinical trial results produce unexpected outcomes sensitivity analyses can be carried out. For example, I used the TWICS conceptual model to further analyse the clinical trial data when results showed that there were more participants reporting exacerbations requiring hospital treatment in the placebo arm. However, there will be a point when additional sensitivity analysis turns into a data fishing exercise; the analysts perform many unplanned sensitivity analyses with the aim of discovering a significant finding from the data. Which leads to the question; 'when do you stop this analysis?' The conceptual model sets out a priori the expected associations in the economic evaluation, the first level of analysis confirms the accuracy of the conceptual models, sensitivity analysis (such as in TWICS) comprises the second level of analysis, I would argue that two levels are enough to gain sufficient understanding of the trial data in this new approach and stops the analysis becomes a data fishing exercise.

Finally, it is important to be clear about the key assumptions made in developing the conceptual model, as discussed in the methodological framework in Chapter 5, so the results can be interpreted correctly.

I believe that the new role achieved its purpose; it has the potential to provide additional insight and more detailed information for decision makers, particularly when the original treatment arm-based economic analysis is uncertain, however the new role should not be seen as limited only in providing further insight into uncertain results. The conceptual model driven analysis should be seen as a supplement to the treatment arm-based analysis, adding value to it.

Strengths and limitations of this novel approach were highlighted in applying it to the case studies, these are discussed in the next sections.

7.6.2 Strengths

This is a new role for conceptual models in the field of economic evaluation, aside informing the structure of a decision analytic model; it provides a novel approach in analysing valuable clinical trial data, going beyond the conventional treatment arm-based analysis where the focus is on the difference in costs and health benefits between treatment arms as measured by the ICER, with no regard to what is driving the costs and health benefits. The new role also considers the trial mechanism driving the results, providing understanding into what is driving the results. It could be argued that this new role brings together elements of treatment arm-based economic evaluation and a decision model based economic evaluation by considering the underlying processes driving the trial mechanism.

The conceptual model driven analysis addresses a sub-set of limitations that result from conducting economic evaluations alongside clinical trials, these include insufficient sample size, suboptimal collection of economic data and protocol-driven costs. However, as mentioned earlier, this new role is not limited to uncertain clinical trial results.

Conducting clinical trials and collecting data for analysis is expensive, time consuming and uses human resources in terms of the trial team and participants, by using the data for additional analysis these valuable resources are further utilised.

In sub-section 2.5.1 approaches that should be used to reduce the burden of resource use collection were discussed, one of these was limiting the resource use categories to those directly affected by the disease or intervention. Developing these conceptual models can help communicate the rationale behind omitting some resource use costs whilst including others.

The case studies showed that more detailed understanding provided by the conceptual model driven analysis could inform the design of future clinical trials, both in terms of identifying resource use categories and in terms of giving understanding into trial mechanisms.

7.6.3 Limitations

The methods described in the case studies are illustrative only and should not be interpreted as the most robust and valid method to conduct conceptual model analyses. The method was chosen because it is easy to understand and simple to apply in this role.

Due to the nature of conceptual models the results are based on a simplification of the trial mechanism, as such identified mediators may be overstated. As the approach uses the clinical trial data there may be mediators that are not measured (unobserved), thereby overemphasising the mediators included in the conceptual model, this may produce spurious results due to underlying mechanisms not being identified and tested. This approach may work better in a clinical trial with a straightforward mechanism, more complex mechanisms may include mediators which are not collected in the clinical trial data leading to issues with unmeasured components.

When identifying the regressions to include in the bootstrapped analysis the regression results showing strong associations should be included, however sometimes only one component is predicting another (for example in the TWICS case study treatment arm linked to exacerbations, and in the BeatIt case study treatment arm linked to activities); if the regression does not show a strong association between these components the regression still needs to be included to facilitate the full conceptual model driven analysis. This is likely to be the case in many conceptual models developed for this new role as the first component will nearly always be the treatment arm and this is likely to only feed into one main component in the trial mechanism, which in turn may drive the rest of the trial mechanism.

It could also be argued that reliance on statistical significance when interpreting the regression results is counterintuitive when many economic evaluations are underpowered due to an insufficient sample size for economic analysis. Due to potential lack of power, there is a possibility that one or more pathways in the conceptual models were rejected when they should have been included. However, the purpose of the regressions is not to conduct a conventional statistical hypothesis test, but as a filter to confirm the logical hypothesis of the

conceptual models. To provide additional evidence to decision makers, I believe this potential trade-off is acceptable.

By following clinical opinion there was a potential for bias in the final conceptual model analysis results for TWICS, caused by omitting the link between treatment arm and non-treatment costs in the conceptual model. The conceptual model found that treatment was a strong predictor of non-treatment costs, however this link was omitted in the final conceptual model analysis. There was no a priori hypothesis that theophylline would have an effect on the severity of exacerbations and therefore the suitable setting for treatment (home or hospital), which is why a link between treatment arm and type of exacerbation cost was not illustrated in the proposed conceptual model (Figure 54). However, when the accuracy of the conceptual model was tested there was strong link between treatment arm and non-treatment costs. Further sensitivity analysis found that this link was driven by the cost of exacerbations treated in hospital – there were more participants with hospitalised exacerbations in the placebo arm than the theophylline arm, and 10 of these participants had lengths of stay greater than 40 days. Clinicians from the TWICS trial put this down to a chance finding; there was no plausible biological mechanism for this, and as the trial was multi-site there may be differing criteria for admitting participants to hospital for treatment compared to treating at home in different sites. Because of this information the link between treatment arm and non-treatment costs was omitted from the final conceptual model analysis. However, this assumption/omission may have resulted in the final conceptual model analysis results not fully reflecting the trial mechanism and being biased, limiting the usefulness of the resulting analysis in the case study. This also highlights a potential limitation of the proposed new role for conceptual models; when testing the conceptual model accuracy results in a conflict between regression results and clinician input. In this situation the modeller would need to decide whether to omit links based on clinical opinion, with resulting potential bias to the final results.

There are limitations in the case studies due to simplifying the new approach to suit an illustrative case study; complete case data has been used without any multiple imputation (only a naive method of replacing missing data with

treatment arm specific means). Furthermore, no adjustments for imbalances were performed, this was to keep the case studies simple, for example BeatIt had an imbalance in quality of life at baseline and omitting this adjustment in the conceptual model driven case study may have affected the results, but I did not want to introduce the complexity of choosing statistical methods when demonstrating the new role.

Whilst the regression method, based on p-values, chosen to illustrate the new role in the case studies was straightforward, it was not intended to be a prescriptive method, there are alternative methods that could be used in this role, and these should be explored in future work. These methods relate to causal inference and include structural equation models (SEM), path analysis and directed acyclic graphs (DAGs). SEM entails numerous linear equations being used to identify causal relationships in observed and unobserved variables(198). Path analysis is similar to SEM, but does not consider unobserved variables, it was developed by Sewall Wright to test whether data is consistent with a hypothesised model (106). The path analysis method is based on hypothesised nested causal relationships in a system which are represented by linear regressions. Data is assessed to identify any effects on one variable being caused by another, testing the hypotheses in the model(199). This format is similar to DAGs, these are mathematical models which represent causal relationships. When developing DAGs and identifying possible causal relationships there are constraints on how they are developed, adding to the complexity of this method(108). The main difference between path analysis and DAGs is that in path analysis the models predetermine linear causal effects, whereas in DAGs the models may be linear or non-linear (108). The regression method chosen for these case studies shows correlation – a linear relationship between two variables, whereas causal inference tests hypothesised links. A step away from the causality methods described above is stepwise variable selection, this method is model driven and does not include any logic imposed on the model, focussing solely on which predictors have the biggest effect on dependent variables(199).

I assessed the appropriateness of using SEM but rejected it as it added an additional layer of complexity to include and consider unobserved variables. A

fundamental concept in this new role is that data from the clinical trial is analysed further by conceptual models, if the data is not available because it is unobserved this would change the focus of the concept. I also applied path analysis methods to the case study data but found model specification was difficult to achieve. The language/terminology, software and assumptions inherent in SEM, path analysis and DAGs were also a barrier for their use as an illustrative method for the case studies, adding extra complexity to the case studies, the purpose of which is to illustrate the new role, not to set prescriptive methods for the new role.

Moving away from causality hypothesis testing and the p-value some have championed using effect size and uncertainty (95% confidence intervals) as a means to measure effects(200). However, for the purpose of the new role decisions would still need to be made on which links in the conceptual model to reject or accept, meaning an acceptable method of interpreting effect size and uncertainty would need to be considered. Another alternative to establishing causality would be to impose a minimum clinically important level as a 'cut-off' for acceptance/rejection, however, again this would need to be considered and data is unlikely to be available for all clinical outcomes.

In summary, in this proposed new role for conceptual models, they are used to illustrate causal links, similar to DAGs, however the relationships are tested using different methods of analysis. Future research should identify differences and overlaps between the methods used in the illustrative examples and SEM (or other methods); however it is unlikely that additional statistically significant results would be found. Furthermore, neither SEM, path analysis nor DAGs are routinely used in economic evaluation and by illustrating the new role for conceptual models using regressions I am not proposing that they are replaced, it may be proven in future work that they are the most valid statistical method, however, as they are not routinely used in economic evaluation, regressions are a simple and widely understood concept to use for illustration purposes.

Variables from the trial data used to represent the components in the conceptual model may not be suitable for that purpose, depending on the variables collected and the purpose of collecting them. For example, utilities are driven by how a person feels and their capabilities, this may not always

relate directly to the disease and changes in the disease, therefore using health utilities can be a blunt instrument in terms of evaluating the expected trial mechanism.

It was not possible to compute the EVPPI for components within the trial mechanism, only the outputs of costs and QALYs, this was because of how the bootstrapping was operationalised, further work needs to be done to identify a suitable mathematical method to extract this data. Furthermore, whilst EVPI was estimated EVSI and ENBS were not, this limits the conclusions on whether future research is worthwhile.

7.7 Summary

This chapter demonstrated the proposed new role for conceptual models to add further insight and understanding into the clinical trial data, this is the third output in this thesis and fulfils the first objective. The case studies were used to evaluate the new role for conceptual models, and highlight the strengths and limitations experienced in applying the new role to the case studies.

The case studies demonstrated that using this conceptual model driven analysis produces additional insight into the valuable clinical trial data, evaluating and confirming which key components of the conceptual model are driving the trial mechanism in relation to the economic evaluation and comparing the original conventional treatment arm-based analysis to the new conceptual model driven analysis.

The inclusion of the case studies was intended to illustrate the concept of the new role for conceptual model and not be an exhaustive presentation of the methods of analysis, therefore the results of the conceptual model driven analysis should be interpreted with caution in terms of understanding the TWICS and BeatIt analyses.

Further research to develop this new approach could include using more case studies to evaluate the new role, in particular in different disease areas, care pathways and in clinical trials reporting certain results. The conceptual model and variables from the clinical trial data could be tested further using path

analysis, and possibly structural equation modelling if latent variables are relevant to the particular trial mechanism. Plus, methods could be investigated to provide complete EVPPI results.

The next chapter is the final chapter of the thesis and presents a discussion and summary of the thesis.

Chapter 8 Discussion and conclusions

8.1 Chapter overview

The aim of this chapter is to discuss the work in the thesis and to provide a conclusion, answering the research question 'how can conceptual modelling enhance health economic evaluation?'

The aim of this thesis was to expand the role of conceptual modelling in health economic evaluation. The two objectives of the thesis were to propose and demonstrate a new role for conceptual models in health economic evaluations, and to propose and demonstrate a methodological framework for developing conceptual models in this new role.

Currently the role of conceptual models in economic evaluation is restricted to a pre-cursor for a decision analytic or mathematical model; these models are routinely used to extract and amalgamate evidence from different sources to estimate cost-effectiveness. This thesis introduces a new role for conceptual models in economic evaluation; to provide decision makers with additional understanding and insight into clinical trial data. Clinical trials are expensive to run and the data collected in them should be explored fully, currently there is robust guidance for carrying out economic evaluations using clinical trial data, however, inherent problems with sample size and conflicting objectives for clinical and cost-effectiveness analyses often lead to uncertain results in economic evaluations alongside clinical trials. A standard treatment arm-based analysis also overlooks important aspects of the trial mechanism relating to costs and quality of life focussing solely on the ICER and not considering what is driving it, in an outcomes driven 'black box' evaluations. These drawbacks provide scope for a new approach to analyse valuable clinical trial data, and thereby giving additional understanding to decision makers.

The contributions of this thesis are summarised by chapter in Section 8.2, the strengths and limitations of the thesis are then discussed in Section 8.3, implications for policy and practice are presented in Section 8.4, recommendations for further work are suggested in Section 8.5, and finally, conclusions are made in Section 8.6.

Chapter 8

8.2 Overview of chapters

8.2.1 Chapter 2 – The role and practices of economic evaluation alongside clinical trials

This chapter considered the role and purpose of economic evaluations; budgets for allocating healthcare services are limited and decision makers need to decide which new and existing health technologies to fund, based on safety, efficacy, effectiveness and cost-effectiveness evidence, the gold standard vehicle for basing these decisions on are clinical trials. Clinical trials follow a pre-determined protocol for determining sample size, recruiting participants, delivering the health technology and establishing the effectiveness of the technology (the treatment effect). Established guidance exists for conducting economic evaluations alongside clinical trials, however there are drawbacks for using clinical trial data for this purpose, these stem from differences in the main objectives and audiences of the clinical effectiveness analysis of a clinical trial and the economic analysis of a clinical trial. Drawbacks resulting from the differences in objectives include lack of generalisability, shortened time horizon, limited comparators and lack of power for economic outcomes. The main consequence of these drawbacks is uncertainty. Guidance exists on how to express and describe uncertainty in economic evaluations; however decision makers are often left with limited evidence on which to base a decision. There is also criticism of the 'black box' nature of economic evaluations, with the analysis driven by treatment arms with no consideration of the mechanisms and associations linking the treatments to the outcomes. This chapter highlighted that although there is well established guidance for conducting economic evaluations alongside clinical trials several potential drawbacks in this field remain. A new role for conceptual models could provide additional information to decision makers, adding understanding of the trial mechanism process and giving an interpretation of the economic evaluation components that goes beyond the conventional focus on the ICER. An overview of this new role was introduced, then a description of the methods in the new role was given: 1) the key economic components of the economic evaluation are identified, helped by a template based on the Weinstein and Stason equations, the trial mechanism is identified and simplified, and the links and associations between the key economic components and trial mechanism are depicted in a conceptual model

diagram; 2) the conceptual model diagram is tested for accuracy using regression techniques, this provides information on components and associations in the conceptual model giving further understanding of the key drivers of the trial mechanism driving economic evaluation results, 3) the confirmed conceptual model is then used as a framework to conduct an additional analysis on the valuable trial data, the results of this additional analysis are compared to the original treatment arm-based analysis results giving further insight and understanding. This chapter provided the reader with an understanding of the role of economic evaluation, guidance used to perform economic evaluations, drawbacks of conducting economic evaluation, and an introduction to a new role for conceptual models that attempts to address the drawbacks and provide further understanding.

8.2.2 Chapter 3 – Case studies: an illustration of current economic evaluation guidance

Chapter 3 introduced the two case studies used for illustration purposes in this thesis. Both cases studies were NIHR funded clinical trials: 1) TWICS investigated the clinical and cost-effectiveness of including theophylline alongside usual care compared to usual care only in patient with COPD to reduce the number of exacerbations needing treatment, and 2) BeatIt investigated the clinical and cost-effectiveness of a behavioural activation therapy (BeatIt) compared to self-help therapy (StepUp) for adults with intellectual disabilities and depression, to reduce depressive symptoms. In both case studies the established methods for conducting economic evaluations alongside clinical trials described in Chapter 2 were applied and the results of a conventional treatment arm-based analysis were presented. Both economic evaluations showed uncertain results, and the uncertainty in the results was illustrated using recommended approaches. By presenting these case studies the scene was set for introducing the proposed novel approach, based on representing the anticipated causality in the trial mechanism (10) using conceptual models. This chapter demonstrated the application of the guidance presented in Chapter 2 and the uncertainty experienced in many economic evaluations.

8.2.3 Chapter 4 – Scoping review of methodological framework development

The second objective of this thesis was to provide guidance for developing conceptual models for the new role in economic evaluation. Methodological frameworks provide a step-wise structured guide for a process, suitable for this objective. However, at the outset of the thesis research there was no guidance on how to develop methodological frameworks, so a new work strand was added to this second objective; to map the existing approaches taken in developing methodological frameworks and amalgamate and summarise these approaches into suggestions for developing methodological frameworks. A scoping review identified 30 methodological frameworks which reported the approaches taken in developing the included methodological frameworks. The approaches taken were extracted and synthesised into phases to make suggestions for developing methodological frameworks. Chapter 4 demonstrated that there are enough similarities in current reported approaches to developing methodological frameworks to group and produce suggestions based on the approaches. This chapter was the first output of the thesis, and partly fulfilled the second objective.

8.2.4 Chapter 5 – A methodological framework for conceptual models in economic evaluation

In this chapter the suggestions for developing a methodological framework from Chapter 4 were used to develop a draft methodological framework for designing conceptual models. The existing guidance for developing conceptual models in economic evaluations focusses on using the conceptual model to design the structure of a decision analytic model, however the aim of this thesis was to explore using conceptual models for guiding trial data analysis, so new guidance was needed. A scoping review identified 16 existing methodological frameworks for developing conceptual models, 18 steps for the process were extracted, these were grouped into phases and amalgamated into stages within the phases. Although the identified methodological frameworks were from seven fields there were enough similarities to create a draft methodological framework, this was then evaluated by comparing it to existing guidance for developing conceptual models in economic evaluation. The draft methodological frameworks was

amended for any gaps or issues highlighted by the evaluation exercise to produce a final methodological framework. This chapter demonstrated that it was possible to produce a methodological framework for developing conceptual models based on reported approaches for development in fields other than economic evaluation; there were similarities in all fields. Also, by applying economic evaluation specific concepts the methodological framework was similar to established conceptual model development guidance. This chapter formed the second output from the thesis and fulfilled the second objective.

8.2.5 Chapter 6 – Applying the methodological framework to developing conceptual models

This chapter applied the methodological framework to the thesis case studies to develop conceptual models to use in the new role. In applying the methodological framework each stage of the methodological framework was described, and the iterations of each conceptual model were presented with a description of how each conceptual model was developed. The template based on the Weinstein and Stason was used to identify the key economic components, these were combined with the trial mechanism to produce the final conceptual models. This chapter demonstrated that following the stages in the methodological framework enables the user to develop a conceptual model, the stages and guidance in the methodological framework were helpful and clear, and several iterations of the diagram were needed before the final conceptual models were developed.

8.2.6 Chapter 7 – Case studies to illustrate the new role for conceptual models

In this chapter the two case studies used throughout this thesis demonstrated the new role for conceptual models. The conceptual model driven analysis delves into the trial data to further understand the trial mechanism beyond a conventional treatment arm-based analysis. The logic of the conceptual models was assessed using regression techniques, this aspect of the new role provided additional understanding in terms of what was driving the economic results, giving insight into the current trial and potentially informing the design of future trials. The results from the regressions fed into the conceptual model driven analysis, results from the analyses showed less uncertainty in the conceptual

model driven analyses compared to the treatment arm-based analysis, overall the conclusions of the original treatment arm-based analysis results did not change but additional understanding was provided. The conceptual model highlighted the key mechanisms within the clinical trial generating costs and health benefits. Decision makers can use the results of the conceptual model driven analysis to better understand clinical trial results and consider the value of future research and how it might incorporate evidence identified in the conceptual model driven analysis. This chapter demonstrated the new concept and illustrated how it could be applied in practice, highlighting its strengths and limitations. This fresh approach for using conceptual models in economic evaluation is the final output in the thesis and fulfils the first objective.

8.3 Strengths and limitations of this research

This section discusses the strengths and limitations of the research, taking one output at a time: 1) the first major methodological piece of work was the scoping review of current reported approaches for developing methodological frameworks; 2) the second major methodological piece of work was presenting a methodological framework for developing conceptual models, and 3) the last output was introducing the new role for conceptual models in economic evaluations.

Strengths

Starting with the first major methodological piece of work, this research identified a gap in the evidence; at the outset of this thesis there was no guidance available for developing methodological frameworks. This piece of novel work addressed this gap by identifying current reported approaches for developing methodological frameworks and grouping and amalgamating them into similar themes to produce suggestions for developing methodological frameworks. Evidence in the scoping review was taken from different fields which naturally leads to a variation in approaches, despite this there were enough similarities in the reported approaches to provide suggestions, and the natural variations resulted in robust and generalisable suggestions. Furthermore, this scoping review identified a lack of consistency in terminology used to describe methodological frameworks and made suggestions for the use of

consistent terminology to aid identification of methodological framework in literature searches in the future. This piece of work has been published in a peer-reviewed journal.

The second major methodological component of this thesis, the development of the methodological framework, was developed using robust suggestions from the scoping review above. The methodological framework for developing conceptual models is the first in the field of economic evaluation that is not for the purpose of developing the structure of a decision analytic model. The evidence for informing the development of the methodological framework came from outwith economic evaluation; it was based on evidence from several fields and for several purposes, there was enough homogenous themes to group into stages and produce a robust methodological framework. The additional step in developing the methodological framework of comparing the draft version to existing conceptual modelling guidance in economic evaluation added an evaluation, however the purpose of my methodological framework and the existing guidance differs. The final methodological framework was successfully applied to the two case studies for demonstration.

The final contribution of the thesis was the new role of the conceptual model, introducing a fresh approach. A strength of this contribution was to provide additional information and evidence to decision makers from valuable clinical trial data, there is also potential for the results from the conceptual model driven analysis to inform future clinical trials. The additional evidence provided by the new role comprises three strands: 1) a communication tool to illustrate (and receive agreement of the associations linking the trial mechanism and economic evaluation components) with interested parties, and explain the rationale behind economic evaluation decisions, for example omitting or including specific resource use categories; 2) confirmation of the underlying trial mechanism presented in the conceptual model using regression techniques, providing understanding of the trial mechanism driving the economic evaluation results, and 3) additional analysis based on the confirmed trial mechanism, when this is compared to the original treatment arm-based analysis additional understanding is provided. Often the only certain results from the treatment arm-based analysis are the known and precise treatment costs, the conceptual

model acts as a filter to remove noise from the original results and adding value as a supplementary analysis. Another method of reducing this noise is to limit to the breadth of the economic data collection at the planning stage of the economic evaluation, by focussing on resources and outcomes likely to be affected by the treatment, this also has the benefit of lessening the burden on participants. However, it is important not to make too many assumptions at this early stage to allow potential genuine differences between arms to be identified in the data. Using a conceptual model at this earlier stage can help focus data collection on resources likely to be affected, whereas the later conceptual model analysis can help filter out potential noise in the data. This conceptual model driven analysis has the potential to address a sub-set of limitations that result from conducting economic evaluations alongside clinical trials; insufficient sample size, suboptimal collection of economic data and protocol-driven costs. The additional information goes beyond the conventional treatment arm-based analysis, where the analyst focusses on the end product of the ICER without considering the trial mechanism, to consider the underlying mechanism driving the clinical trial and economic evaluation.

Limitations

Starting with the scoping review of reported approaches in methodological frameworks, whilst a scoping review was the correct method for a first step in identifying current reported approaches in an area with scarce evidence, the resulting suggestions produced can only be described as suggestions, not recommendations to inform practice, for this a full systematic review would be needed with external validation. A full systematic outwith the scope of this thesis due to time and resource constraints. However, the scoping review has been published in a peer reviewed journal and the interest in the paper shows there is an appetite for this information. The search terms of the scoping review were restricted to make the search manageable and pragmatic for the timescale of the thesis, these may have led to the exclusion of some relevant methodological frameworks. The lack of consistent terminology may also have led to the exclusion of relevant evidence. Whilst published guidance for grey literature searches was followed there are limitations in searching the internet, these are the transient nature of the internet and the personalised nature of

internet searches making replication difficult. A fifth of all identified methodological frameworks were rejected because they did not report the methods used in their development, or approaches were not clear, limiting the amount of evidence I could include in the suggestions. Finally, as the purposes of the methodological frameworks were varied and none were for developing conceptual models, it is possible that the suggestions extracted were not all suitable for methodological frameworks for developing conceptual models.

For the second output of the methodological framework, similar to the scoping review above, the search terms were limited to make the search pragmatic, therefore some evidence may have been missed, plus the internet search will have been subject to the same limitations as described above. Particularly evident were the exclusion of logic models, which are very similar to the purpose of conceptual models proposed in this thesis. Both can be used a priori to identify suitable outcome measures, provide transparency (184), and are visual representations of expected outcomes of an intervention in the system they depict. Extracting the themes and amalgamating them into stages was iterative and based partially on personal opinion making the methodological framework subject to personal interpretation. The methodological framework was applied to the two case studies as an illustration, however the case studies do not represent the full breadth of circumstances and possibilities of all economic evaluations, so it is possible that the methodological framework may not be suited to all economic evaluation situations. Due to time constraints, it was not possible to externally validate the methodological framework, which may have enhanced it.

Limitations of the new role for conceptual models include no opportunity to present the results of the conceptual model driven analyses case studies to the relevant clinical trial teams. Had this been possible there would be three possible benefits: 1) more complete understanding of the economic evaluation results for the clinical trial team, 2) feedback from the clinical trial teams might have enhanced interpretation of the conceptual model driven analyses for further refinement of the proposed application, and 3) their feedback might have fed back into the methodological framework. The nature of conceptual models is a simplification of a real-life system or process, therefore it is possible

that some important mediators may be missed, and/or identified mediators overstated. Data for some components may not be available from the clinical trial data, therefore limiting the analysis. The purpose of this third output of the thesis was to propose a new role for conceptual models, part of this demonstration were examples of possible analyses, as these are illustrative only they have been kept simple so may not stand up to mathematical rigour. The purpose of these case studies is not to rigidly prescribe methods or definitive guidance, but to introduce the concept of conceptual model analysis, highlight key issues and stimulate debate and conversation around applying the new role for conceptual models. The methods used in the case studies were kept simple in order to illustrate the new role clearly. The TWICS case study highlighted a limitation of the new role; there is potential for conflict between the results when testing the accuracy of the conceptual model and clinical opinion/biological plausibility, and potential biases when assessing the effects of interventions on economic outcomes if clinical opinion is followed.

8.4 Policy and practice implications and recommendations

When existing guidance is applied to an economic evaluation, the focus is on the ICER for estimating the cost-effectiveness in clinical trials. However, this thesis has shown that looking at the underlying process driving the trial mechanism can provide additional evidence and interpretation to decision makers when funding decisions are made; this additional analysis considers where the costs and QALYs were generated along the trial mechanism pathway.

The new role for conceptual models has implications at several stages of the HTA process, these are discussed below and then recommendations are made.

When a funding bid is released, funders should encourage applications to build in time for conceptual model development, this should allow sufficient time for the development of a conceptual model, sharing it with the wider clinical trial team, and conducting a planned conceptual model driven analysis.

If the conceptual model is developed early on it has the potential to input into the trial design by confirming key components including mediators and

covariates, this can inform the collection of variables important to the economic evaluation. A well-developed conceptual model, which has had input from the clinical trial team, could prevent unnecessary data collection and enable a less burdensome reporting for participants.

In its conventional role as a communication tool the conceptual model can be used to help understanding of the economic evaluation with the wider clinical trial team throughout the HTA process, and to emphasise the importance of economic data collection and minimising missing data.

Results from the conceptual model driven analysis will help interpretation of the treatment arm-based results and provide confirmation of the trial mechanism important to the economic evaluation. The conceptual model provides a framework for discussions of cost-effectiveness results, which in turn can inform future research, either from testing the accuracy of the conceptual model or from the final analysis.

If a body of conceptual models are developed and stored in a public repository, these could provide better information about important and relevant economic outcomes in different disease areas, providing evidence to developers of future trials.

Based on the research in this thesis the following recommendations are made:

1. The suggested approaches presented in this thesis should be followed to develop methodological frameworks, for transparency and consistency.

2. When publishing a methodological framework, the authors should use the term ‘methodological framework’ in the title and keywords to aid future literature searches.

3. The methodological framework included in this thesis should be used when developing a conceptual model for the new role; it provides detailed stages to follow a standardised approach based on existing methodological frameworks.

4. A conceptual model should be developed early on in the trial to reduce bias, input from the trial team and other relevant parties is encouraged as it improves the accuracy of the conceptual model.

5. The conceptual model should be used as a communication tool throughout the clinical trial, this helps with understanding of the economic evaluation, highlighting the key drivers of costs and health benefits.

6. There is potential to use a conceptual model driven analysis of the clinical trial data to provide additional information to decision makers as a supplement to the conventional treatment arm-based analysis, as illustrated in the case studies in this thesis. However, given the potential limitations of the regression method used in the case studies (as discussed previously), further work on alternative methods is needed before the new role for conceptual models can be recommended for routine use.

8.5 Future work

8.5.1 Methodological framework

The suggestions in Chapter 4 for developing methodological frameworks were extracted from a scoping review, this was the first step in developing guidance. To produce more robust recommendations the terminology extracted in the scoping review should be used to develop and conduct a full systematic review. This should be followed by validation of the recommendations with a group of experts to produce best practise guidance.

Particular attention should be given to develop a standardised procedure for collecting qualitative data in phase one, this would add consistency and transparency to the evidence gathering.

8.5.2 Evaluating and validating the methodological framework for developing conceptual models

Within the scope of this thesis the methodological framework for developing conceptual models was evaluated by comparing it to existing guidance for developing conceptual models. Further work needs to be done to validate the

methodological framework; it should be scrutinised by experts and consensus on the contents of the methodological should be reached using suitable methods, for example a Delphi panel or focus groups. The methodological framework should also be evaluated by users in additional case studies.

8.5.3 New role evaluation

This thesis introduced a new role for conceptual models in economic evaluation and demonstrated it with two case studies. This new role should be evaluated and developed further: it should be evaluated by applying the new role to more case studies in a variety of disease areas, research problems and levels of certainty in the treatment arm-based results, this will highlight any potential weaknesses of the new role or situations when the role is not suitable. Further development should assess more sophisticated mathematical methods of analysis such as path analysis and DAGs to confirm the model structure and structural equation modelling if a latent variable is expected to drive the conceptual model.

Future work should also assess the value of qualitative research to inform the conceptual model, understand the trial mechanism and interpret results; in developing the case studies and in my experience of conducting economic evaluations, qualitative research carried out as part of the clinical trial has been crucial for informing the economic evaluation, particularly in understanding trial mechanisms, services provided, and participant behaviour. The role of realist evaluations should be considered and explored in this assessment of qualitative research.

8.6 Conclusions

The research question posed in this thesis was 'How can conceptual modelling enhance health economic evaluation?' This thesis found that it is possible to use conceptual modelling to enhance economic evaluations using the proposed new role. The thesis showed that conducting an analysis, guided by a conceptual model, does provide additional understanding and further interpretation to conventional economic evaluation, and has the potential to inform future clinical trial designs. Overall, the thesis concluded that performing the

supplementary conceptual model driven analysis is worth doing and should be routinely conducted.

The two objectives of the thesis were to:

- Propose and demonstrate a new role for conceptual models in health economic evaluations.

- Propose and demonstrate a methodological framework for developing conceptual models in this new role.

The first objective was fulfilled by developing and presenting the new role and demonstrating it with two case studies. The new role has three components: 1) developing a conceptual model depicting key components in the economic evaluation, with associations linking the key components to illustrate the trial mechanism driving the results, this can be used as a communication tool; 2) testing the associations in the conceptual model for accuracy to provide additional insight into what is driving the economic evaluation results, and 3) conducting an additional analysis based on the strong associations in the conceptual model, using the conceptual model as a framework, the results of which are compared to the original treatment arm-based evaluation to provide further insight.

The second objective was fulfilled in two parts: 1) using scoping review methodology to identify approaches used in developing methodological frameworks and provide suggestions based on these approaches, and 2) applying these suggestions to develop a methodological framework for developing conceptual models; a scoping review identified methods used for developing conceptual models outwith economic evaluation, and these methods were grouped into stages to develop a methodological framework. The methodological framework was demonstrated with two case studies; conceptual models were developed using the methodological framework, these conceptual models were used to demonstrate the new role.

Appendices

Appendix 1: NIHR HTA reports published in 2020 (Volume 24) (*Chapter 2)*

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Allotey J, et al. (201)	Validation and development of models using clinical, biochemical and ultrasound markers for predicting pre-eclampsia: an individual participant data meta-analysis	No						
Brealey S, et al. (202)	Surgical treatments compared with early structured physiotherapy in secondary care for adults with primary frozen shoulder: the UK FROST three-arm RCT	Yes	Within trial	No	No, 95%CI difference in costs does not cross 0	Mention of ACR being more expensive but unclear if this relates to total or treatment costs	No	No
Hagen S, et al. (203)	Basic versus biofeedback-mediated intensive pelvic floor muscle training for women with urinary incontinence: the OPAL RCT	Yes	Within trial	Yes, total costs	Yes, 95% CI crosses 0	Results mention that treatment costs are higher in one arm (SS) and remaining costs not SS, overall total costs	Yes	No

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
						not SS. Discussion section talks about what the treatment costs were driven by		
Shalhoub J, et al. (204)	Compression stockings in addition to low-molecular-weight heparin to prevent venous thromboembolism in surgical inpatients requiring pharmacoprophylaxis: the GAPS non-inferiority RCT	No						
McRobbie HJ, et al. (205)	Nicotine replacement treatment, e-cigarettes and an online behavioural intervention to reduce relapse in recent ex-smokers: a multinational four-arm RCT	Yes	QoL only, trial curtailed					

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Simpson CR, et al. (206)	Vaccine effectiveness of live attenuated and trivalent inactivated influenza vaccination in 2010/11 to 2015/16: the SIVE II record linkage study	No						
Medina-Lara A, et al. (207)	Cancer diagnostic tools to aid decision-making in primary care: mixed-methods systematic reviews and cost-effectiveness analysis	Yes	DAM					
Barker KL, et al. (208)	Outpatient physiotherapy versus home-based rehabilitation for patients at risk of poor outcomes after knee arthroplasty: CORKA RCT	Yes	Within trial	Yes, total costs for two perspe ctives	Yes, 95% CI crosses 0 for all cost categories	Results section has presentation of costs and differences between arms for cost categories, no discussion	No	No
Batchelor JM, et al. (209)	Home-based narrowband UVB, topical corticosteroid or combination for children and adults with vitiligo: HI-Light Vitiligo three-arm RCT	Yes	Within trial (3 arm)	No	No, 95%CI difference in costs does not cross 0	Discussion section discusses that higher treatment costs of combination treatment not offset by NHS cost savings	Yes	Yes

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Taylor AH, et al. (210)	Adding web-based behavioural support to exercise referral schemes for inactive adults with chronic health conditions: the e-coachER RCT	Yes	Within trial	Yes, total costs	Yes, 95%CI crosses 0	Discussion section includes discussion of difference in costs for different cost categories, but not treatment costs	Yes	No
Crawford F, et al. (211)	Risk assessments and structured care interventions for prevention of foot ulceration in diabetes: development and validation of a prognostic model	Yes	DAM					
Pickard R, et al. (212)	Open urethroplasty versus endoscopic urethrotomy for recurrent urethral stricture in men: the OPEN RCT	Yes	Within trial and DAM	Yes, total costs	No, 95%CI difference in costs does not cross 0 and urethrotomy is dominant	None	No	No
Clarkson JE, et al. (213)	Risk-based, 6-monthly and 24-monthly dental check-ups for adults: the INTERVAL three-arm RCT	Yes	Within trial CUA, CBA and WTP	Yes, total costs for CUA, CBA	Yes, 95% CI crosses 0 in all analyses	None	No	No

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
				and WTP				
Abel KM, et al. (214)	An intervention to improve the quality of life in children of parents with serious mental illness: the Young SMILES feasibility RCT	No						
Appleton RE, et al. (215)	Levetiracetam as an alternative to phenytoin for second-line emergency treatment of children with convulsive status epilepticus: the EcLiPSE RCT	No						
Gilbert R, et al. (216)	Antimicrobial-impregnated central venous catheters for preventing neonatal bloodstream infection: the PREVAIL RCT	Yes	Cost study and DAM					
Stephenson J, et al. (217)	An interactive website to aid young women's choice of contraception: feasibility and efficacy RCT	No						

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Duffy S, et al. (218)	Annual mammographic screening to reduce breast cancer mortality in women from age 40 years: long-term follow-up of the UK Age RCT	No						
Rodgers H, et al. (219)	Robot-assisted training compared with an enhanced upper limb therapy programme and with usual care for upper limb functional limitation after stroke: the RATULS three-group RCT	Yes	Within trial and DAM	Yes, total costs	Yes for one comparison. No, 95% CI does not cross 0 with usual care and robot, but yes uncertain between usual care and EULT	Results section has brief discussion cost categories and then adding treatment costs reverses trend	Yes	Yes
Van den Bruel A, et al. (220)	Non-contact infrared thermometers compared with current approaches in primary care for children aged 5 years and under: a method comparison study	No						
Dias J, et al. (221)	Surgical fixation compared with cast immobilisation for adults with a bicortical fracture of the scaphoid waist: the SWIFFT RCT	Yes	Within trial and DAM	No	No, 95% CI does not cross 0	No, only discussion about total costs	No	No

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Armstrong N, et al. (222)	Avatrombopag and lusutrombopag for thrombocytopenia in people with chronic liver disease needing an elective procedure: a systematic review and cost-effectiveness analysis	Yes	Systematic review and DAM					
Bray N, et al. (223)	Powered mobility interventions for very young children with mobility limitations to aid participation and positive development: the EMPoWER evidence synthesis	Yes	Cost analysis only					
Foster NE, et al. (224)	Stratified versus usual care for the management of primary care patients with sciatica: the SCOPiC RCT	Yes	Within trial, two perspectives	Yes, total costs for each perspective	Yes, 95% CI crosses 0	No, only discussion about total costs	No	No

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Walton M, et al. (225)	Selective internal radiation therapies for unresectable early-, intermediate- or advanced-stage hepatocellular carcinoma: systematic review, network meta-analysis and economic evaluation	Yes	Systematic review and DAM					
Gilson R, et al. (226)	Imiquimod versus podophyllotoxin, with and without human papillomavirus vaccine, for anogenital warts: the HIPvac factorial RCT	Yes	Within trial, 4 arms	Yes, total costs	No, 95% CI does not cross 0	Results section some discussion on most costly category	Yes	No
Leaviss J, et al. (227)	Behavioural modification interventions for medically unexplained symptoms in primary care: systematic reviews and economic evaluation	Yes	Systematic review					
Baker P, et al. (228)	Occupational advice to help people return to work following lower limb arthroplasty: the OPAL intervention mapping study	No						

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Crawford MJ, et al. (229)	Switching antipsychotic medication to reduce sexual dysfunction in people with psychosis: the REMEDY RCT	Yes	Descriptive statistics					
Melton H, et al. (230)	Interventions for adults with a history of complex traumatic events: the INCiTE mixed-methods systematic review	No						
Lewis AL, et al. (231)	Urodynamics tests for the diagnosis and management of bladder outlet obstruction in men: the UPSTREAM non-inferiority RCT	Yes	Within trial, using 3 perspectives	No	Hard to tell as there is no table with incremental costs	No	No	No
Worthington J, et al. (232)	Thulium laser transurethral vaporesection versus transurethral resection of the prostate for benign prostatic obstruction: the UNBLOCS RCT	Yes	Within trial with 2 perspectives	No	Hard to tell as no incremental costs for CEA, only CCA- one treatment weakly dominated	Results section small discussion on what was driving the costs	Yes	No
Earl H, et al. (233)	Six versus 12 months' adjuvant trastuzumab in patients with HER2-positive early breast cancer: the PERSEPHONE non-inferiority RCT	Yes	Within trial and DAM	Yes	No, 95% CI does not cross 0	No, only vial sharing	No	No

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Corbett M, et al. (234)	Point-of-care creatinine tests to assess kidney function for outpatients requiring contrast-enhanced CT imaging: systematic reviews and economic evaluation	Yes	DAM					
Costa ML, et al. (235)	Negative-pressure wound therapy compared with standard dressings following surgical treatment of major trauma to the lower limb: the WHiST RCT	Yes	Within trial	Yes total costs only	No, 95% CI does not cross 0 for base-case analysis	No	No	No
Hamdy FC, et al. (236)	Active monitoring, radical prostatectomy and radical radiotherapy in PSA-detected clinically localised prostate cancer: the ProtecT three-arm RCT	Yes	Within trial, three arm comparison	No	No for RT v AM comparison and yes for RP v RT comparison	In Results section discussion of resource use and distribution of costs over time	No	No
Jones AP, et al. (237)	Different corticosteroid induction regimens in children and young people with juvenile idiopathic arthritis: the SIRJIA mixed-methods feasibility study	No						

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Rodgers M, et al. (238)	Interventions based on early intensive applied behaviour analysis for autistic children: a systematic review and cost-effectiveness analysis	Yes	DAM					
Hernández Alava M, et al. (239)	Mapping clinical outcomes to generic preference-based outcome measures: development and comparison of methods	No						
Coomarasamy A, et al. (240)	Progesterone to prevent miscarriage in women with early pregnancy bleeding: the PRISM RCT	Yes	Within trial	Yes total costs only	Yes, 95% CI crosses 0	Key driver only	Yes	No
Lobban F, et al. (241)	A web-based, peer-supported self-management intervention to reduce distress in relatives of people with psychosis or bipolar disorder: the REACT RCT	Yes	Within trial	Yes total costs only	Yes, 95% CI crosses 0	Discussion in Results section of greatest cost categories, in Discussion section discussion about which treatment had the highest treatment costs	Yes	No

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Fraser H, et al. (242)	Rapid antigen detection and molecular tests for group A streptococcal infections for acute sore throat: systematic reviews and economic evaluation	Yes	DAM					
Hounsome J, et al. (243)	Prophylactic removal of impacted mandibular third molars: a systematic review and economic evaluation	Yes	Systematic review and DAM					
Davis S, et al. (244)	Denosumab, raloxifene, romosozumab and teriparatide to prevent osteoporotic fragility fractures: a systematic review and economic evaluation	Yes	Systematic review and DAM					
Winkley K, et al. (245)	Psychological interventions to improve self-management of type 1 and type 2 diabetes: a systematic review	Yes	DAM					

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Witham MD, et al. (246)	Sodium bicarbonate to improve physical function in patients over 60 years with advanced chronic kidney disease: the BiCARB RCT	Yes	Within trial	Yes total costs only	Mix of certain and uncertain results	None	No	No
Bhatnagar R, et al. (247)	Thoracoscopy and talc poudrage compared with intercostal drainage and talc slurry infusion to manage malignant pleural effusion: the TAPPS RCT	Yes	Within trial, although not clear	Yes total costs only	Yes, 95% CI crosses 0	None	No	No
Poolman M, et al. (248)	Carer administration of as-needed subcutaneous medication for breakthrough symptoms in people dying at home: the CARiAD feasibility RCT	No						
Shaw L, et al. (249)	An extended stroke rehabilitation service for people who have had a stroke: the EXTRAS RCT	Yes	Within trial	Yes, total costs only	Yes, 95% CI crosses 0	Results section discusses cost categories and reductions/increases in these between arms. Discussion section discusses 'cost shifting' where	Yes	Yes

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
						cost saving may accrue to service/organisation not paying for the intervention - disincentive/barrier to adopting cost-effective intervention		
Tume LN, et al. (250)	Routine gastric residual volume measurement to guide enteral feeding in mechanically ventilated infants and children: the GASTRIC feasibility study	No						
Dennis M, et al. (251)	Fluoxetine to improve functional outcomes in patients after acute stroke: the FOCUS RCT	Yes	Within trial	Yes total costs only	Yes, 95% CI crosses 0	None	No	No

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Davidson B, et al. (252)	Liver resection surgery compared with thermal ablation in high surgical risk patients with colorectal liver metastases: the LAVA international RCT	No						
Beard DJ, et al. (253)	Total versus partial knee replacement in patients with medial compartment knee osteoarthritis: the TOPKAT RCT	Yes	Within trial	Yes, total costs	No, 95% CI does not cross 0	Results discusses key cost drivers. Summary discusses index surgery and health-care costs	Yes	No
Palmer R, et al. (254)	Computerised speech and language therapy or attention control added to usual care for people with long-term post-stroke aphasia: the Big CACTUS three-arm RCT	Yes	Within trial and DAM		No, 95% CI does not cross 0	Discussion of total costs in each arm, and discussion of training costs	Yes	No
Dorling J, et al. (255)	Two speeds of increasing milk feeds for very preterm or very low-birthweight infants: the SIFT RCT	Yes	Within trial	Yes, total costs	Yes, 95% CI does cross 0	No	No	No

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Mallucci CL, et al. (256)	Silver-impregnated, antibiotic-impregnated or non-impregnated ventriculoperitoneal shunts to prevent shunt infection: the BASICS three-arm RCT	Yes	Within trial, three arm comparison	No	Yes, both 95% CI cross 0	Discussion section discusses that higher upfront cost of antibiotic-impregnated shunt catheters could be justifies by associated cost savings of further surgery and hospital care. Also talks about CMA results if they had been appropriate	Yes	Yes
Surr CA, et al. (257)	Dementia Care Mapping to reduce agitation in care home residents with dementia: the EPIC cluster RCT	Yes	Within trial	Yes, total costs	No, 95% CI does not cross 0	One treatment more costly due to one cost category with 6 participants with very high costs. No other discussion about breakdown of costs	Yes	No

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Francis NA, et al. (258)	C-reactive protein point-of-care testing for safely reducing antibiotics for acute exacerbations of chronic obstructive pulmonary disease: the PACE RCT	Yes	Within trial	Yes, total costs	No, 95% CI does not cross 0	Results section discusses cost categories with higher costs and cost savings, and displays bar charts with breakdown of costs per category. Summary section discusses savings in one specific cost category for intervention arm, but not other categories and QoL	Yes	No
Day C, et al. (259)	An intervention for parents with severe personality difficulties whose children have mental health problems: a feasibility RCT	No						
Hemming C, et al. (66)	Surgical interventions for uterine prolapse and for vault prolapse: the two VUE RCTs	Yes	Within trial and DAM	Yes, total costs	No, 95% CI does not cross 0	Discussion about why the intervention was more expensive	Yes	No

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
						than the comparator		
Snowden C, et al. (260)	Preoperative behavioural intervention to reduce drinking before elective orthopaedic surgery: the PRE-OP BIRDS feasibility RCT	No						
Stevenson M, et al. (261)	Interventions to reduce the risk of surgically transmitted Creutzfeldt Jakob disease: a cost-effective modelling review	Yes	DAM					
Rake C, et al. (262)	High-dose oral vitamin D supplementation and mortality in people aged 65 84 years: the VIDAL cluster feasibility RCT of open versus double-blind individual randomisation	No						

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Davies NM, et al. (263)	Varenicline versus nicotine replacement therapy for long-term smoking cessation: an observational study using the Clinical Practice Research Datalink	No						
Costa ML, et al. (264)	Plaster cast versus functional bracing for Achilles tendon rupture: the UKSTAR RCT	Yes	Within trial	Yes, total costs		Results section presents treatment costs per treatment arm, and other costs lower in treatment but did not specifically say that treatment costs are only known costs. Discussion section included same presentation of cost savings in higher cost treatment	Yes	Yes

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Riley P, et al. (265)	Oral splints for patients with temporomandibular disorders or bruxism: a systematic review and economic evaluation	Yes	DAM					
Froggatt K, et al. (266)	A group intervention to improve quality of life for people with advanced dementia living in care homes: the Namaste feasibility cluster RCT	Yes	Cost analysis and feasibility of resource use collection					
Edwards SJ, et al. (267)	Implantable cardiac monitors to detect atrial fibrillation after cryptogenic stroke: a systematic review and economic evaluation	Yes	Systematic review and DAM					
Lincoln NB, et al. (268)	Group cognitive rehabilitation to reduce the psychological impact of multiple sclerosis on quality of life: the CRAMMS RCT	Yes	Within trial	Yes, total costs	Yes, 95% CI crosses 0	Results section discusses the participants who were cost outliers. Discussion section discusses the outliers and drivers but no discussion of treatment costs	Yes	No

Appendix 1

Authors	Title	Includes EE	EE type	CE plane	Uncertain cost results	Discussion of treatment costs separately	Different cost categories discussed	Treatment costs offset by cost savings in other cost categories discussed
Duarte R, et al (269)	Lead-I ECG for detecting atrial fibrillation in patients with an irregular pulse using single time point testing: a systematic review and economic evaluation	Yes	Systematic review and DAM					
Fleeman N, et al. (270)	Lenvatinib and sorafenib for differentiated thyroid cancer after radioactive iodine: a systematic review and economic evaluation	Yes	Systematic review and DAM					
Maguire A, et al. (271)	Best-practice prevention alone or with conventional or biological caries management for 3- to 7-year-olds: the FiCTION three-arm RCT	Yes	Within trial	No	Yes, 95% CI crosses 0	Discussion section only discusses total costs and no separate categories	No	No

Appendix 2: Scoping review search terms (*Chapter 4)*

OVID Medline search - September 2018

1. "develop*".m_titl.

2. methodological framework.m_titl.

3. "design*".m_titl.

4. "creat*".m_titl.

5. 1 or 3 or 4

6. 2 and 5

Web of Science search - September 2018

#1 TITLE: (methodological framework)

DocType=All document types; Language=All languages

#2 TITLE: (develop*) OR TITLE: (creat*) OR TITLE: (design*) OR TITLE: (writ*)

DocType=All document types; Language=All languages;

#3 #1 AND #2

Appendix 3: Results of scoping review – study characteristics (*Chapter 4)*

Author(s)	Title	Year of publication	Type of record	Subject area	Source origin
Achillas, C; Aidonisb, D; Iakovouc, E; Thymianidisa, M; Tzetzisa, D (35)	A methodological framework for the inclusion of modern additive manufacturing into the production portfolio of a focused factory	2015	Journal article	Manufacturing	Greece
Anagnostou, A; Taylor, S (34)	A distributed simulation methodological framework for OR/MS applications	2017	Journal article	Simulation modelling	UK
Battini, D.; Faccio, M.; Persona, A.; Sgarbossa, F (33)	New methodological framework to improve productivity and ergonomics in assembly system design	2011	Journal article	Manufacturing	Italy
Brondizio, E; Vogt, N; Mansur, A; Anthony, E; Costa, S; Hetrick, S (272)	A conceptual framework for analysing deltas as coupled social-ecological systems: an example from the Amazon River Delta	2016	Journal article	Ecology	US
Chesson, L; Howa, J; Lott, M; Ehleringer, J (168)	Development of a methodological framework for applying isotope ratio mass spectrometry to explosive components	2016	Journal article	Forensic chemistry	US
Dean, E; Taylor, M; Francis H; Lisboa, P; Appleton, D; Jones, M (163)	A methodological framework for Geographic Information Systems development	2017	Journal article	GIS	UK
George, H.; Bosc, P. M.; Even, M. A.; Belieres, J. F.; Bessou, C. (31)	WAW proposed methodological framework to monitor agricultural structural transformations and their contributions to sustainable development	2012	Conference proceedings	Agricultural systems	Italy
Halbe, J; Pahl-Wostl, C; Adamowski, J (32)	A methodological framework to support the initiation, design and institutionalization of participatory	2018	Journal article	Ecology modelling	Canada

Appendix 3

Author(s)	Title	Year of publication	Type of record	Subject area	Source origin
	modeling processes in water resources management				
Ianni, M; de Leon, M (273)	Applying Energy Performance-Based Design in Early Design Stages A methodological framework for integrating multiple BPS tools	2013	Journal article/conference proceedings	Architecture	Spain
Kallio, H; Pietila, A; Johnson, M; Kangasniemi, M (38)	Systematic methodological review: developing a framework for a qualitative semi-structured interview guide	2016	Journal article	Qualitative research	Finland
Kumar, A; Singh, A; Deng, Y; He, X; Kumar, P; Bansal, R (274)	A novel methodological framework for the design of sustainable rural microgrid for developing nations	2018	Journal article	Electricity supply in rural countries	China
Kumke, M; Watschke, H; Vietor, T (167)	A new methodological framework for design for additive manufacturing	2016	Journal article	Manufacturing	Germany
Lee, J; Jang, S (30)	A methodological framework for instructional design model development: critical dimensions and synthesized procedures	2014	Journal article	Education	Korea
Linek, S; Schwarz, D; Bopp, M; Albert, D (157)	When playing meets learning: methodological framework for designing educational games	2010	Book chapter	Educational games	Austria
Lopes, A; Ruiz-Cecilia, R (161)	Designing technology-mediated tasks for language teaching: A methodological framework	2017	Journal article	Education	Portugal
Nicod, E; Kanavos, P (165)	Developing an evidence-based methodological framework to systematically compare HTA coverage decisions: A mixed methods study	2016	Journal article	Healthcare	UK

Appendix 3

Author(s)	Title	Year of publication	Type of record	Subject area	Source origin
Pahl-Wostl, C; Holtz, G; Kastens, B; Knieper, C (37)	Analyzing complex water governance regimes: the Management and Transition Framework	2010	Journal article	Ecology	Germany
Panagiotopoulou, M; Stratigea, A (275)	A participatory methodological framework for paving alternative local tourist development paths-the case of Sterea Ellada Region	2014	Journal article	Tourism	Greece
Procházka, J: Melichar, J (36)	Methodological framework for operational risk assessment	2017	Journal article	Defence	Czech Republic
Reed, M; Kenter, J; Bonn, A; Broad, K; Burt, T; Fazey, I; Fraser, E; Hubacek, K; Nainggolan, D; Quinn, C; Stringer, L; Ravera, F (28)	Participatory scenario development for environmental management: A methodological framework illustrated with experience from the UK uplands	2013	Journal article	Ecology	UK
Reidsma, P; Konig, H; Feng, S; Bezlepkina, I; Keulen, H. van; Ittersum, M van; Brouwer, F (29)	A methodological framework for sustainability impact assessment of land use policies in developing countries: re-using and complementing approaches	2009	Conference proceedings	Ecology	The Netherlands

Appendix 3

Author(s)	Title	Year of publication	Type of record	Subject area	Source origin
Rijke, J; Brown, R; Zevenbergen, C; Ashley, R; Farrelly, M; Morison, P; van Herk, S (276)	Fit-for-purpose governance: A framework to make adaptive governance operational	2012	Journal article	Ecology	The Netherlands
Rodgers, M; Thomas, S; Harden, M; Parker, G; Street, A; Eastwood, A (42)	Developing a methodological framework for organisational case studies: a rapid review and consensus development process	2016	Journal article	Organisation	UK
Schmitt, J; Apfelbacher, C; Spuls, P; Thomas, K; Simpson, E; Furue, M; Chalmers, J; Williams, H (277)	The Harmonizing Outcome Measures for Eczema (HOME) roadmap: A methodological framework to develop core sets of outcome measurements in dermatology	2015	Journal article	Healthcare	Germany
Stratigea A; Papadopoulou, C-A (158)	Foresight analysis at the regional level - A participatory methodological framework	2013	Journal article	Regional studies	Greece
Stremke, S; Van Kann, F; Koh, J (162)	Integrated visions (Part I): Methodological framework for long-term regional design	2012	Journal article	Regional design	The Netherlands
Squires, H; Chilcott, J; Akehurst, R; Burr, J; Kelly, M (22)	A framework for developing the structure of public health economic models	2016	Journal article	Health economics	UK

Appendix 3

Author(s)	Title	Year of publication	Type of record	Subject area	Source origin
Sun, Y; Strobel, J (164)	Elementary Engineering Education (EEE) adoption and expertise development framework: An inductive and deductive study	2013	Journal article	Education	US
Tappenden, P; Chilcott, J; Brennan, A; Squires, H; Stevenson, M (160)	Whole disease modeling to Inform resource allocation decisions in cancer: A methodological framework	2012	Journal article	Health economics	UK
Tondel, K; Niederer, S; Land, S; Smith, N (159)	Insight into model mechanisms through automatic parameter fitting: a new methodological framework for model development	2014	Journal article	Biology models	UK

NB: Where the country of employment varies between authors, the country of employment of the lead author, at the time of publication is used.

Appendix 4: PRISMA Scoping review checklist (Publication from Chapter 4)

Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) Checklist

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a scoping review.	Page 1
ABSTRACT			
Structured summary	2	Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	Pages 2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	Page 4
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	Pages 4
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	Page 5
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	Page 6
Information sources*	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	Page 6
Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.	Additional file 1
Selection of sources of evidence†	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	Page 6
Data charting process‡	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	Page 6/7
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	Page 6/7
Critical appraisal of individual sources of evidence§	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).	N/A

Appendix 4

Synthesis of results	13	Describe the methods of handling and summarizing the data that were charted.	Page 7
RESULTS			
Selection of sources of evidence	14	Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	Page 8
Characteristics of sources of evidence	15	For each source of evidence, present characteristics for which data were charted and provide the citations.	Page 8 and Additional File 3
Critical appraisal within sources of evidence	16	If done, present data on critical appraisal of included sources of evidence (see item 12).	N/A
Results of individual sources of evidence	17	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	Pages 8-12
Synthesis of results	18	Summarize and/or present the charting results as they relate to the review questions and objectives.	Pages 12-13
DISCUSSION			
Summary of evidence	19	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.	Page 14-15
Limitations	20	Discuss the limitations of the scoping review process.	Page 15
Conclusions	21	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	Page 16
FUNDING			
Funding	22	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	Page 17

JBI = Joanna Briggs Institute; PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews.

* Where *sources of evidence* (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites.

† A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused with *information sources* (see first footnote).

‡ The frameworks by Arksey and O'Malley (6) and Levac and colleagues (7) and the JBI guidance (4, 5) refer to the process of data extraction in a scoping review as data charting.

§ The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision. This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion, and policy document).

From: Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. ;169:467–473. doi: 10.7326/M18-0850

Appendix 5: Scoping review search terms (*Chapter 5)*

Web of Science - November 2016 (with alert to October 2018)
TITLE: Conceptual model* NEAR framework

OR

TITLE: Conceptual model* NEAR method

OR

TITLE: Conceptual model* NEAR guide*

Scopus - November 2016 (with alert to October 2018)
(TITLE "conceptual model*" W/5 "framework")

OR

(TITLE "conceptual model*" W/5 "method")

OR

(TITLE "conceptual model" W/5 "guide*")

Appendix 6: Results of scoping review (*Chapter 5)*

Authors	Title	Journal/source	Type of record	Year	Search	Discipline
Abdelmegid, MA Gonzalez, VA Naraghi, AM O'Sullivan, M Walker, CG Poshdar, M (12)	Towards a conceptual modeling framework for construction simulation	2017 Winter Simulation Conference	Proceedings paper	2017	WoS	Construction
Brassington, FC; Younger, PL (15)	A proposed framework for hydrogeological conceptual modelling	Water And Environment Journal	Article	2010	WoS	Ecology (groundwater)
Chwif, L; Banks, J; de Moura, JP; Santini, B(21)	A framework for specifying a discrete-event simulation conceptual model	Journal Of Simulation	Article	2013	WoS	Operations research
Furian, N; O'Sullivan, M; Walker, C; Vossner, S; Neubacher, D (18)	A conceptual modeling framework for discrete event simulation using hierarchical control structures	Simulation Modelling Practice And Theory	Article	2015	WoS	Engineering
Gentile, JH; Harwell, MA; Cropper, W; Harwell, CC; DeAngelis, D; Davis, S; Ogden, JC; Lirman, D (172)	Ecological conceptual models: a framework and case study on ecosystem management for South Florida sustainability	Science Of The Total Environment	Article	2001	WoS	Ecology
Authors	**Title**	**Journal/source**	**Type of record**	**Year**	**Search**	**Discipline**

Appendix 6

Gray, K; Sockolow P (16)	Conceptual models in health informatics research: A literature review and suggestions for development	JMIR informatics	Article	2016	Google CM steps	Healthcare
Gross, J (20)	Developing conceptual models for monitoring programs	NPS Website	NPS inventory and monitoring program	2003	Google CM steps (cited by NSW website)	Ecology
Montevechi, JAB; Friend, JD (13)	Using a soft systems methodology framework to guide the conceptual modeling process in discrete event simulation	2012 Winter Simulation Conference	Proceedings Paper	2012	WoS	Industrial engineering
NSW website (178)	Building a conceptual model	NSW website	Office of environment and heritage	2011	Google CM steps	Ecology
Pace, D (176)	Ideas about simulation conceptual model development	John Hopkins Applied Physics Laboratory Technical Digest	Article	2000	Google CM development	Applied physics
Pereira, TF; Montevechi, JAB; Miranda, RD; Friend, JD (179)	Integrating soft systems methodology to aid simulation conceptual modeling	International Transactions In Operational Research	Article	2015	WoS	Manufacturing
Robinson, S (43)	Conceptual modelling for simulation Part II: a framework for conceptual modelling	Journal Of The Operational Research Society	Article	2008	WoS	Operations research
Authors	**Title**	**Journal/source**	**Type of record**	**Year**	**Search**	**Discipline**

Appendix 6

Tako, AA; Kotiadis, K; Vasilakis, C (175)	A participative modelling framework for developing conceptual models in healthcare simulation studies	Proceedings Of The 2010 Winter Simulation Conference	Proceedings Paper	2010	WoS	Healthcare
van der Zee, DJ (44)	An integrated conceptual modeling framework for simulation linking simulation - modeling to the systems engineering process	2012 Winter Simulation Conference	Proceedings Paper	2012	WoS	Engineering
Hesch, W (19)	Conceptual model development for MODFLOW or FEFLOW models	FEFLOW conference HydropGeoBuilder (Finite element flow simulator)	Presentation	2009	Google CM development	Ecology (software)
Morgan, A (WWF) (14)	Basic guidance for cross-cutting: Conceptual models	WWF Website	Charity guide	2005	Google CM steps	Ecology

Appendix 7: Terminology allocated to themes (*Chapter 5)*

Theme	Author's name for theme	Included framework
Objectives	'Defining the model objectives'	Abdelmegid (12)
	'Defining the objectives'	Brassington (15)
	'Defining the OCIR (*objectives*, complexity, input/outputs and runs)'	Chwif (21)
	'Define objectives'	Hesch (19)
	'Identification of modelling and general objectives'	Furian (18)
	'Define the environmental goals and objectives'	Gentile (172)
	'Clearly state the goals of the conceptual models'	Gross (20)
	'Developing simulation project objectives through SSM'	Montevechi (13)
	'Clearly state goals for developing the model'	NSW (178)
	'Objectives and system definition'	Pereira (177)
	'Determining the modelling and general project objectives'	Robinson (43)
	'Study objectives'	Tako (175)
	'Determine objectives – modelling objectives/general project objectives'	Van der Zee (174)
Model scope	'Designing conceptual model: Individual model behaviour'	Abdelmegid (12)
	'Defining the OCIR (objectives, *complexity*, input/outputs and runs)'	Chwif (21)
	'Model content (scope, level of detail)'	Furian (18)
	'Delineate the spatial, temporal and ecological scales and boundaries'	Gentile (172)
	'Identify bounds of the system of interest'	Gross (20)
	'Identify the scope of what is under investigation'	NSW (178)

Appendix 7

Theme	Author's name for theme	Included framework
	'Address relationships among simulation relationships'	Pace (176)
	'Determining the model content (*scope* and level of details), identifying any assumptions and simplifications'	Robinson (43)
	'Determine model contents, *scope* and level of detail'	Van der Zee (174)
	'Scope and vision'	WWF (14)
Model content	'Defining the topography and surface water drainage', 'defining the geology', 'defining the aquifer framework and boundaries', defining groundwater flow directions', 'defining the aquifer relationships' and 'water balance'	Brassington (15)
	'Build conceptual model'	Hesch (19)
	'Allow the conceptual model to influence the research design'	Gray (16)
	'Developing model inputs, outputs and *content* through SSM'	Montevechi (13)
	'Develop simulation elements'	Pace (176)
	'Construction of the conceptual model'	Pereira (177)
	'Determining the model content'	Robinson (43)
	'Model content'	Tako (175)
System behaviour	'Designing conceptual model: Model control'	Abdelmegid (12)
	Model control	Furian (18)
	'Develop risk hypotheses and stress-effects causal pathways'	Gentile (172)
	'Identify key model components, subsystems, and interactions', develop control models of key systems and subsystems', and 'articulate key questions or alternative approaches'	Gross (20)
	'Identify relationships between components of the model'	NSW (178)
	'Identify entities and processes for representation'	Pace (176)
	'Define the system'	Tako (175)

Appendix 7

Theme	Author's name for theme	Included framework
	'Targets'	WWF (14)
Understanding the problem	'Problem formulation'	Abdelmegid (12)
	'Understanding of the problem situation'	Furian (18)
	'Acknowledge the conceptual models of contributing domains'	Gray (16)
	'Developing system understanding through SSM'	Montevechi (13)
	'Collect authoritative simulation context information'	Pace (176)
	'Understand the problem situation'	Robinson (43)
	'Initiate study'	Tako (175)
	'Understanding the problem and candidate solutions'	Van der Zee (174)
Model inputs	'Determining the conceptual model *inputs* and outputs'	Abdelmegid (12)
	'Defining the OCIR (objectives, complexity, *input*/outputs and runs)'	Chwif (21)
	'Collect data'	Hesch (19)
	'Defining input factors'	Furian (18)
	'Developing model *inputs*, outputs and content through SSM'	Montevechi (13)
	'Identify the model inputs (experimental factors)'	Robinson (43)
	'*Inputs* and outputs'	Tako (175)
	'Identifying the model inputs'	Van der Zee (174)
Model outputs	'Determining the conceptual model inputs and *outputs*'	Abdelmegid (12)
	'Defining the OCIR (objectives, complexity, input/*outputs* and runs)'	Chwif (21)
	'Defining output responses'	Furian (18)

Appendix 7

Theme	Author's name for theme	Included framework
	'Developing model inputs, *outputs* and content through SSM'	Montevechi (13)
	'Identifying the model outputs (responses)'	Robinson (43)
	'Inputs and *outputs*'	Tako (175)
	'Identify the model outputs'	Van der Zee (174)
Documentation	'Describing the conceptual model'	Brassington (15)
	'Revision timetable and adjustments'	Chwif (21)
	'Explicating a conceptual model *verbally* and graphically'	Gray (16)
	'Documentation'	NSW (178)
	'Documentation of the conceptual model'	Pereira (177)
	'Determine model contents: scope and level of detail'	Van der Zee (174)
Assumptions and simplifications	'Designing conceptual model: Individual model behaviour'	Abdelmegid (12)
	'Process description, including model assumptions'	Chwif (21)
	'Documentation'	NSW (178)
	'Determining the model content (scope and level of detail), *identifying any assumptions and simplifications*'	Robinson (43)
	'Assumptions and simplifications'	Tako (175)
	'Determine model contents, scope and level of detail'	Van der Zee (174)
Model detail	'Designing conceptual model: Individual model behaviour'	Abdelmegid (12)
	'Defining the OCIR (objectives, *complexity*, input/outputs and runs)'	Chwif (21)
	'Model individual behaviour'	Furian (18)
	'Identify entities and process for representation'	Pace (176)

Appendix 7

Theme	Author's name for theme	Included framework
	'Determining the model content (scope and *level of detail*), identifying any assumptions and simplifications'	Robinson (43)
	'Determine model detail (attributes) for all components included'	Van der Zee (174)
Diagram	'Describing the conceptual model'	Brassington (15)
	'Develop graphical conceptual model'	Gentile (172)
	'Explicating a conceptual model verbally and *graphically*'	Gray (16)
	'Transform the SSM models into a visual abstraction'	Montevechi (13)
	'Determine model contents: scope and level of detail'	Van der Zee (174)
Entities	'Designing conceptual model: Model structure'	Abdelmegid (12)
	'Model structure'	Furian (18)
	'Identify ecological receptors and at-risk components'	Gentile (172)
	'Identify key components of the model and pressures'	NSW (178)
	'Identify entities and processes for representation'	Pace (176)
Validation	'Suitable?'	Hesch (19)
	'Seek critical feedback on the conceptual model from multiple perspectives'	Gray (16)
	'Conceptual model validation through SSM'	Montevechi (13)
	'Validation of the conceptual model'	Pereira (177)
Stressors	'Inventory resource use and other human activities', 'describe sources of natural and anthropogenic stressors', 'identify the primary and secondary stressors of concern' and 'describe stressor mechanisms and routes of exposure'	Gentile (172)
	'Identify natural and anthropogenic stressors' and 'describe relationships of stressors, ecological factors, and responses'	Gross (20)
	'Context and stakeholder'	WWF (14)

Appendix 7

Theme	Author's name for theme	Included framework
Review and refine	'Revisit the conceptual model in light of the research findings'	WWF (14)
	'Review, revise, refine models'	Gross (20)
	'Review and refine the model'	NSW (178)
Model structure	'Designing conceptual model: Model structure'	Abdelmegid (12)
	'Model structure'	Furian (18)
Team identification	'DES study initiation'	Abdelmegid (12)
	'Initiate study'	Tako (175)
Use previous conceptual model	'Review conceptual models already used in health informatics'	Gray (16)

Appendix 8: Ordering and sequence of themes in each phase (*Chapter 5)*

Below, the order of the themes included within each phase is explained, and the theme descriptions (if reported) are summarised with advice from the methodological framework.

Phase 1 – Getting familiar with the research problem and setting the objectives of the model

This first phase mainly includes getting to know the problem situation and setting the objectives of the conceptual model. It comprises four stages;

- Understanding the problem
- Use previous conceptual model
- Identify the team
- Setting objectives

Understanding the problem

'Understanding the problem' is the first stage in this phase of the draft methodological framework. 'Understanding the problem' is nearly always the first stage when it is included in any of the methodological frameworks. Understanding the problem is key to developing a conceptual model; without a good and detailed understanding of the problem it would be impossible to develop an accurate conceptual model (173). And the more complete and clear a problem area is the easier it is to understand (176).

Seven of the eight frameworks that include this stage mention stakeholders (not Montevechi), either coming to an agreement with them on the understanding of the research problem or using the stakeholders to explain the problem (12, 16, 18, 43, 44, 175, 176). Furian and Robinson also suggest that the modeller introduces alternative hypotheses in order to come to a consensus with the

stakeholders(18, 43). Abdelmegid comments that limitations in knowledge at this stage will result in assumptions being made (12).

Furian states that the outcome of this stage is an informal description of the problem, including assumptions used in reaching an understanding of the problem(18). Robinson advises that assumptions made in coming to an understanding of the problem situation are recorded(43). Gray and Robinson both recommend using conceptual models of the problem situation as a tool to describe the problem (16, 43).

Four methodological frameworks recommend using soft systems methodology to gain an understanding of the problem situation (12, 18, 43, 173), in particular the Purposeful Activity Model. Robinson also suggests the use of cognitive mapping and causal loop diagrams to help understanding the problem.

Use previous conceptual model

This stage was only included in one methodological framework; Gray, it was included as a second stage in the framework (16). Gray suggests identifying previous conceptual models that could be adapted for use in the current project. I have included it in this phase however, it may be good practice to review previous conceptual models during Phase 2 too, in relation to the model structure and content.

Identify the team

This stage was included in two methodological frameworks, Tako and Abdelmegid (12, 175). Tako placed this stage with the 'Understanding the problem' stage in the initial section of the framework, and Abdelmegid placed it at the initiation of the study.

The majority of included methodological frameworks mentioned working with stakeholders in various stages, and it seemed to be taken as read that the modeller would not work alone in developing the conceptual model. Tako suggests a maximum of 12 people on the project team, comprising modellers and stakeholders, adding that stakeholders are experts in the problem situation, and

should include decision makers who ought to be the key stakeholders. Tako also recommends appointing a 'project champion' who would oversee the conceptual model. Abdelmegid suggests meeting with the problem owner to collect information on the system and understand it, and a stakeholder list should be developed.

Setting objectives

The final stage in this phase is 'Setting objectives'. This stage has the highest number of mentions in the methodological frameworks; 13 frameworks included this stage. This stage is the second one in this phase because out of the 13 methodological frameworks which included 'Setting objectives' as a stage, seven had it as a first stage, and five as a second stage after 'Understanding the problem'. Therefore, it is included early on in the draft methodological framework to mirror its position in the included frameworks and because of its importance in guiding the development of the conceptual model, as discussed below.

Brassington states that objectives should be agreed and set in writing early on with clients (15). Gentile reiterates this saying that this stage is essential and objectives are important in defining model content and endpoints (172). Robinson confirms this; objectives are key and drive the modelling process (43).

Brassington suggests that objectives should focus on key questions needing to be answered (15). Chwif advises that objectives are often linked to performance measures (21) and van der Zee states that objectives are linked to the project problem (44). Gross maintains that different goals require different model structures and level of detail (20).

Abdelmegid, Furian and Robinson split objectives between general objectives and modelling objectives (12, 18, 43). General objectives include wider project objectives such as time frame and visualisation, and should be kept in mind when deciding on complexity of the model. Modelling objectives are what it is hoped the model will achieve, such as increasing throughput or reducing cost and also includes modelling constraints such as budget. Robinson adds that

modelling and general objectives should be agreed with the client so that expectations can be managed.

Montevechi, Pereira and Tako all suggest the use of soft systems methodology to set the objectives; explaining that the five steps in developing a Performance Measurement Model will generate the objectives (13, 175, 177).

Phase 2 – Model content; what is included and excluded

This middle phase includes 10 stages identified from the included methodological frameworks; model scope, model detail, model outputs, model inputs, model content, entities, system behaviour, stressors, model structure and assumptions and simplifications. These are discussed one-by-one below, in the order in which they should be considered in the draft methodological framework.

Model scope

Ten methodological frameworks include 'Model Scope' as a stage. Five included 'Scope' as the first or second stage in the framework, the remaining five methodological frameworks position 'Model scope' in the second half of the framework.

Once the problem situation is understood and goals or objectives for the conceptual model set, the scope of the model should be considered. The scope will guide the development of the conceptual model. It is this stage that the modeller decides what is included and excluded from the bigger picture of the problem situation to answer the research question.

Many of the methodological frameworks that include the 'Model scope' stage give definitions of scope; Chwif; the extent of the project, for example restricting it to only certain parts of a manufacturing process (21). Furian; which entities will be included in the model (18). Gentile; 'bounding the decision problem', what is included, excluded and outside the scope of the conceptual model (172). Gross; bounds of system and what components should be included (20). NSW defines the scope as boundaries, and Robinson describes

defining the scope as deciding which entities, activities, queues and resources are included in the conceptual model (43, 178). In summary the scope of the conceptual model is what is included and excluded.

Abdelmegid, Gross and Robinson recommend that the modeller and the project team decide the scope of the project and document this (12, 20, 43).

Robinson and van der Zee consider similar methods for this stage (43, 174);

1. Identify the conceptual model boundary

2. Identify all the components in the real system lying within that boundary

3. Assess whether to include or exclude the components identified. Each component should be assessed for importance in 'validity, credibility, utility and feasibility' of the model in this step.

I think the 'relevance' of the components to the research question is an important factor here too.

Model detail

Once the breadth of the conceptual model is decided in the 'Model scope' stage above, the depth of the conceptual model should be considered. This is the depth or detail the conceptual model goes into. Six frameworks include this stage, two frameworks position it in the first half of the framework, the remaining four in the second half of the framework. It is always included as a step within the framework that includes two or more stages, so these positions are open to interpretation.

Descriptions of 'Model detail' include; Chwif; the depth of the model is subtle (21). Pace; whether the system is represented at a higher/broader or drilled down to greater detail (176). Furian explains that looking at the detail of the conceptual model may mean that elements of it are simplified to a higher level (18).

Robinson suggests that the modeller and project team should decide on the detail and then document the level of detail for each component in the conceptual model (43). The level of detail should be determined for each component depending on its effect on the validity, credibility, utility and feasibility of the conceptual model. Again, as in 'Model scope' above, I think that the relevance of the level of detail to the research question is important in this stage.

Model content

Eight methodological frameworks include 'Model content'. Three have 'Model content' as a stage in the first half of the framework (13, 15, 177), the remaining five frameworks position it in the second half of the framework.

I include 'Model content' as an umbrella 'catch all' stage that includes all the components and their relationships within the conceptual model. Stages that related to the content of the conceptual model but were not clearly linked to a particular stage were included in this stage, along with specific 'Model content' stages.

The methodological frameworks which included less specific details were: Brassington - ecology based steps relating to components and their relationships were included in this stage (15); Hesch - 'build conceptual model' (19); Gray - 'allow the model to influence the research design' (16); Montevechi's step includes inputs, outputs and activities (13); Pace includes entities and process (176); and Pereira's step is 'construction of the conceptual model' (177).

Robinson and Tako are more specific and label the step 'determining model content' and 'model content' respectively (43, 175).

Montevechi and Tako recommend the use of soft systems methodology to determine the contents of the model (13, 175). Tako also recommends using a patient flow diagram to establish contents, however, depending on the research question this may not always be relevant.

Appendix 8

Model outputs

Seven methodological frameworks include 'Model outputs'. Four frameworks position the 'Model output' stage in the first half and the remaining three place it in the second half of the framework.

Furian describes model outputs as responses which can be numerical (means) or streamed data (time series). Montevechi defines model outputs as the output from the model's system transformation process (patient to treated patient).

Abdelmegid, Furian, Montevechi, Robinson, Tako and van der Zee all state that model outputs are linked to objectives; they are used to evaluate whether the objectives of the model have been met, and if the objectives have not been met, why not. Tako suggests that once objectives are set, the next stage should be determining the inputs and outputs.

Abdelmegid, Furian, Robinson and van der Zee all recommend considering how the outputs should be presented (tables or graphs).

Both Montevechi and Tako suggest identifying inputs and outputs using soft systems methodology, in particular Performance Measurement Model.

Model inputs

Eight methodological frameworks included 'Model inputs' as a stage, the majority (five) placed this stage in the middle of the framework. Model inputs are also called experimental factors. They are the components of the model that can be altered to represent different scenarios.

Some frameworks included descriptions and examples of Model inputs; Furian and Robinson state that inputs are also known as experimental factors; they can change in different scenarios and should a range around them should be included. Furian adds that in healthcare inputs examples may include policies such as best way to dispatch orderlies. Montevechi gives examples of Model inputs; workstations, machines and individuals, continuing that objectives affect how inputs are assessed. Robinson and van der Zee both state that inputs are

either qualitative or quantitative and their ranges should be determined. Robinson goes on to say that the modeller should vary inputs the client has little control over, adding that identifying inputs is an iterative process.

Three methodological frameworks suggest methods for identifying conceptual model inputs; Tako suggests that stakeholders should be consulted on how much the inputs can vary, Robinson states that identification is driven by objectives, Tako and Montevechi suggest the use of SSM. Tako specifically advises the use of a PMM, Montevechi suggests using CATWOE mnemonic components 'customers' and 'actors', these are individuals who will use resources such as workstations and equipment; from this inputs are identified.

Whilst I feel it is important to identify which components are inputs at this stage, I would argue that it is not appropriate to be deciding numerical values and ranges, this would relate to a mathematical model, not a conceptual model. Taking into account the descriptions of inputs above I would argue that inputs for the purposes of my methodological inputs relate to components at the start of the system or process, such as treatment or patient.

Model entities

Five frameworks include the stage of 'Entities', all but one (NWS) come in the second half of the framework.

Abdelmegid describes entities as core components of the model structure. Furian describes two types of entities; active and passive. Active can change their role and passive are fixed and not related to the flow of a system. NSW also describes different entities; they can be 'pressure, state or response'. Pace suggests identifying entities needed to achieve the objectives. For the purpose of the proposed framework I will be calling entities 'components'.

Appendix 8

System behaviour

Eight methodological frameworks included a system behaviour stage, mainly in the middle of the framework.

The system behaviour is how the components of the system are linked, or relate, to each other. Furian reports that the components included in this stage are control units (combined entities), events and activities. Abdelmegid comments that 'control units represent different levels of decision making within the system' ('determining the set of rules and their relations'). Gentile links end-points to physical characteristics and stressors in their system behaviour stage.

Gross recommends identifying components of the system and how they are linked, this can be done by breaking systems down into habitats for the ecology discipline. NSW and Pace both reflect Gross' recommendation of identifying relationships between components of the model. Gross adds that questioning the system and suggesting alternative hypotheses can help develop an understanding of the system behaviour.

Tako suggests using SSM methodology, specifically the CATWOE and root definition, to define system behaviour. However, it could be argued that this is methodology that would aid in the 'Understanding the problem' stage as well as the 'System behaviour' stage, applying it to understand the bigger or rich picture.

Model stressors

The three methodological frameworks containing 'Stressors' position this stage in the first half, the second half and right in the middle of the framework.

All the methodological frameworks including this stage were from the ecology discipline. Stressors are specific to ecology and are the threats to, or constraints on, the ecosystem. They are mainly related to human activity, (pollution, timber extraction and hunting), but can also be natural (fires or storms). Gross advises that only the stressors relating to the project are included, and should be kept simple.

Appendix 8
Model structure

Two methodological frameworks include 'Model structure' as a stage, and it is placed half way through both of these frameworks. Furian states that this stage is closely linked to the Entity stage because the Model structure is characterised by the entities and their aggregation. Abdlemegid states that designing the model structure begins with defining the entities. For the purpose of my proposed framework I will use the description of 'Model layout' for this stage.

Assumptions and simplifications

Five methodological frameworks have 'Assumptions and simplifications' as a stage and all of these are in the second half of the framework.

Whilst I do not consider this to be a separate discrete step in the framework it is important. It can be applied to 'Understanding the problem' if when understanding assumptions and simplifications are used to gain an understanding of the problem. However, assumptions and simplification are mostly applied to the model contents so I will include it in this stage of the proposed framework. At each stage of considering the model content the modeller will make assumptions about the contents and how the components interact. The modeller will also need to make simplifications to make the conceptual models understandable and not overcomplicated.

Chwif and Robinson explain that assumptions relate to a limited knowledge or uncertainty of the problem situation/real world, and simplifications result from keeping the model simple and easy to use. Chwif adds that the more assumptions made, the less detail included in the conceptual model. Robinson explains that when identifying scope and detail of the conceptual model assumptions and simplifications will be made. Adding that assumptions and simplifications are referenced by components and detail omitted from the model, Abdelmegid adds that the assumptions and simplifications define the scope and detail of the model. Robinsons suggests that the impact of assumptions and simplifications should be assessed in agreement with the stakeholders. Van der Zee adds that assumptions and simplifications are

determined regarding scope and detail, and the effect on outputs should be considered. Chwif recommends that assumptions are recorded.

Phase 3 – Documentation and validation

This final section includes documenting and validating the conceptual model after it has been developed. Four stages are included in this section;

- Diagram
- Documentation
- Validation
- Review, revise, refine

Diagram

Five methodological frameworks include a 'Diagram' stage, all but one of these have it as the last stage of the framework.

Brassington advises including a diagram with the documentation. Gentile explains the diagram visually links the process and components. Gray states that the diagram acts as a check for the modeller; the detail in the diagram should match the detail in the conceptual model, the modeller should ask themselves 'does the diagram leave the user making assumptions?' Montevechi explains there is no formal way of producing visual representation of conceptual model, and adds that the diagram should be used as a communication tool and can take the form of a chart, image or diagram, and that using SSM can help the development of the diagram. Van der Zee comments that visualisation is useful for validation and relevance.

Appendix 8
Documentation

Six methodological frameworks included 'Documentation' as a stage, five of these had this stage in the second half, of these, four had it as the last stage (or in a step that was the last stage).

The first stage in this phase is to produce a diagram of the conceptual model, this is closely linked to producing a document of the conceptual model, they should be presented together. The document should include explanations, assumptions, simplifications and a description of the model, this is closely linked to the diagram stage above.

Brassington recommends that once the conceptual model is agreed by stakeholders it should be documented as a written description. Chwif suggests the document should include any conceptual model revisions. Gray advises documenting the conceptual model diagram, including justification of the choice of illustration. NSW proposes the document should record sources of evidence, key questions, assumptions and limitations. Van der Zee suggests to document the conceptual model and to justify scope and detail.

In my draft methodological framework I will combine the diagram and document stages as they are so closely linked.

Validation

Four methodological frameworks include 'Validation', all of them position this stage in the second half of the framework. Two of these methodological frameworks have 'Validation' as the final stage.

FEFLOW suggests that if the conceptual model is not suitable the modeller should revisit earlier stages, rebuild the conceptual model and revisit this stage. Gray suggests that the modeller uses the conceptual model as a communication tool in this stage and get feedback from the stakeholders. It is likely that the stakeholders will have a different perspective of the problem than the modeller and should sense check the conceptual model. Gray adds that this stage is important and should be informal. Feedback from the stakeholders should be

used to refine and strengthen the supporting arguments for the conceptual model. Gray goes on to say that this stage is important to confirm whether the conceptual is robust and will stand up to scrutiny from experts. Montevechi suggests this stage should be done in parallel with other stages whilst developing the conceptual model. Montevechi adds that this stage is used to test the theories and assumptions underlying the conceptual model, that it is suitable for its intended purpose, and that if SSM methods were used to develop the conceptual model the documented methods will act as a validation method. Pereira advises that if the model is not validated the modeller should revisit the development of the conceptual model. Pereira suggests that validation is done in conjunction with experts.

Review, revise, refine

Three methodological frameworks include this stage, two include it as the last stage and the other second to last.

Gray suggests revisiting the conceptual model when the results from the project are available, assessing whether these results support the original conceptual model and if any modifications are needed. Gross believes that conceptual models are an 'incomplete abstraction of reality' and will need revisions as new evidence is available and goals alter. The conceptual model should be reviewed periodically to ensure it reflects current knowledge. NSW suggests revising the original conceptual model; it was based on the best knowledge at the time of development, however, more up to date evidence come to light.

Appendix 9: Methodological framework diagrams from identified studies and early drafts of diagram to include in final methodological framework (*Chapter 5)*

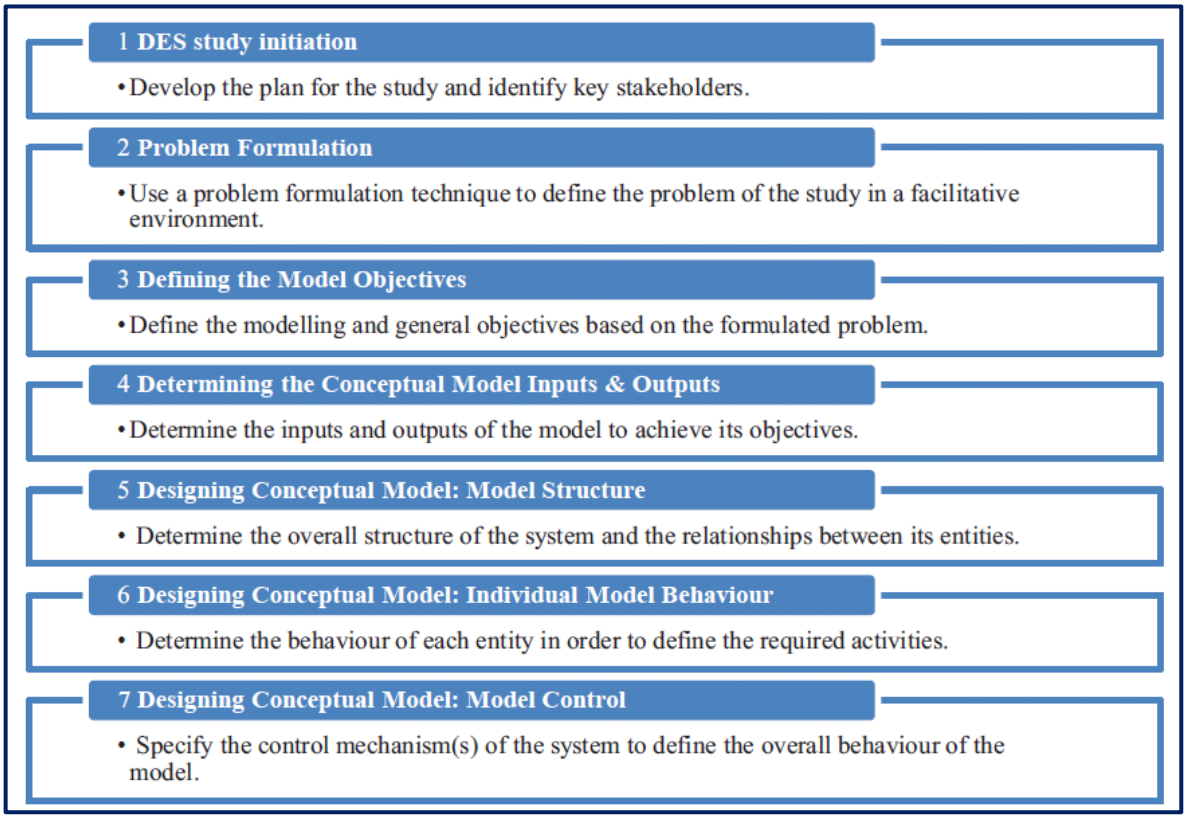


Figure A9.70: Abdelmegid et al. methodological framework diagram (12)

Appendix 9

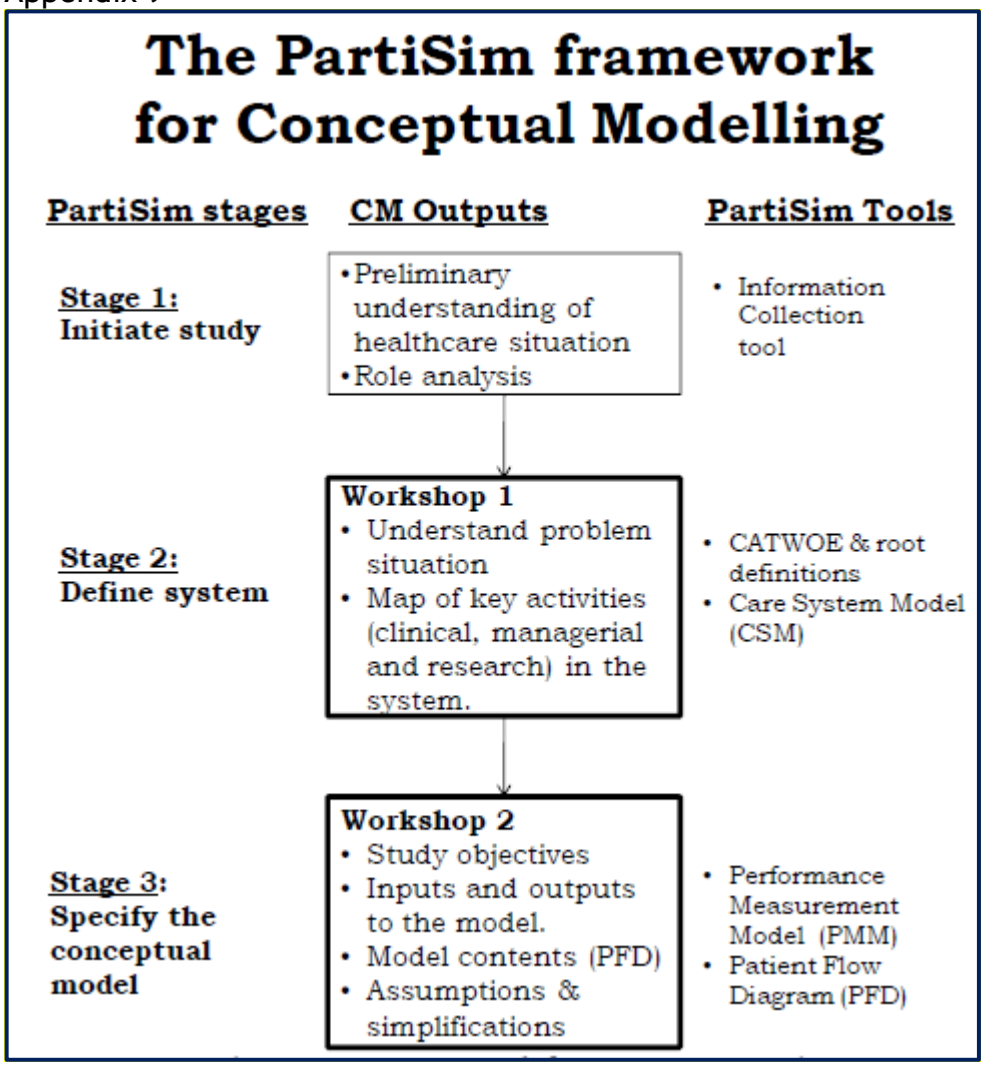


Figure A9.71: Tako et al. methodological framework diagram (175)

Appendix 9

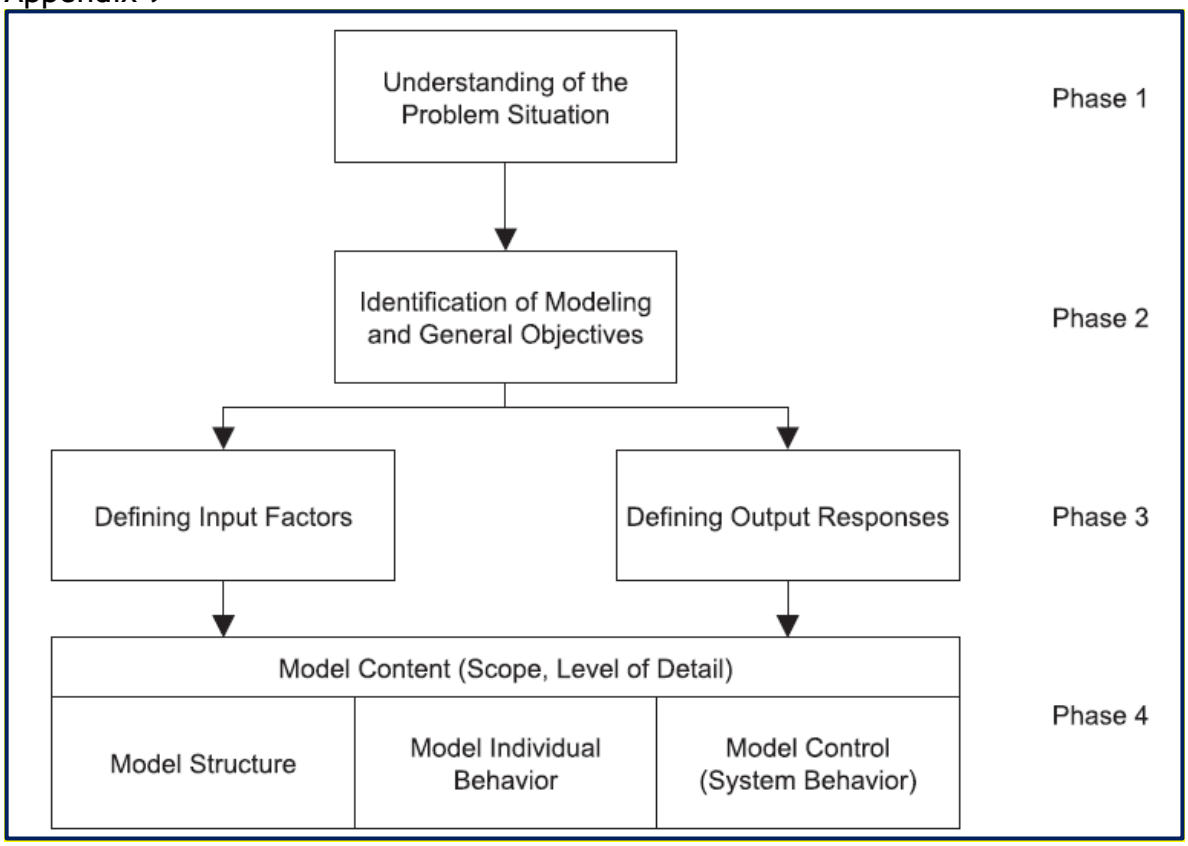


Figure A9.72: Furian et al. methodological framework diagram (18)

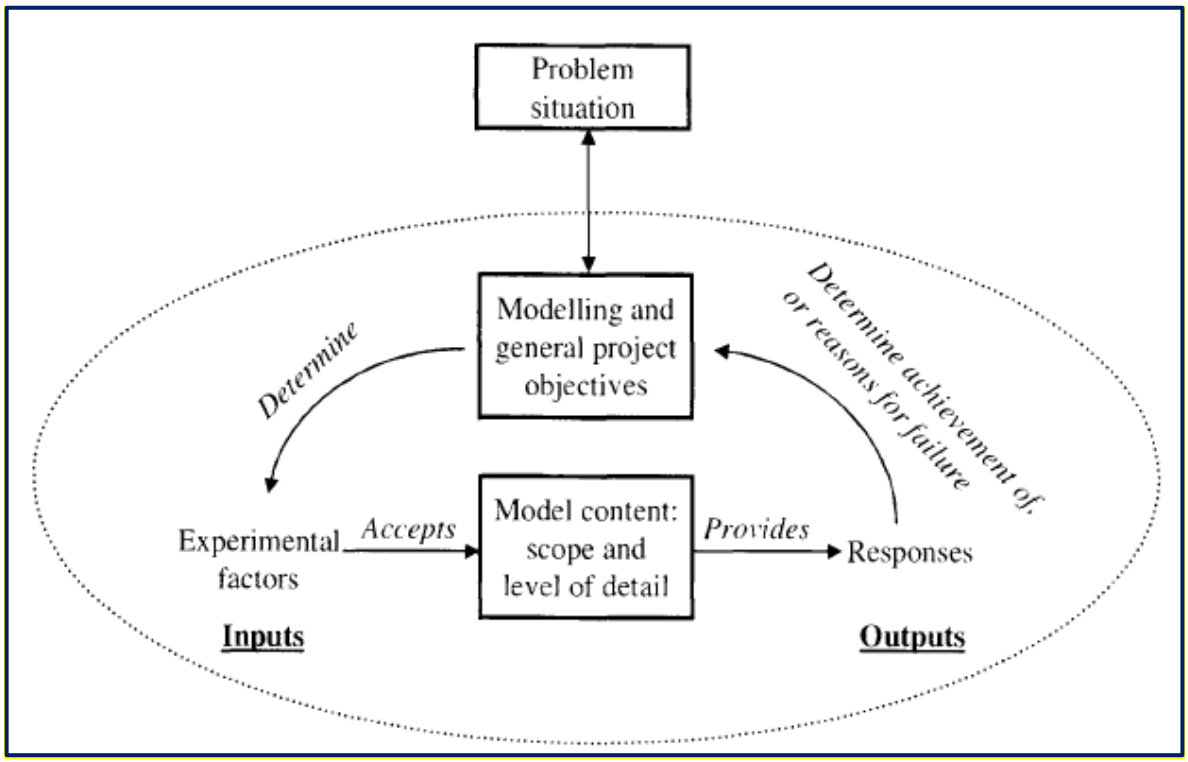


Figure A9.73: Robinson methodological framework diagram (43)

Appendix 9

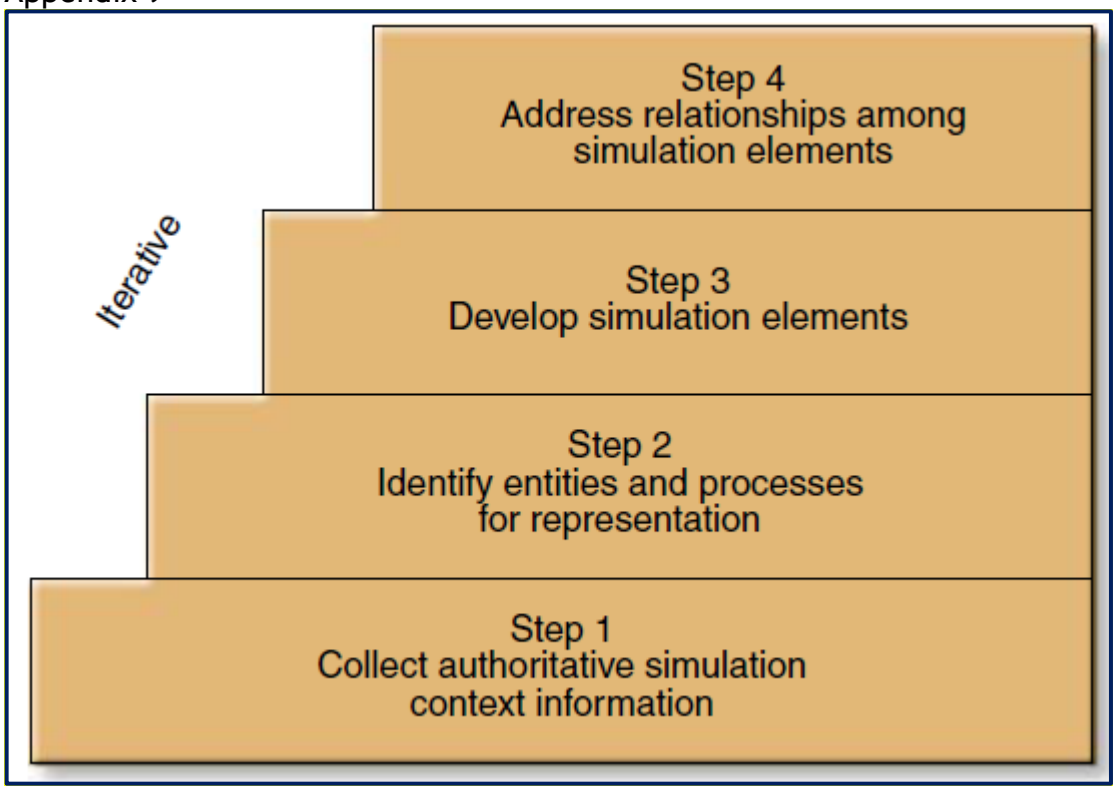


Figure A9.74: Pace methodological framework diagram (176)

Appendix 9

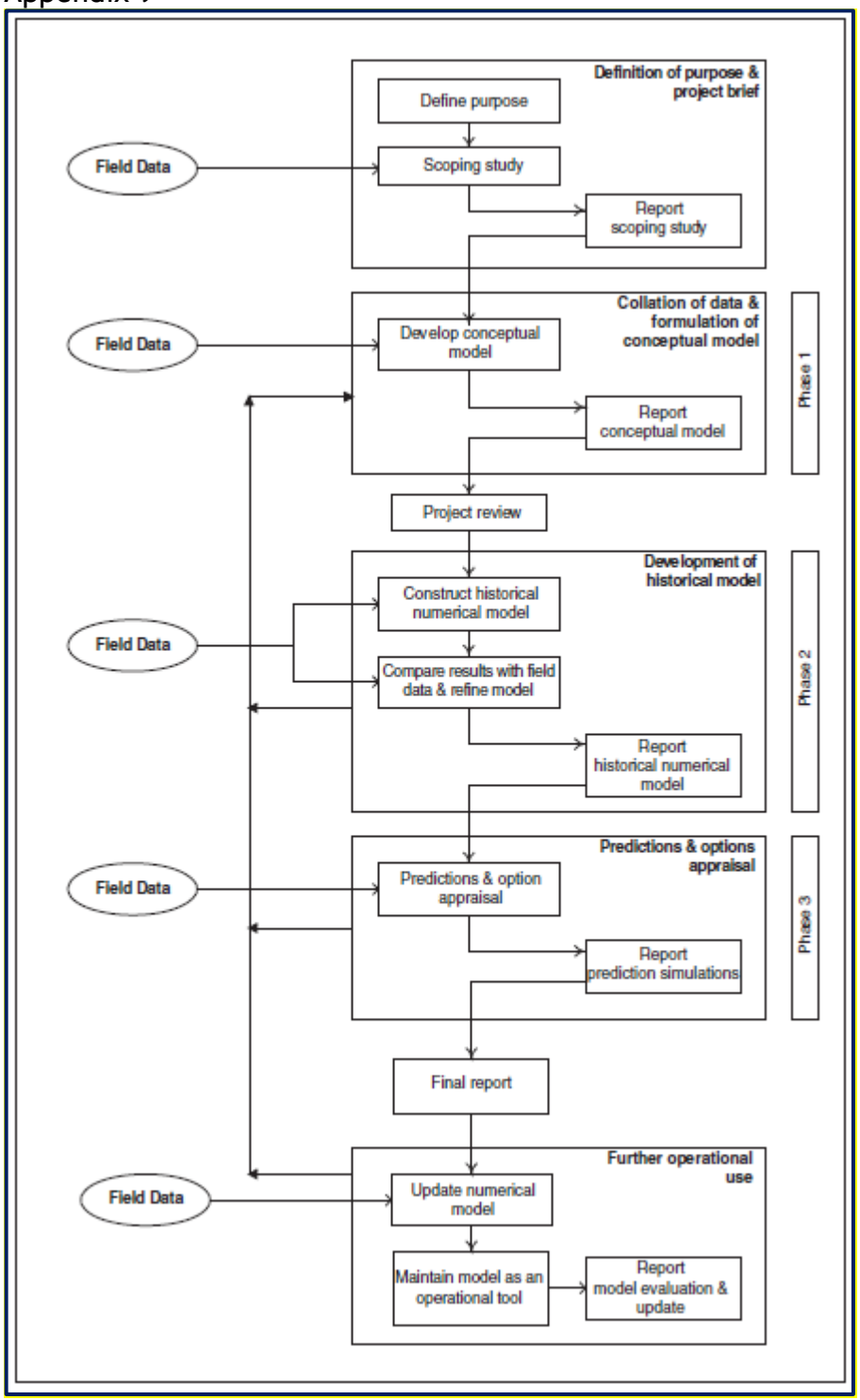


Figure A9.75: Brassington et al. methodological framework diagram (15)

Appendix 9

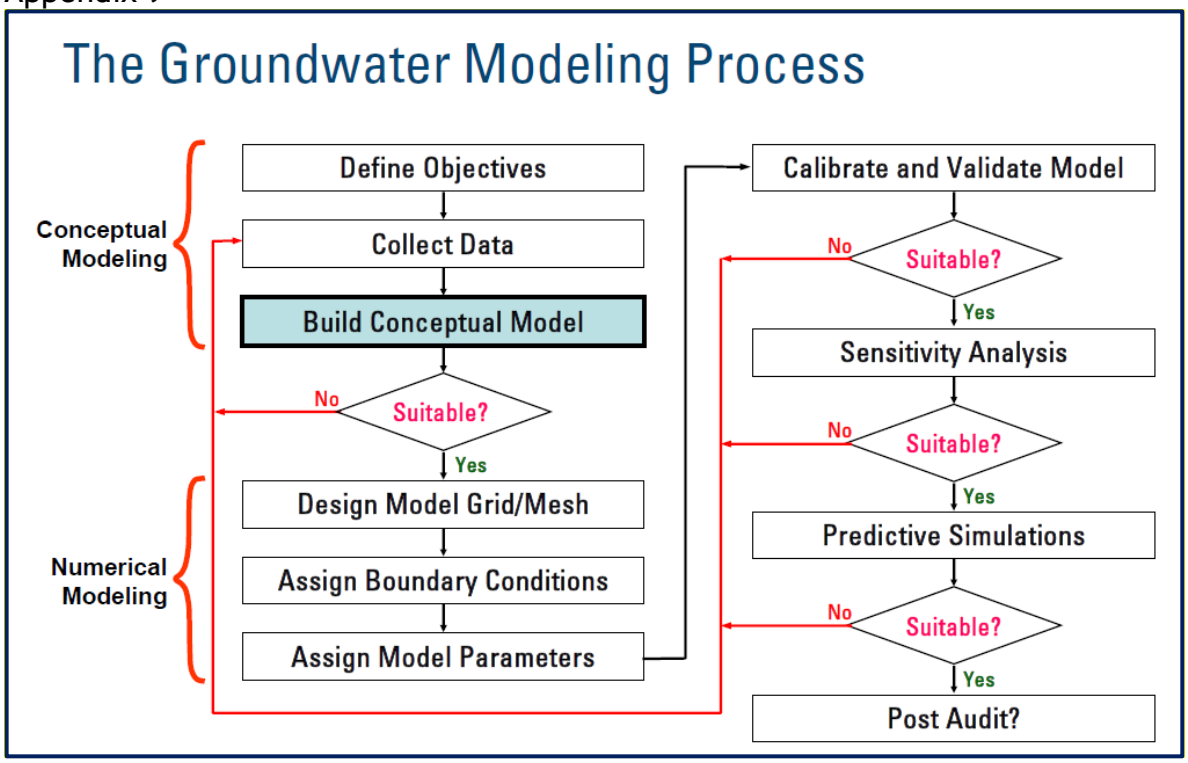


Figure A9.76: Hesch FEFLOW methodological framework diagram (19)

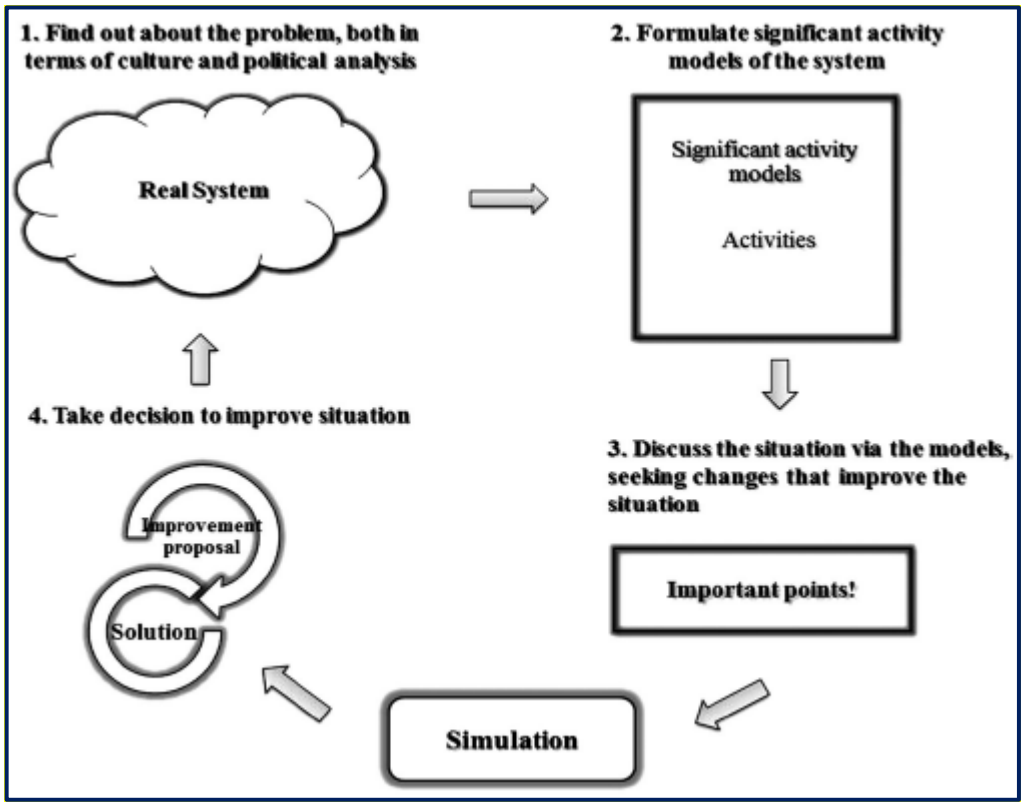


Figure A9.77: Pereira et al. methodological framework diagram (177)

Appendix 9

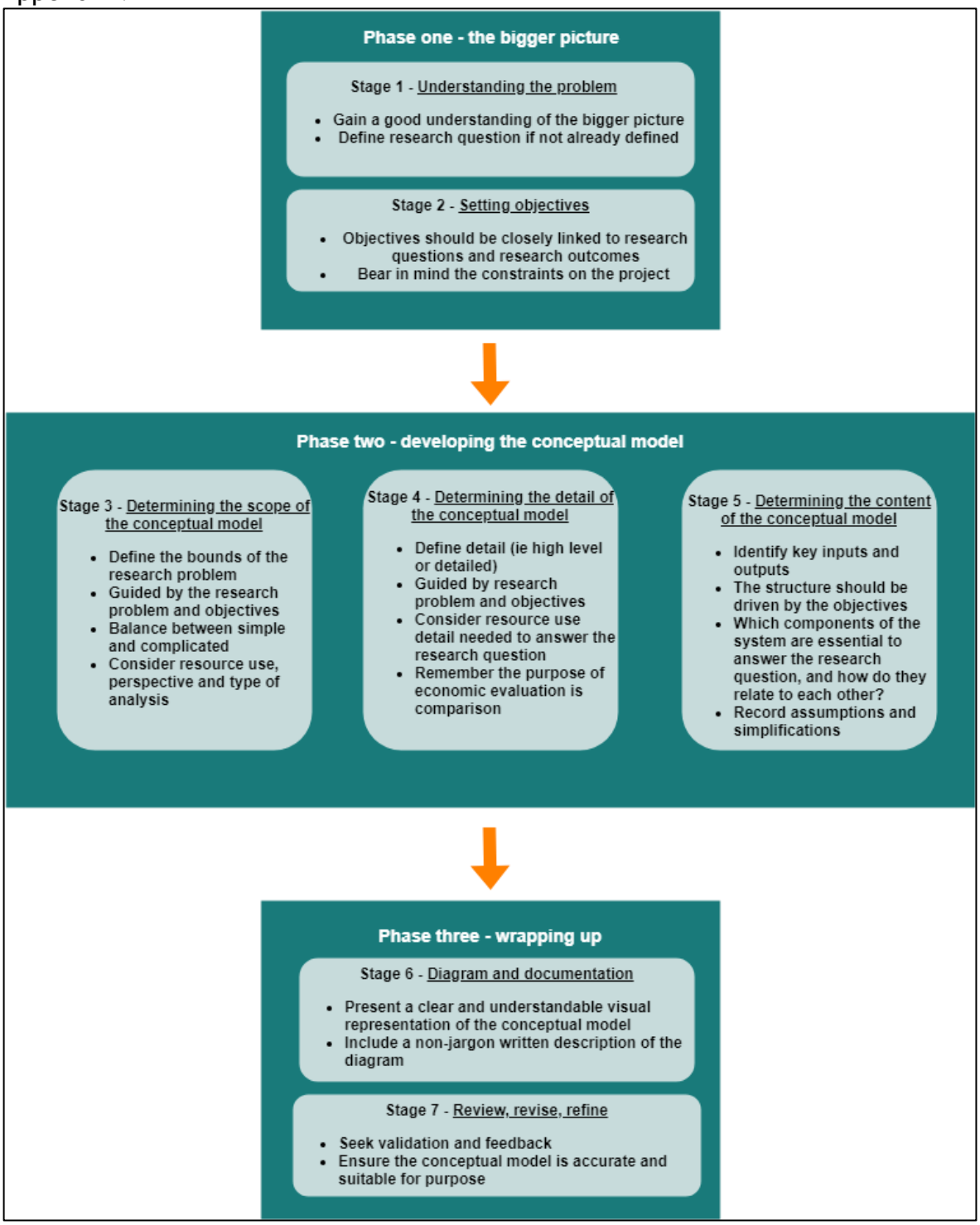


Figure A9.78: Potential methodological framework diagram v1

Appendix 9

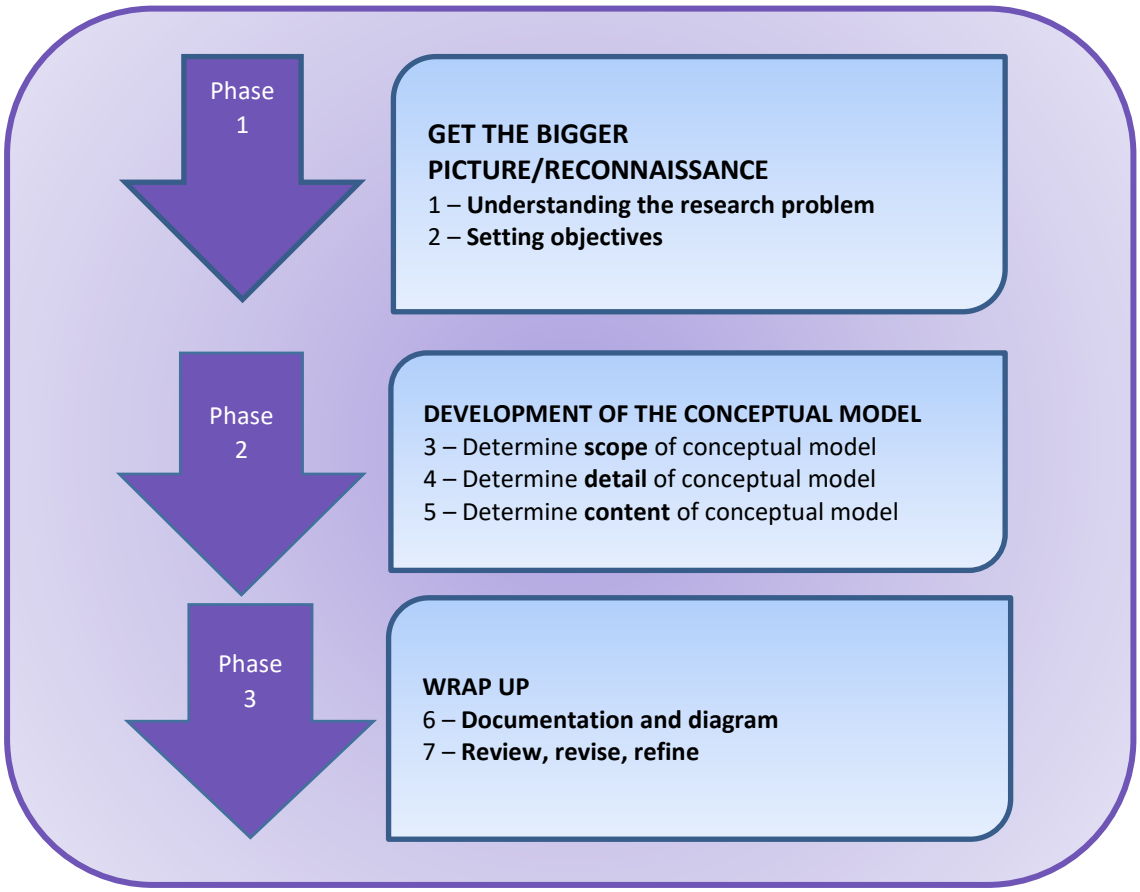


Figure A9.79: Potential methodological framework diagram v2

Appendix 9

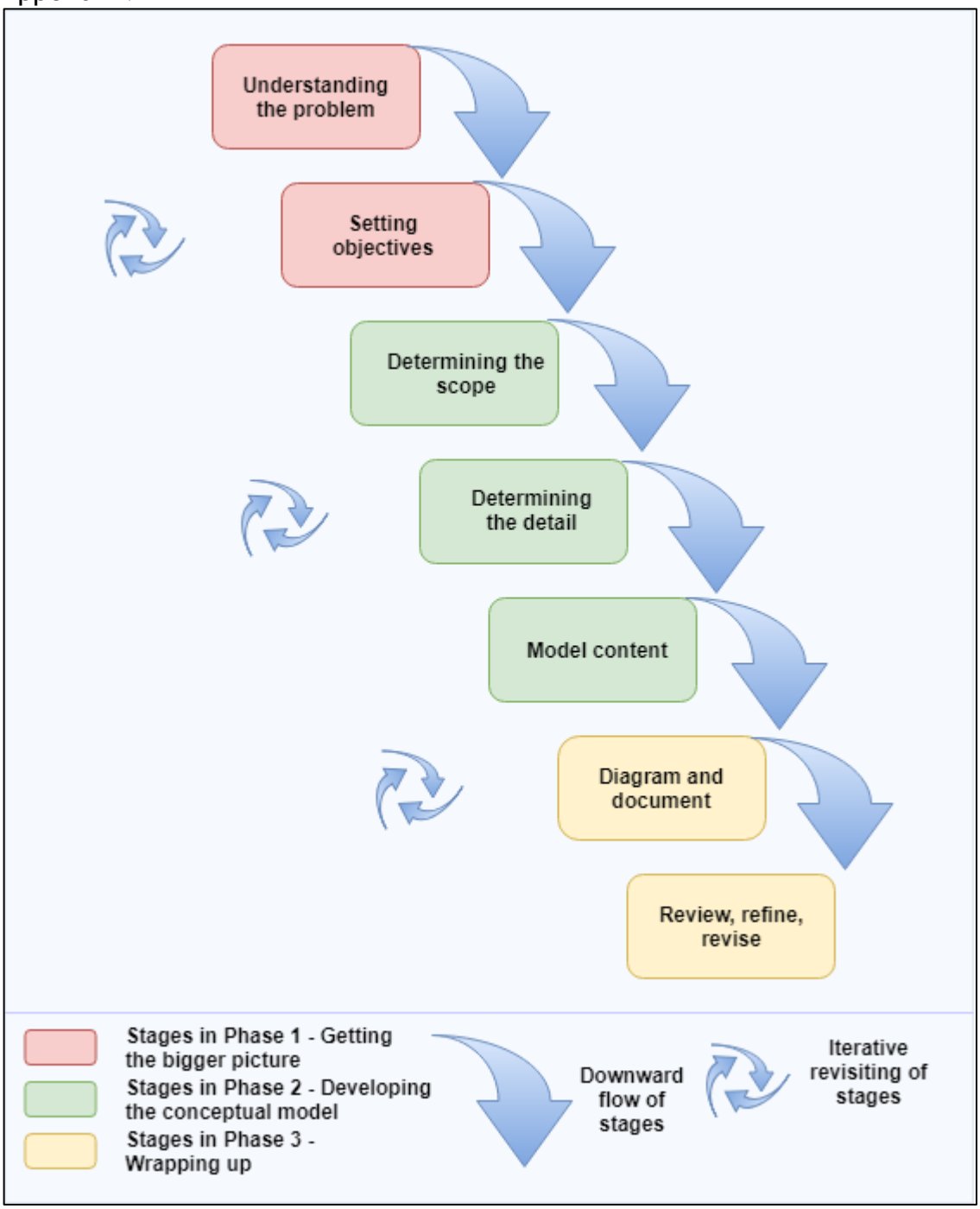


Figure A9.80: Potential waterfall diagram for methodological framework

Appendix 9

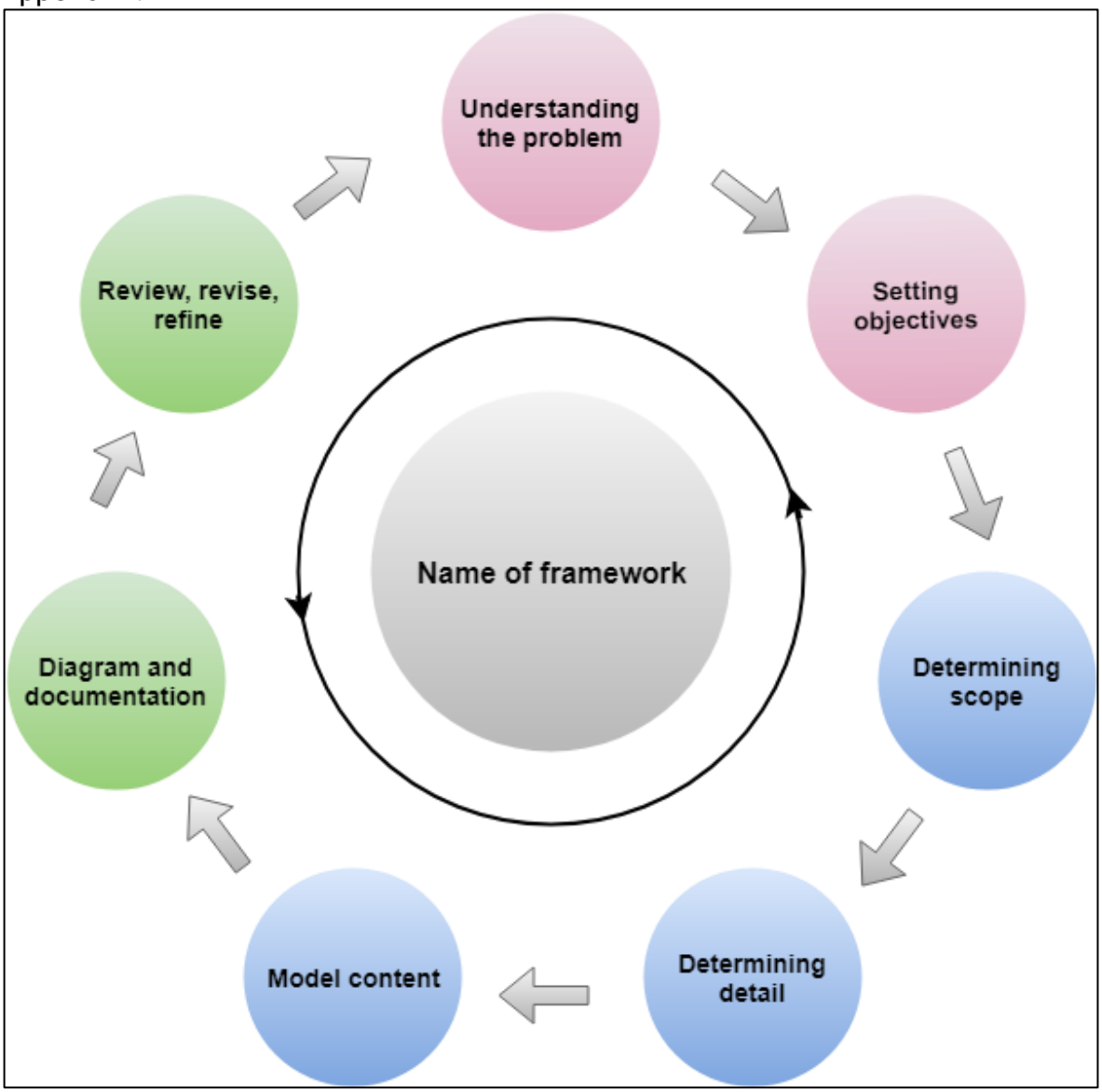


Figure A9.81: Potential circular methodological framework diagram

Appendix 10: Draft methodological framework (*Chapter 5)*

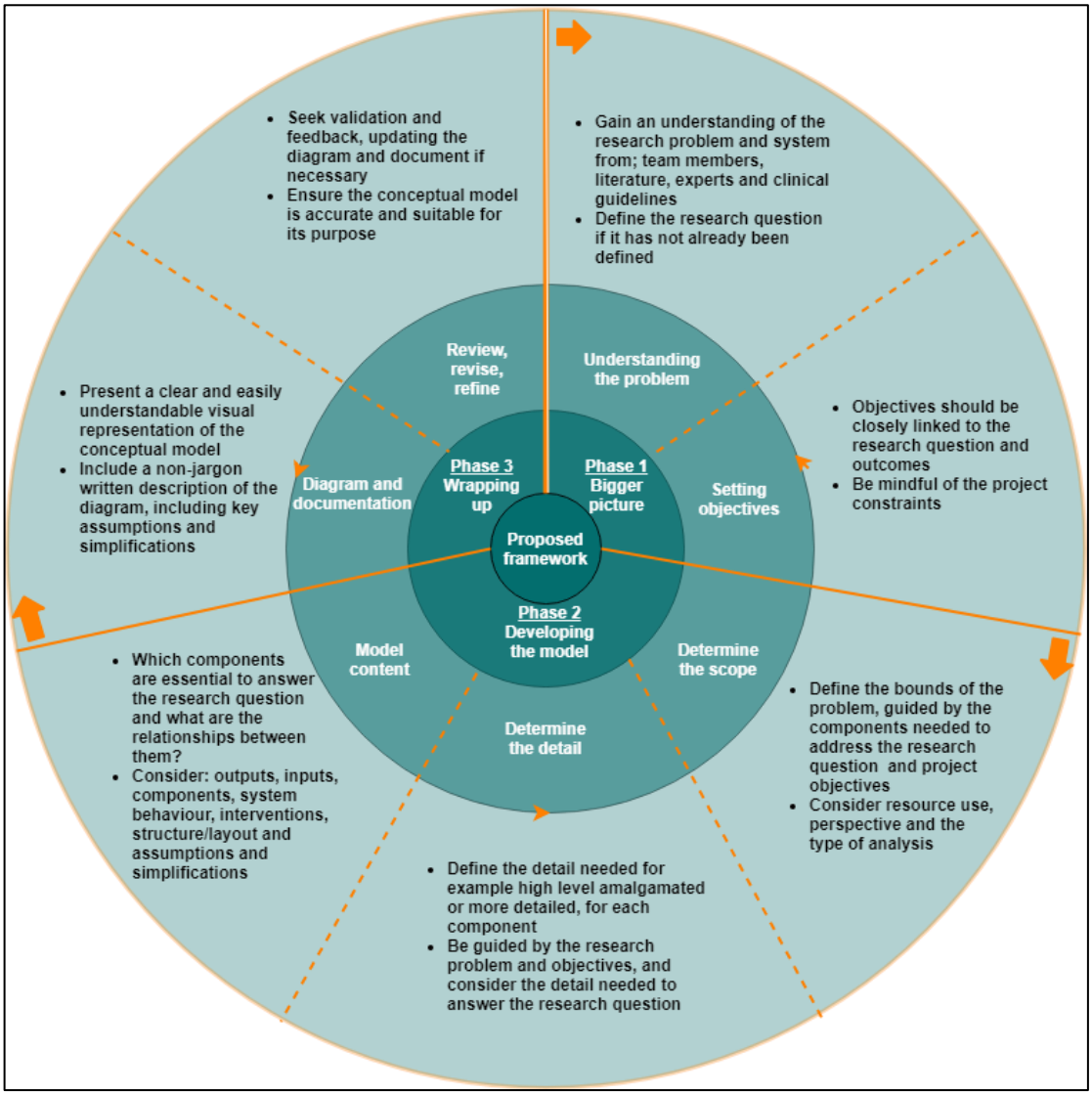


Figure A10.82: Draft methodological framework diagram

General advice

This proposed framework is a generic guide to developing conceptual models for use in economic evaluations. It should be applied to each research project in a pragmatic way and tailored to each situation.

The framework is based on two key concepts. These two key concepts are fundamental in developing all conceptual models: the iterative process and keeping the conceptual model simple. The modeller iteratively revisits stages in

the framework until the conceptual model is completed. The conceptual model should be kept simple; it should contain enough detail to answer the research question, but not too much that would make it unwieldy, containing unnecessary detail. Both of these elements should be borne in mind by the modeller throughout the conceptual modelling process.

This conceptual modelling framework is intended for when the main/key purpose is to develop a conceptual model; for guiding the analysis, interpretation and presentation of results.

Because the research theme and question may be predefined (as is often the case in clinical trials), some aspects of the proposed framework may not be relevant to all circumstances.

The framework is split into three phases: Getting the picture, Development of the conceptual model and Wrapping up. Within these phases are discrete stages, making up the steps of the framework. The outputs of this framework are a diagram and document depicting and explaining the conceptual model (Figure A10.82).

Throughout the conceptual model development process the modeller will add information to a document, this will be described in more detail in Stage 6 ‘Diagram and documentation’.

Phase 1 - Getting the picture

In this phase the modeller immerses themselves in the research problem, gaining an understanding of the problem and research question, and setting objectives.

Stage number 1: Understanding the problem

Recommendation

The modeller should gain a good understanding of the bigger picture of the research problem. This involves looking at the wider systems and subject area in context to the search problem. If the research question has not already been

defined it should be determined during this stage. Without a sound understanding of the research problem, it is not possible to develop an accurate conceptual model. The problem may require identifying patient pathways, patient behaviour or disease pathways. The project team should also be identified during this stage if they have not been identified already.

Explanation

Understanding can come from literature, stakeholders, decision makers, experts, existing conceptual models, trial protocol (if relevant) and clinical guidelines. Questions useful to help in understanding the problem are presented in Table A10.26. These questions are split into two sections; understanding the problem and understanding the systems within the problem area.

Specific elements of soft systems methodology can be used in this stage to come to an understanding of the problem systems. These elements of soft systems methodology involve developing a rich picture (an informal, often hand drawn diagram of the problem area) and using the CATWOE mnemonic to understand important aspects of the problem system. The CATWOE mnemonic is a checklist used to identify the purpose and key elements of a system. The rich picture and CATWOE mnemonic are used to produce a description of the system.

The makeup of the project team will vary depending on the research question. The project team will typically comprise; the modeller, other health economists, clinicians with an expert knowledge of the disease area, patient representatives and study or trial team members (i.e. trial manager). Whilst it is good practice for the project team to be involved at this stage, it may be practical to only include them at the later 'Review, revise, refine' stage. If the project team are involved at this stage the project team and the modeller should come to a consensus on the understanding of the research problem; the modeller can suggest alternative hypotheses to the project team at this stage to help reach an understanding of the research problem.

A structured description of the research problem and system (including activities within the system) should be developed to include in the conceptual model document. The modeller will need to make assumptions in coming to an

understanding of the problem; any key assumptions made in reaching this understanding should also be recorded in the conceptual modelling document.

Table A10.26: Questions to help understanding the problem

Questions to guide understanding the problem
What is the problem?
Why is it a problem?
Who are the target population?
What are the treatments to be included?
What are the outcomes?
What is the policy context?
Questions to guide understanding the problem systems
Who benefits in the system? (ie patients)
Who carries out the activities in the system? (ie clinicians)
What is the purpose of the system?
How does the system fit into the bigger picture?
Who has formal power over the system?
What constraints are there on the system?

Stage number 2: Setting objectives

Recommendation

The objectives of the conceptual model should be set in this stage. Setting objectives is key in guiding the development of the conceptual model and defining the model outputs and content.

Explanations

The objectives will be closely linked to the research question(s) and what the model should achieve. The objectives will be related to outcomes, including hospital services and throughput, disease progression or patient behaviour. The modeller should ask themselves; what are the project performance measures or outcomes? Can these performance measures or outcomes be broken down further?

The modeller should also bear in mind the constraints on time and budget of the project; making sure the objectives are realistic and expectations are managed.

The modeller should agree the objectives with the model team (if applicable) and record them in the conceptual model document.

Phase 2 - Development of the conceptual model

In the second phase the modeller decides what to include and exclude in the model, which components best represent the research problem, and how these components are linked. During this phase the modeller should keep in mind the objectives set in Stage 2 and the project hypothesis if there is one, to ensure the conceptual model is relevant to the research question. The development of the conceptual model will be an iterative process, with the modeller revisiting stages until it is complete.

Stage number 3: Determine the scope of the conceptual model

Recommendation

In the first stage of this phase the modeller determines the scope of the conceptual model. The scope bounds the research problem, restricting the conceptual model to only the elements of the bigger picture needed to address the research question.

The modeller should see the 'Understanding the problem' stage as getting to know the 'bigger picture' of the problem; understanding the wider subject area that includes the research problem. The modeller then decides on the narrower scope of the conceptual model in the 'Determine the scope of the conceptual model' stage. This guides the development of the model, helping the modeller decide what should be included and excluded from the bigger picture to answer the research problem. The relationship between the bigger picture and the scope of the conceptual model is illustrated in Figure A10.83.

Appendix 10

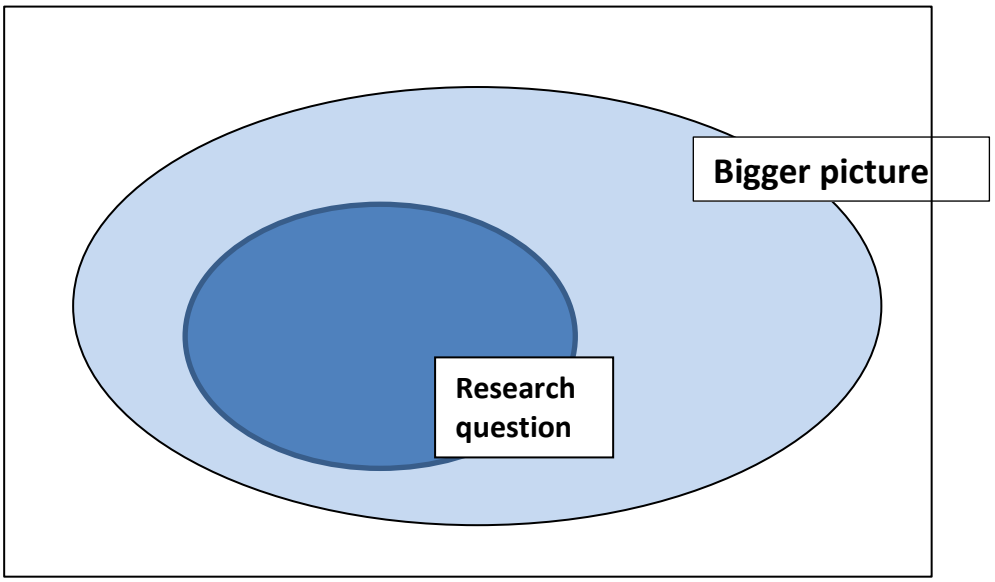


Figure A10.83: Bigger picture v. research question

Explanation

When determining the scope the modeller should concentrate on the research questions, project hypothesis (if there is one) and conceptual model objectives. Only components important and relevant to the research question, hypothesis and objectives are included. If components that are not relevant to the research question etc are included in the conceptual model it will become too complicated and unwieldy, creating 'noise' and making interpretation difficult. If there is a trial protocol available this can also help determine the scope of the conceptual model.

The modeller should consider resource use, the perspective taken for the costs and outcomes and the type of economic evaluation (i.e. cost-utility, cost-effectiveness or cost-benefit), as these considerations will determine components to be included in the scope of the conceptual model.

Each of the components identified as within the scope of the model should be assessed by evaluating the relevance of these components to answering the research question. The modeller should bear in mind that the purpose of economic evaluations is the comparison of health technologies and should consider which components are important in completing this comparison.

The modeller and project team (if appropriate) should come to a consensus on the scope of the research problem and record this in the conceptual model document.

Stage number 4: Determine the detail of the conceptual model

Recommendation

Once the scope of the conceptual model has been determined the depth of the conceptual model should also be considered. The depth determines how far the modeller will drill down into individual elements of the conceptual model. Depending on the objectives the detail may be high level and simplified or could involve precise details of all or some of the components.

Explanation

The modeller should determine the detail in relation to the research question, project hypothesis and conceptual model objectives. The modeller should bear in mind resource use data needed to answer the research question and the detail needed to capture this. The modeller should take care to only include detail relevant in testing the hypothesis and comparing the health technologies.

Similar to 'Determining the scope of the conceptual model' above, the detail of each component should be assessed on their relevance and validity in answering the research question.

The decided detail of the conceptual model should be reported in the conceptual model document.

Stage number 5: Determine the content of the conceptual model

Recommendation

The final stage in this phase is deciding on the content of the model. There are several elements that the modeller should consider in this stage. These are included below, but not in any particular order.

Model outputs are important aspects of this stage. Outputs can be used to check that the objectives of the conceptual model have been reached. **Model inputs** should also be considered if appropriate, they are components within the model that can be altered to represent different scenarios.

The **components** within the scope of the conceptual model should be considered, along with how they interact or are linked to each other (**system behaviour)**. Components can either be passive (static) or active within the flow of the system. **Stressors** are the interventions; they alter or are hypothesised to alter the flow of the system. The overall **structure/layout** of the model should be considered. **Assumptions and simplifications** will be used in most of the stages in model development and are also relevant to understanding the research problem, scope and detail. Assumptions relate to limited knowledge or evidence of the research problem bigger picture. The more assumptions made, the less detail included in the conceptual model. Simplifications result from keeping the conceptual model straightforward and uncomplicated.

Explanation

Model outputs will be closely linked to the research question and hypothesis. Key outputs are costs and effects; examples include a patient becoming a treated patient, quality of life measures, clinical effects (such as strokes avoided or cancer cases detected) or patient behaviour such as increase in physical activity. Examples of **Model inputs** include health resources and patients. The **Model structure** will be driven by the objectives and research problem. For example a simple patient pathway may be used, or disease progression, however the structure of any planned mathematical model would not be relevant here. When determining the model structure the modeller should consider resource constraints and capturing resource use.

As in the 'Understanding the problem' stage the modeller can use soft systems methodology to define the root definition of the problem system within the scope of the conceptual model. Again, suitable methods to do this include drawing a rich picture and using the CATWOE mnemonic to develop a description of the system. For example; who benefits in the system? Who carries out the

activities in the system? What is the purpose of the system? Who has formal power over the system? What constraints are there on the system?

In determining the content of the model the modeller should be constantly reminding themselves of the objectives of the project, research question and hypothesis, if applicable.

Model content, and assumptions and simplifications made in determining the model content, should be recorded in the conceptual model document.

Phase 3 - Wrapping up

The final phase of the framework includes diagram, documentation, and review, revise, refine. The conceptual model is presented in a visual form and documented. The conceptual model is then validated and, if necessary, revisions are made.

Stage number 6: Diagram and documentation

Recommendation

A visual representation of the conceptual model should be presented. This should link the components and processes in a clear way. The diagram used should be clear, accurate and relevant to the research problem. As a communication tool the diagram should not leave the user having to make assumptions. For the modeller carrying out the economic evaluation the conceptual model will guide the analysis, interpretation and presentation of the research project.

The document, a non-jargon written description of the conceptual model, should help the reader understand the model, and should be used as a communication tool, along with the diagram.

Appendix 10
Explanation

The diagram should be clear and understandable; it is likely that the modeller will need several iterations to develop the diagram.

The purpose of the document is to help the user to understand the model, it does not need to be extensive or overly comprehensive.

Information recorded at each stage of the conceptual model development included in the document will consist of;

- evidence used in understanding the bigger picture,
- an explanation of the research problem, including funder, population, outcomes and interventions,
- conceptual model objectives,
- descriptions of the scope and detail of the conceptual model,
- the key section of the document should be a description of the conceptual model, describing the components included, the relations between them, and the general ‘flow’ of the conceptual model
- assumptions and simplifications made in the development of the conceptual model; only key assumptions should be recorded, unimportant assumptions and simplifications should be omitted
- finally, methods of validation and any revisions should be documented

A summary of the diagram and document can be included in the Health Economics Analysis Plan, if applicable.

Appendix 10
Stage number 7: Review, revise, refine

Recommendations

The final stage of the framework is 'Review, revise and refine'. In this stage the conceptual model should be validated and, if appropriate, the modeller should seek agreement and feedback on the conceptual model from the project team. The project team should be shown the conceptual model diagram and document to check they reflect the team's understanding of the research question and the research system. The conceptual model should also be checked in terms of logic, presentation and ease of understanding. This is an important step, without which the conceptual model may not be accurate or useful. Any suggested changes during this stage should be used to refine the model, strengthening it and making it robust and suitable for its intended purpose.

Explanations

If changes are made to the conceptual model the diagram and documentation should be updated in this stage. If new information comes to light after the development of the conceptual model the conceptual model should be updated if relevant.

The modeller and project team should check the conceptual model to ensure it accurately depicts the research area and question. Table A10.27 lists questions the modeller and project team can ask themselves during this stage to review and validate the conceptual model.

Appendix 10

Table A10.27: Validation questions

Diagram:
- Is the diagram well defined, logical and transparent? - Does the conceptual model reflect the research question and subject area system?
Document:
- Is there a clear description of the research question, including background information to help understand the problem? - Are there clear objectives reported? - Are the scope and detail considered relevant to the research question and justified? - Is there a clear and understandable description of the content of the conceptual model? - Are the key assumptions and simplifications made explicit? - Does the user need to make assumptions about the conceptual model to understand it? - Does the conceptual model include all the components needed to complete an economic evaluation, accurately representing resource use, outcomes and perspective? - Are changes made to the conceptual model during the 'Review, revise, refine' stage recorded in the document?

Appendix 11: Iterative development of TWICS case study conceptual model diagram (*Chapter 6)*

The first rich picture developed illustrated both the increasing disease severity and progressive nature of COPD and the effects of exacerbations (Figure A11.84).

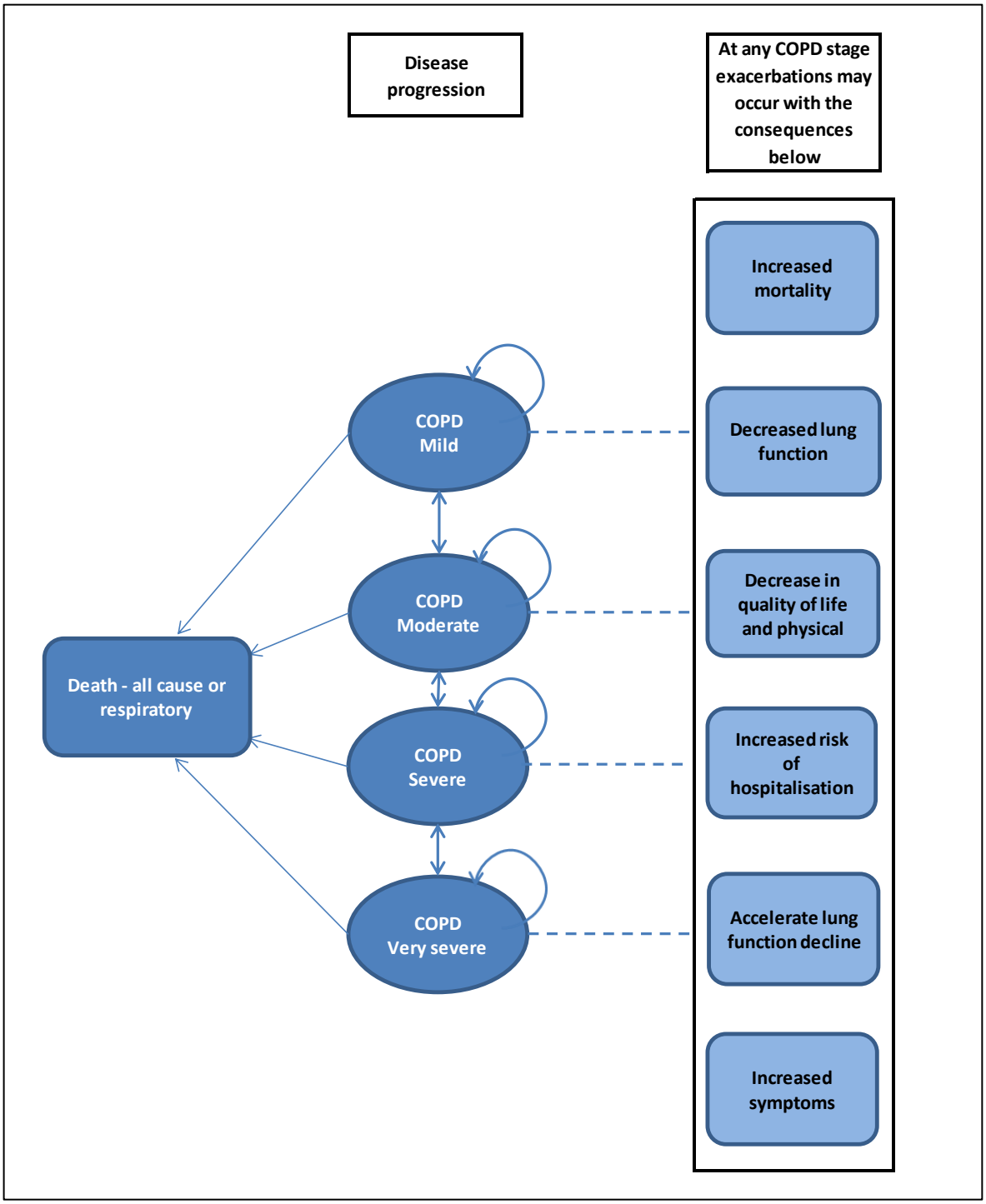


Figure A11.84: Rich picture of COPD

Appendix 11

The first version of the conceptual model is presented in Figure A11.85. I decided to concentrate on the conceptual model depicting the intervention arm incorporating the theophylline intervention initially. Working from the left-hand side of the conceptual model to the right-hand side of the model starting with the COPD patient, the patient is anticipated to experience serious adverse events resulting from co-mobilities, COPD related resource use and exacerbations. The intervention (theophylline) is anticipated to affect the number of exacerbations the patient experiences during the trial, and serious adverse events are unknown as a result of the intervention. Exacerbations are one of the outcomes used in the economic analysis, and therefore is depicted by a beige oval with two outlines and linked by an 'analysis' line to the 'cost per exacerbation avoided'. The third column from the left consisting of all blue 'entity' ovals represent measures that feed into the economic evaluation analysis, such as quality of life and mortality data. The fourth column of beige 'output' ovals depict the two main elements of an economic evaluation; namely costs and effects, in this case the effects are quality adjusted life-years. The final column containing the green rectangles illustrates the analysis measures that were used to report the cost-effectiveness results, for example an incremental cost-effectiveness ratio of cost per quality adjusted life-year gained.

Appendix 11

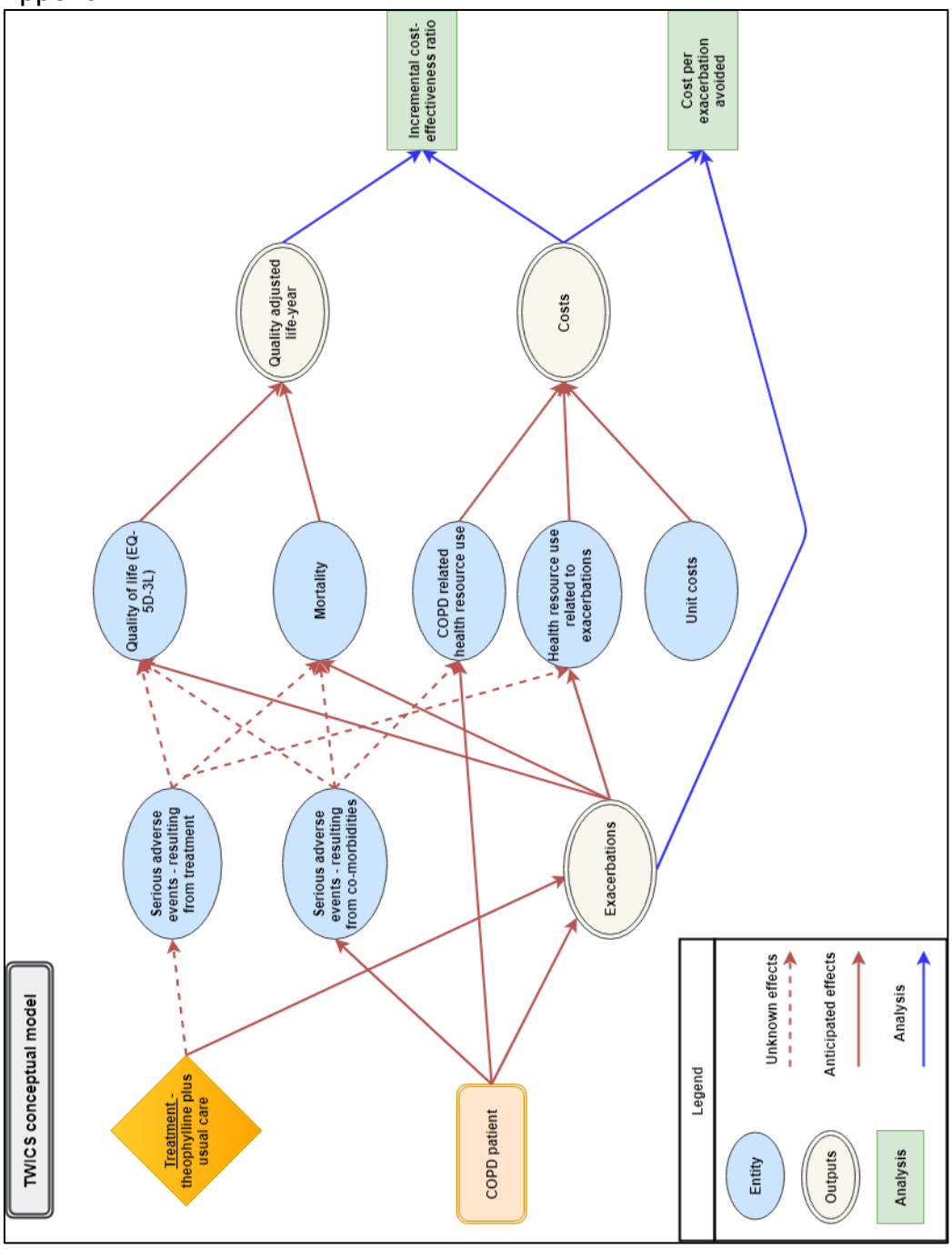


Figure A11.85: Version 1 ‘intervention’ TWICS conceptual model

Appendix 11

A ‘comparator’ version of version 1 was then developed as depicted in Figure A11.86. The columns in this diagram are more clearly defined and are categorised into: ‘disease area’, ‘changes in health’, ‘results of changes in health’, ‘summary of results’ and ‘estimation’. Serious adverse events have been removed as it was not considered a good representation of the economic evaluation. Instead in the ‘Changes in health’ column the entities are ‘non-exacerbation COPD complications’, ‘non-COPD related changes in health’ and the original output of ‘exacerbations’. In this diagram ‘comparator’ routes are added in black; it was considered after reviewing version 1 ‘intervention’ that there would be more links from the COPD patient to entities in the conceptual model, not necessarily mitigated through the ‘changes in health’ events, such as: quality of life and mortality. For example, a patient will use COPD related non-exacerbation health resources (maintenance therapy) whether they experience a health event or not and they will have a quality of life level/measure too. The red lines between the ‘changes in health’ column and ‘results of changes in health’ indicate that health events in the former column will result in the use of health services resources, change in quality of life measure and potentially length of life. As in the previous conceptual model these feed into the outcome measures which, in turn, feed into the analysis results measures.

Appendix 11

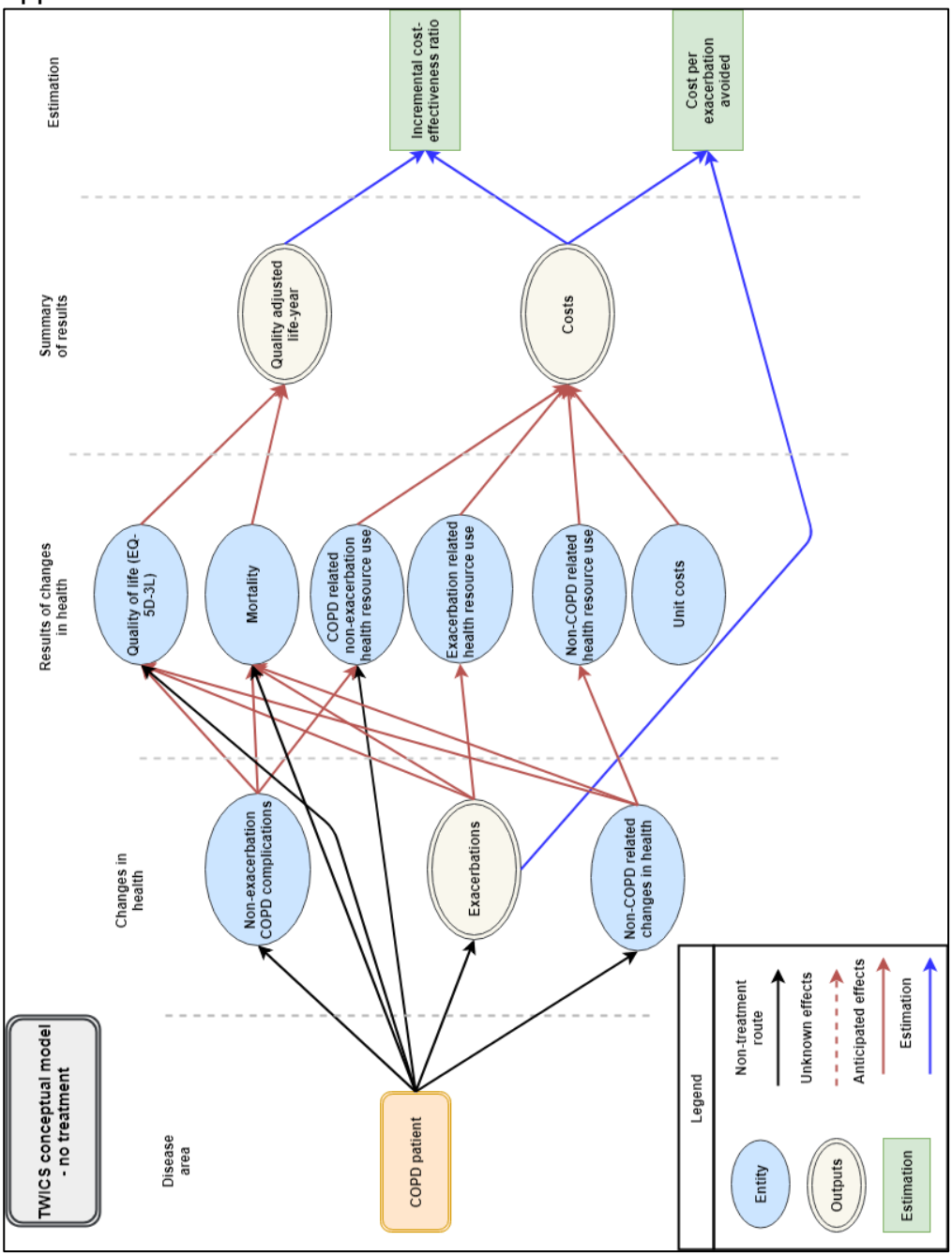


Figure A11.86: Version 1 'comparator' arm TWICS conceptual model

Appendix 11

Version 2 of the ‘intervention’ TWICS conceptual model builds on the ‘comparator’ version 1 adding in the intervention Figure A11.87. The anticipated consequences of the treatment effect the number of exacerbations, and unknown effects on ‘non-exacerbation COPD complications’ and ‘non-COPD related changes in health’.

Appendix 11

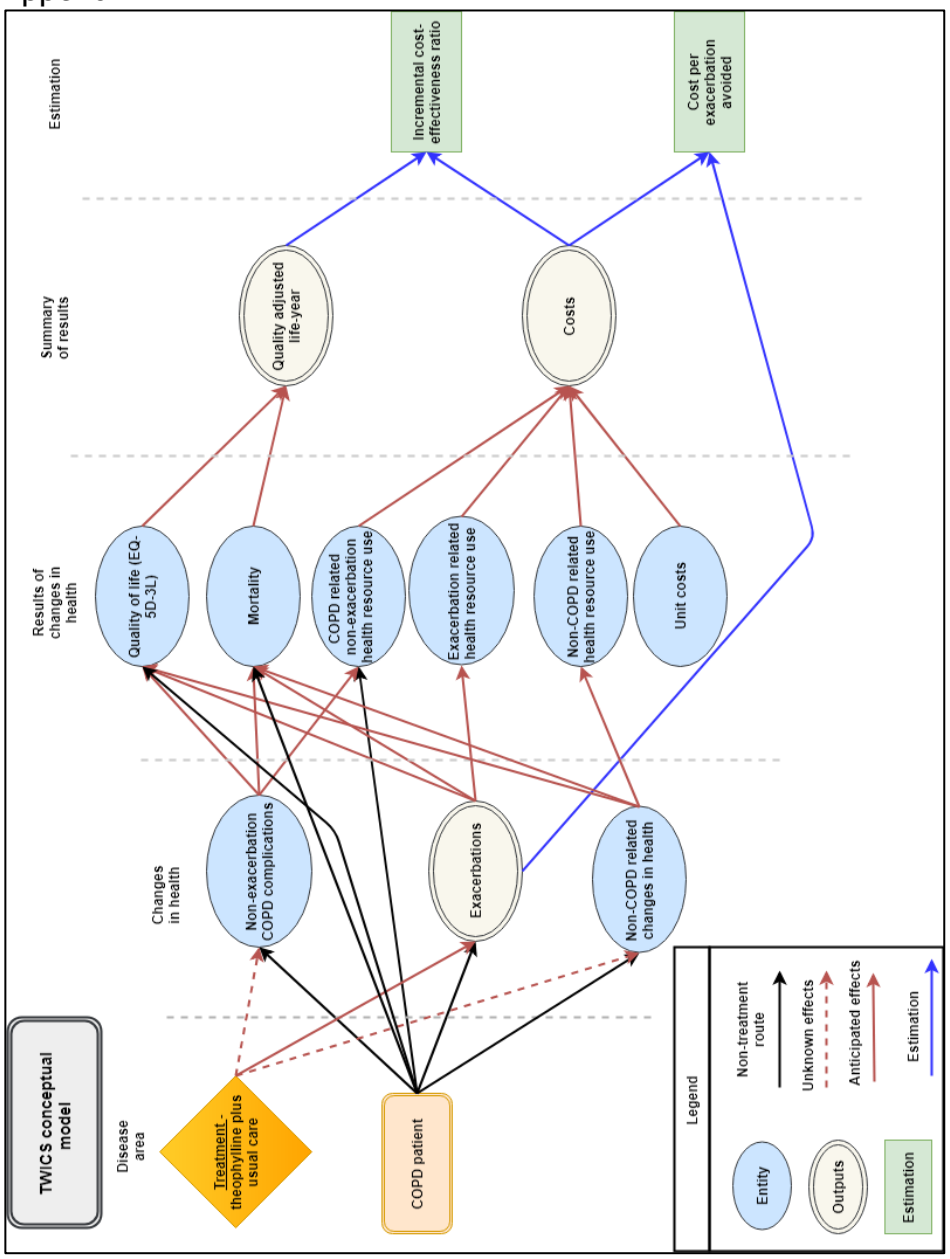


Figure A11.87: Version 2 'intervention' TWICS conceptual model

Appendix 11

The ‘comparator’ version of version 2 amalgamates the 3 health resource use entities into one to make the conceptual model simpler and easier to interpret (Figure A11.88).

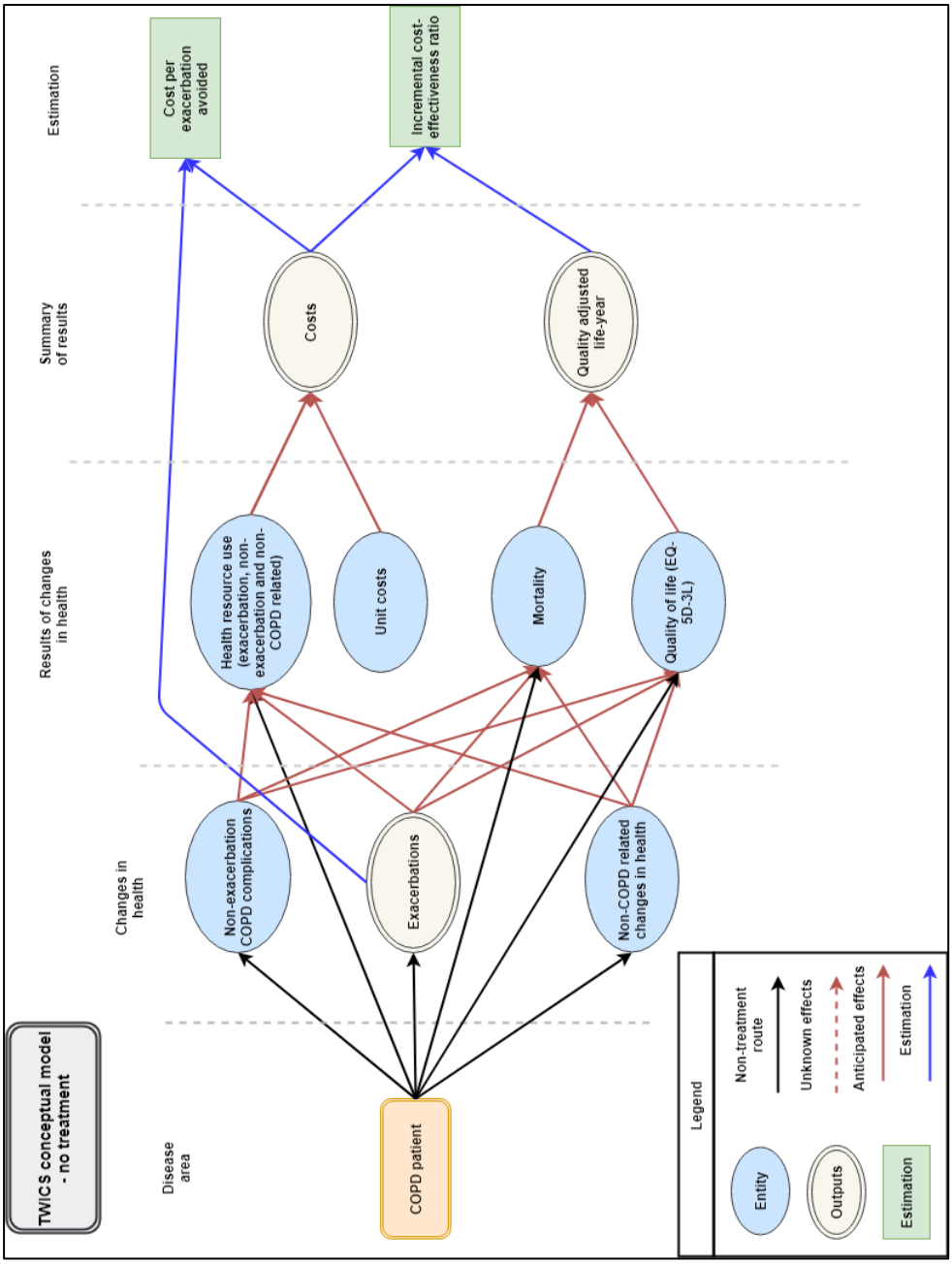


Figure A11.88: Version 2 'comparator' TWICS conceptual model

Appendix 11

Version 3 ‘intervention’ is similar to version 2 ‘comparator’ but with the intervention added Figure A11.89.

Appendix 11

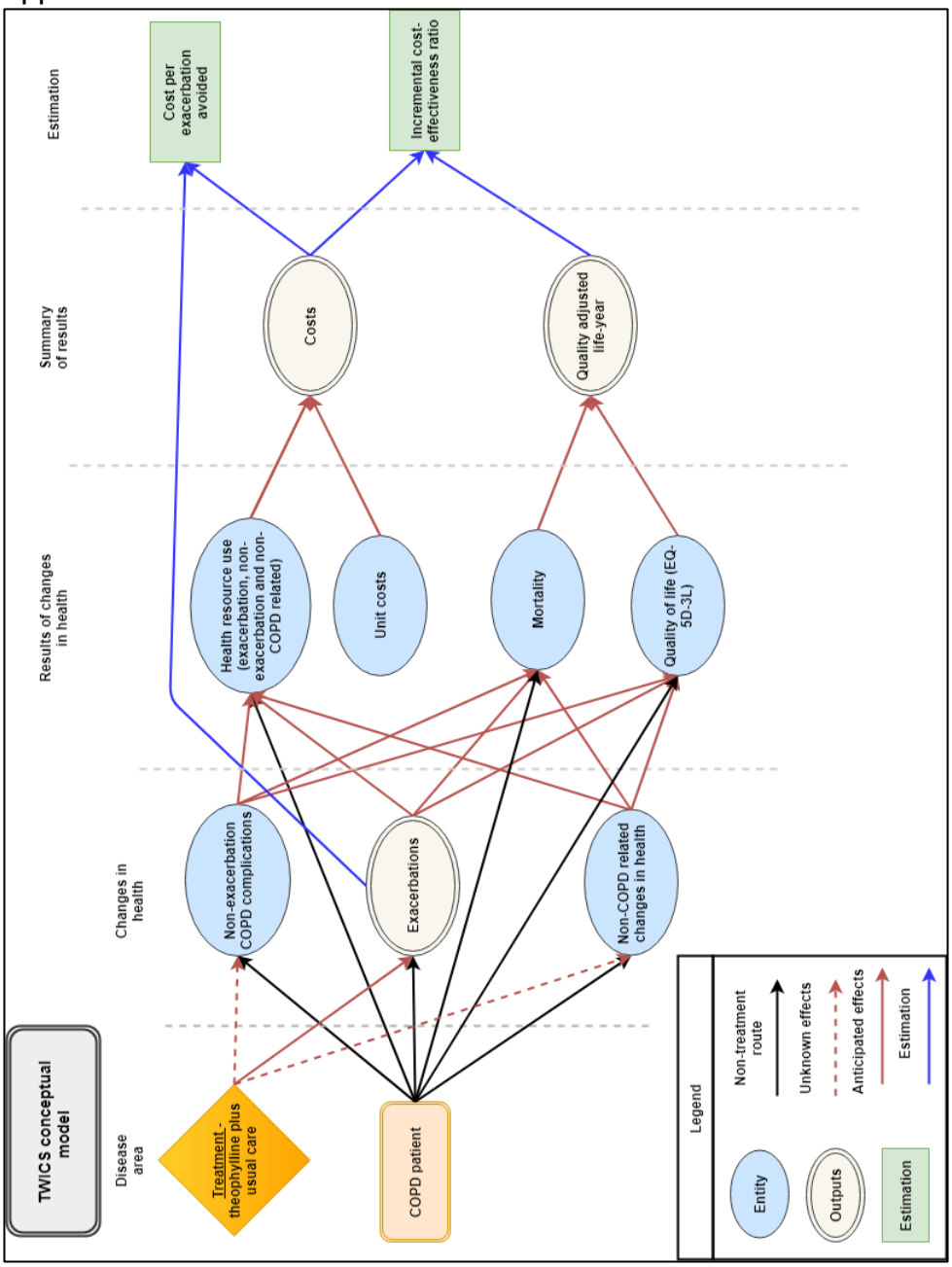


Figure A11.89: Version 3 'intervention' TWICS conceptual model

Appendix 11

Version 3 ‘comparator’ TWICS conceptual model (Figure A11.90) has the same layout and components, with a small change in descriptions: the ‘columns’ have changed from ‘changes in health’, results of changes in health’, ‘summary of results’ and ‘estimation’ to ‘events’, ‘measures’, and ‘summary of results’. These latter descriptions/labels were considered to be clearer and a better explanation of the components in the columns labelled. Some labels in the legend were also updated: ‘entity’ to ‘component’, ‘outputs’ to ‘output’, ‘estimation’ to ‘summary of results’ and the blue ‘estimation’ arrow to ‘summary of results’.

Appendix 11

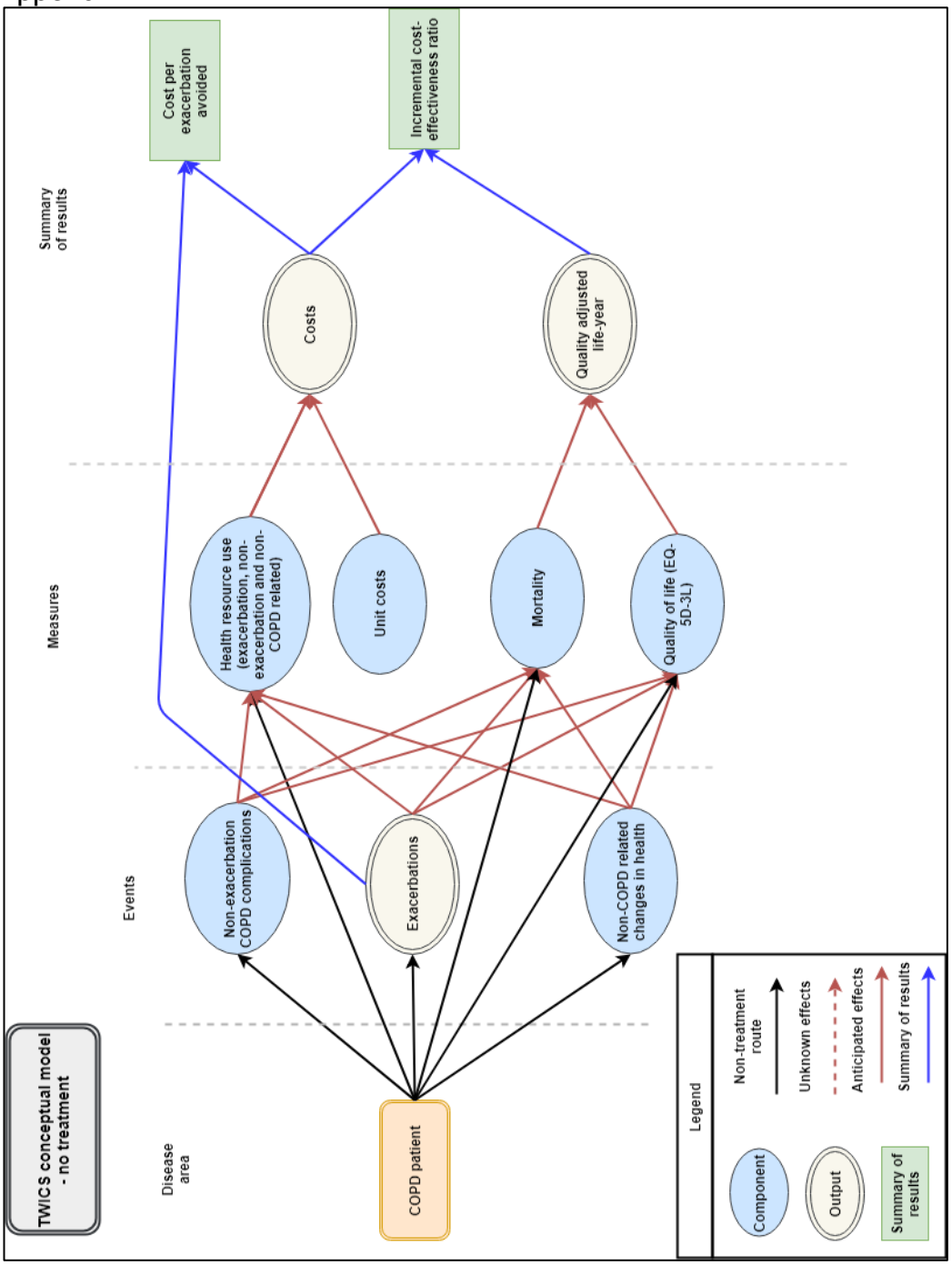


Figure A11.90: Version 3 'comparator' TWICS conceptual model

Appendix 11

Version 4 ‘intervention’ of the conceptual model utilises the changes to descriptions made in version 3 ‘comparator’, layering on the intervention (Figure A11.91).

Appendix 11

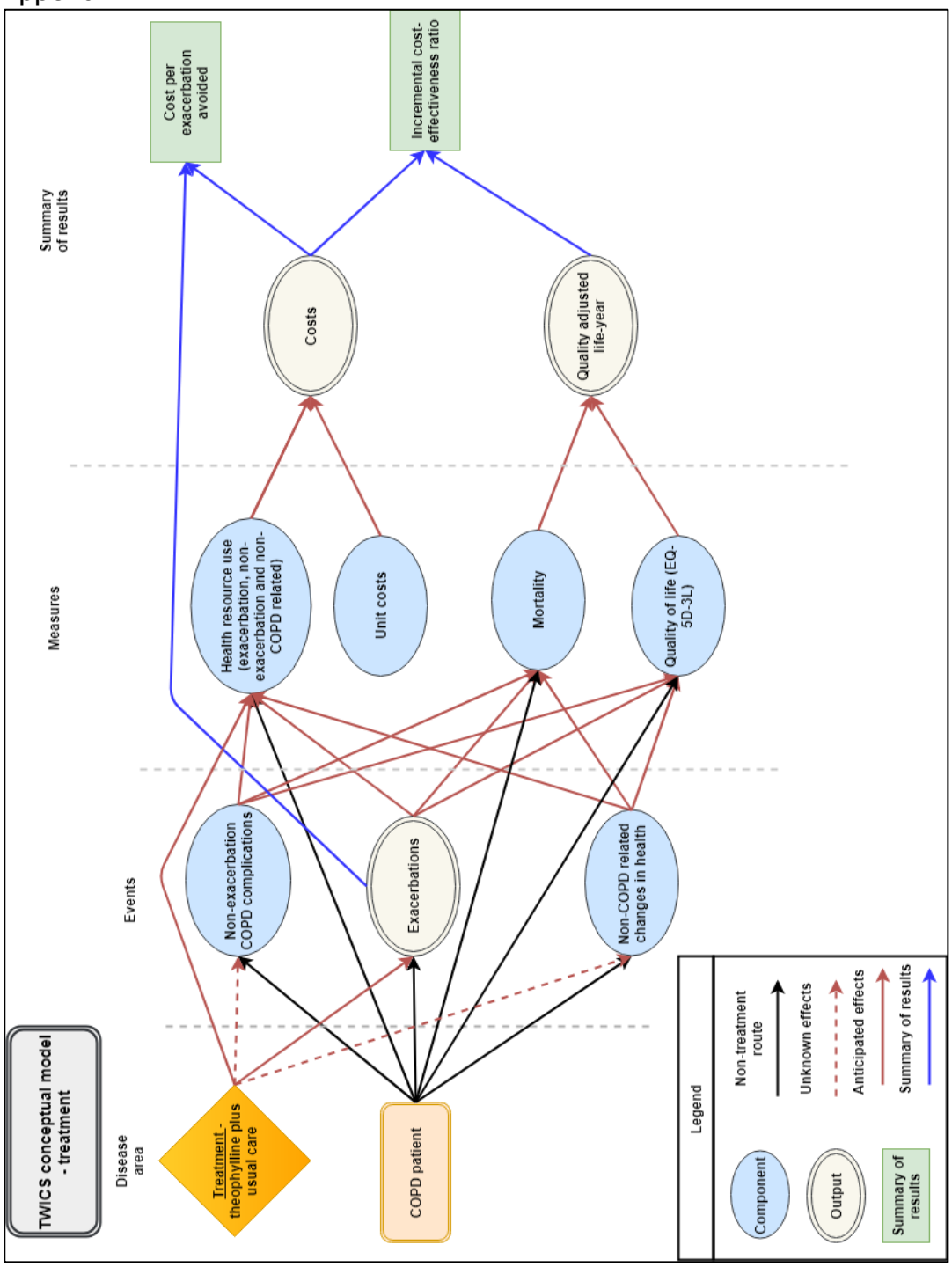


Figure A11.91: Version 4 'intervention' TWICS conceptual model

Appendix 11

At this stage in the development of the conceptual model it was presented to and discussed with the lead clinician in the TWICS trial. Changes made as a result of this consultation are explained next. To be more accurate about the description of the patient or participant the description should be changed to 'High risk COPD patient' from simply 'COPD patient'. (All participants are classified as GOLD stages C&D, as they have all had at least 2 exacerbations in year prior to recruitment). Instead of splitting the COPD related health events into exacerbations and non-exacerbations, the clinician advised splitting COPD related events into 'causal complications' (including exacerbations, pneumonia and collapsed lung) and 'associated complications' (such as CVD, depression, osteoporosis; these are lifestyle factors of COPD patients, side effects of medication or linked to COPD – depression - rather than a direct causal link - comorbidities linked to lifestyle). This adds an extra layer to divide the direct causal COPD events into exacerbations and non-exacerbations (pneumonia and collapsed lung), this makes more clinical sense than having COPD related complication and exacerbations separately in the first column.

The 'causal complications' description was split into exacerbations and 'COPD causal non-exacerbations' (pneumonia and collapsed lung). Then another component was included for events not linked to COPD events such as cancers (not lung cancers), falls, etc. We also took out the direct link between patient and mortality because mortality will always occur as a result of an event, whether it is COPD related or not.

To make the conceptual model clearer the 'events' column was renamed 'patient-related events'.

Side effects of treatment may increase health resource use (primary care visits for example) and treatment side-effects may also be the cause of pulling out of study, so may find the patients who pull out of study in the intervention arm have higher health resource use prior to dropping out – side effects may cause the patient to stop and re-start taking the intervention more than once. Increase use of ICS is linked to increase in pneumonia.

After making the changes suggested by the clinician the legend was moved across to the right-hand side to allow more room on the left-hand side. The

arrows were made narrower as they were so wide that they were dominating the model.

All these changes were made in version 5 ‘comparator’ (Figure A11.92).

Version 5 ‘intervention’ uses different colours for each column to make it easier to see that each column relates to different aspects of the economic evaluation (Figure A11.93). The ‘measures’ description the third column has been changed to ‘data’ to make it clearer what the component relate to in this column. It also incorporates a bracket to cut out the arrows from the ‘patient-related events’ column to the ‘data’ column.

Appendix 11

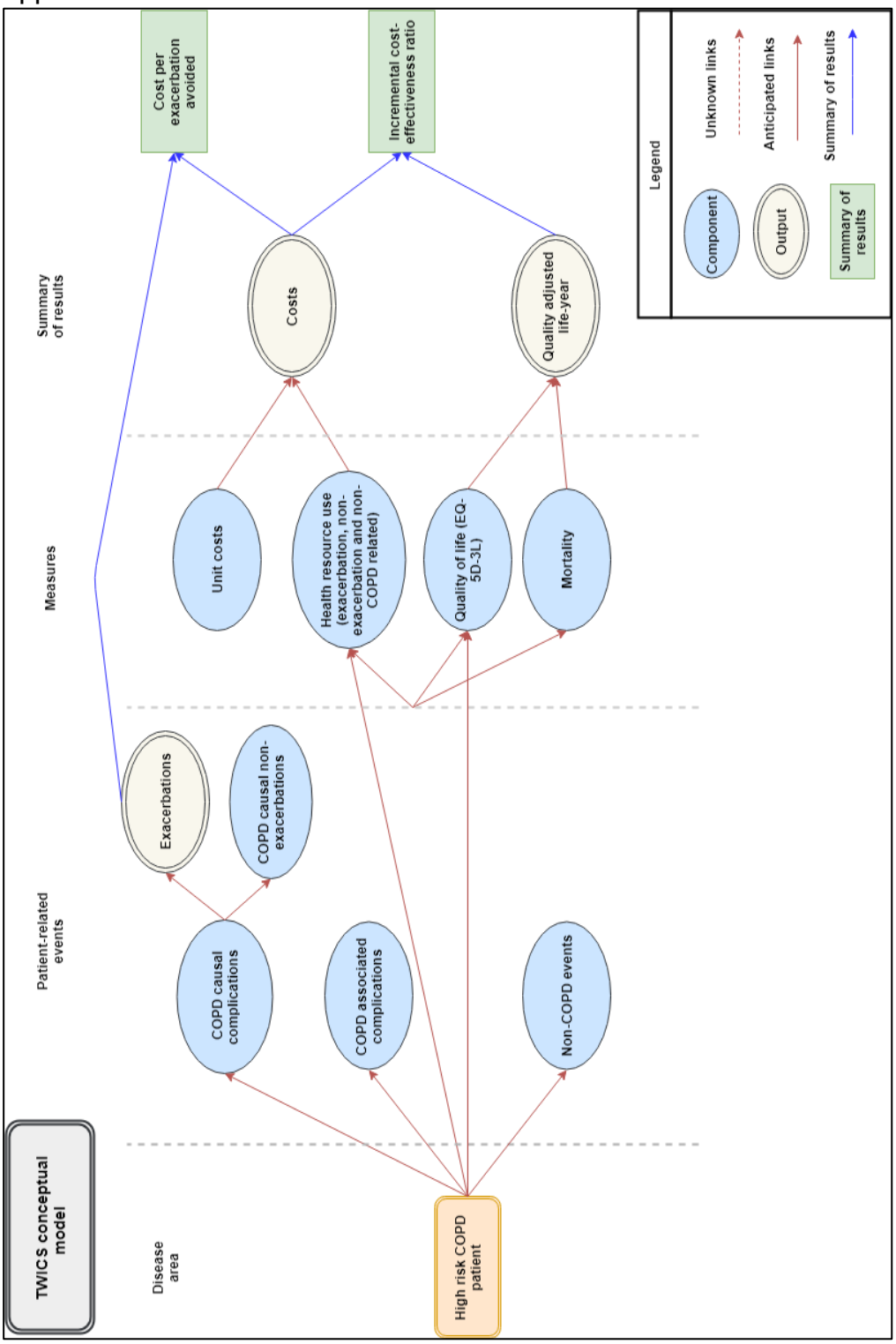


Figure A11.92: Version 5 'comparator' TWICS conceptual model

Appendix 11

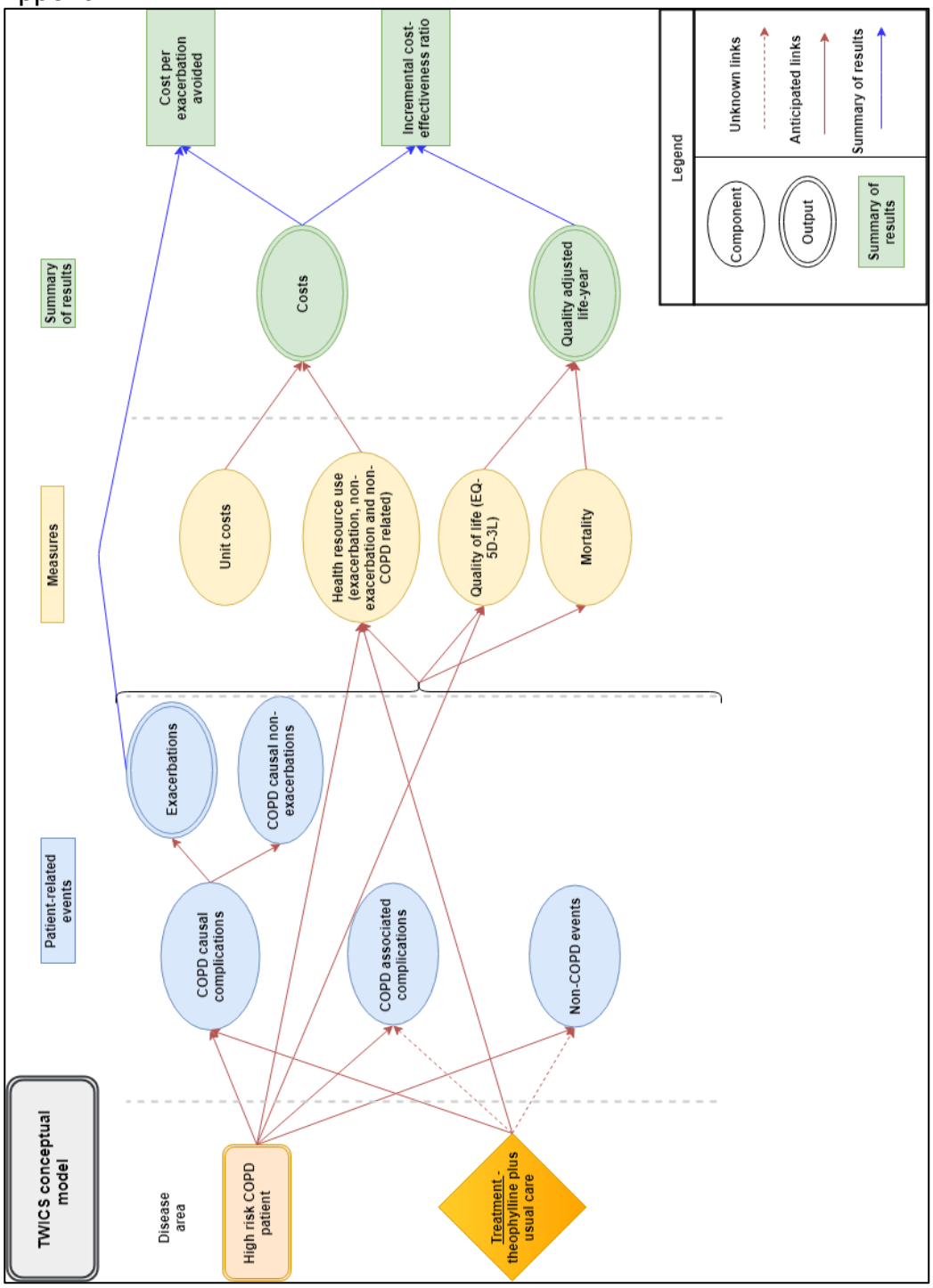


Figure A11.93: Version 5 'intervention' TWICS conceptual model

Appendix 11

V5 part 4 takes out direct links between patient and resource use and EQ-5D, whilst there is still a link between resource use and EQ-5D directly from the patient that does not have to be mediated through a patient-related event, the addition of the bracket to the right of this column represents that all previous components are able to link to the three relevant components in the ‘data’ column (Figure A11.94).

Appendix 11

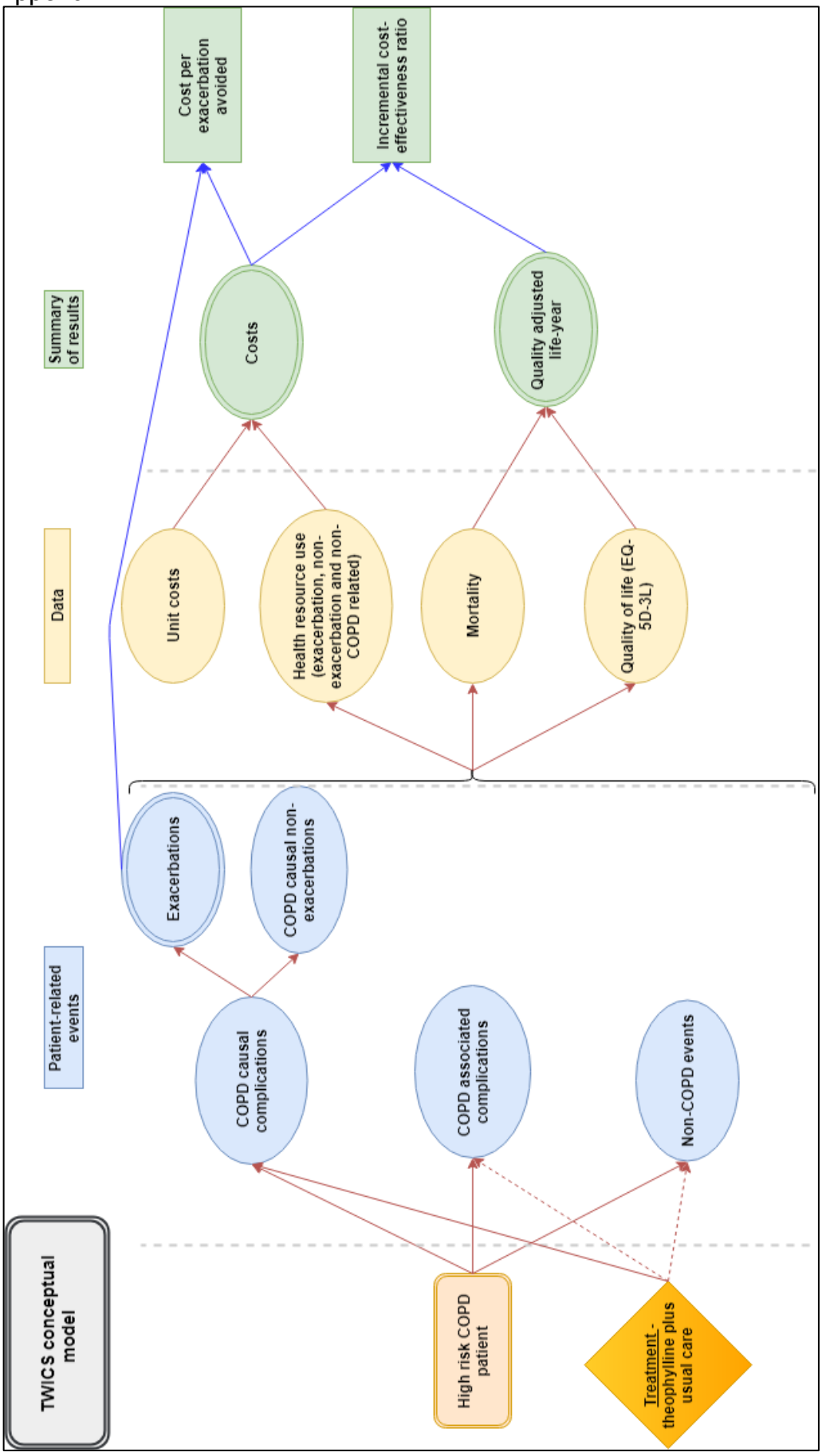


Figure A11.94: Version 5 part 4 ‘intervention' TWICS conceptual model

Appendix 11

V5 part 4 was presented to TWICS investigators as part of the 'Review, revise, refine' stage. Whilst presenting the conceptual model it was apparent that removing the lines from patient to data (quality of life and health resource use) and 'intervention' to 'health resource use' did not correctly represent the links between patients and data, it made the conceptual model too simplified, these lines were added back in. The description of 'Data' in the third column was changed back to 'Measures' as this is a better description of the components in this column. Therefore, the conceptual model reverted to V5. This was the last version in this sequence, after this the conceptual model was based on the disaggregated cost-effectiveness template.

Figure A11.95 is an early version of a generic conceptual model based on the disaggregated cost-effectiveness template; it has three panels relating to costs, the process or system, and health benefits.

Appendix 11

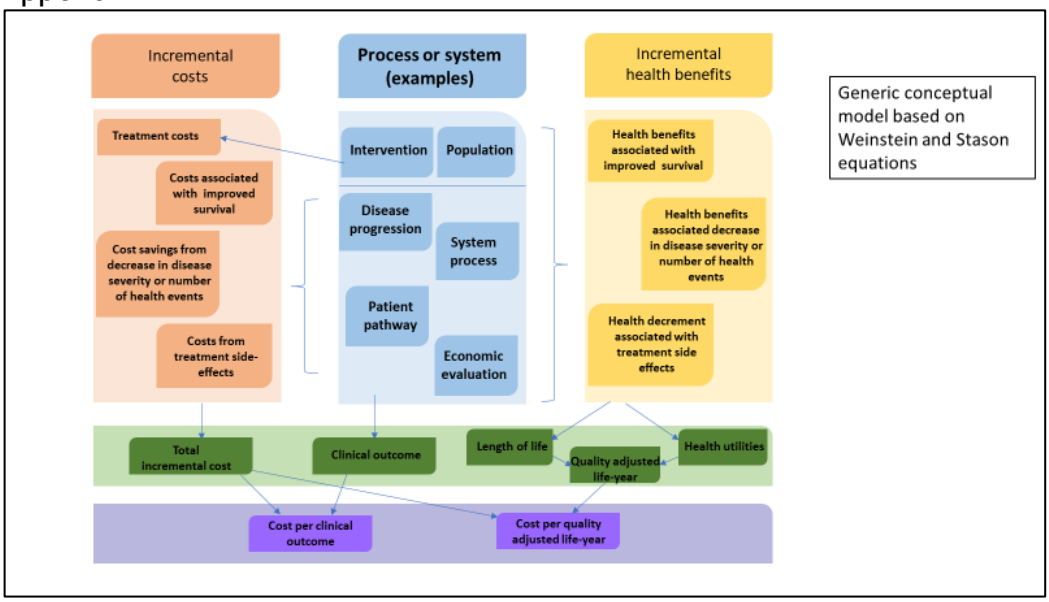


Figure A11.95: Early generic conceptual model – disaggregated cost-effectiveness template

Figure A11.96 is an early version of a COPD disease conceptual model based on the disaggregated cost-effectiveness template.

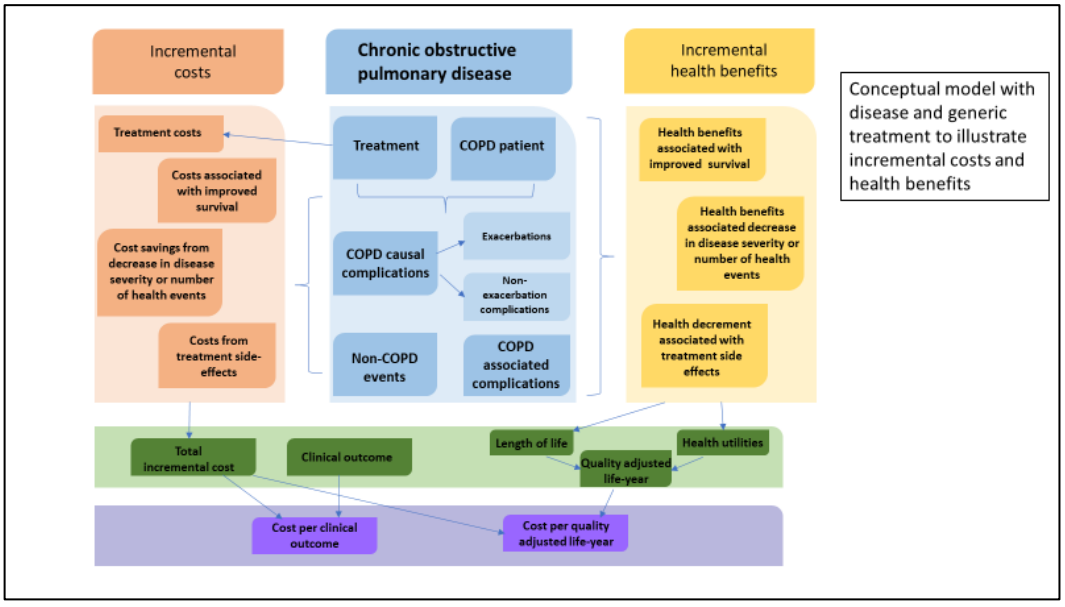


Figure A11.96: Early COPD and generic treatment conceptual model – disaggregated cost-effectiveness template

Appendix 12: Iterative development of BeatIt case study conceptual model diagram (*Chapter 6)*

This appendix includes early drafts of diagrams that feed into the final rich picture and conceptual model of the BeatIt case study.

The first diagram is an early rich picture of depression that depicts the characteristics of depression, showing that a patient with depression may achieve remission, however they may also experience a recurrence of depression Figure A12.97.

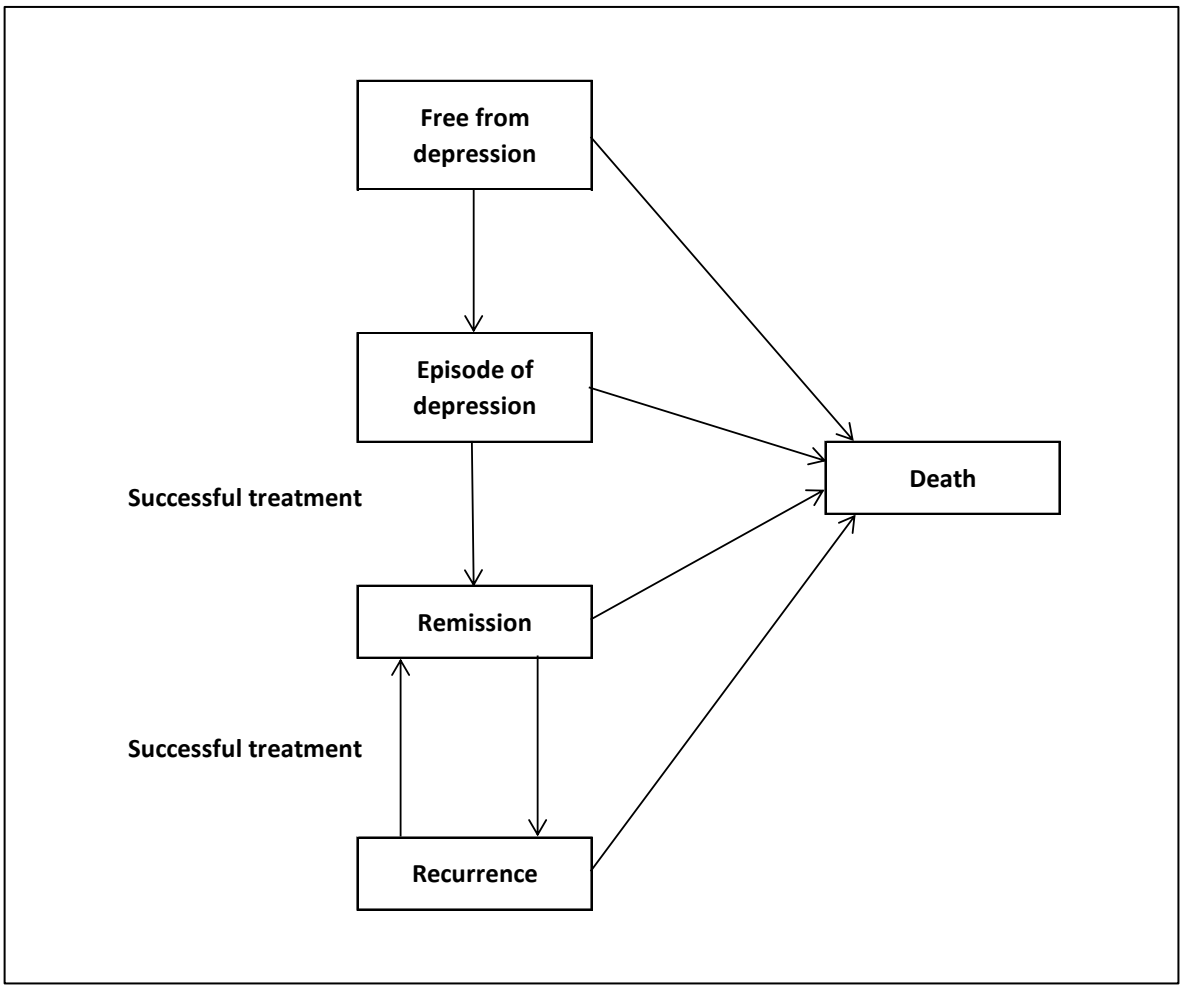


Figure A12.97: BeatIt draft rich picture

Appendix 12

The next figure is an early conceptual model showing the BeatIt therapy and its expected effect on activity (Figure A12.98), and then adding in the economic data and finally the analysis summary measure of incremental cost-effectiveness ratio (Figure A12.99).

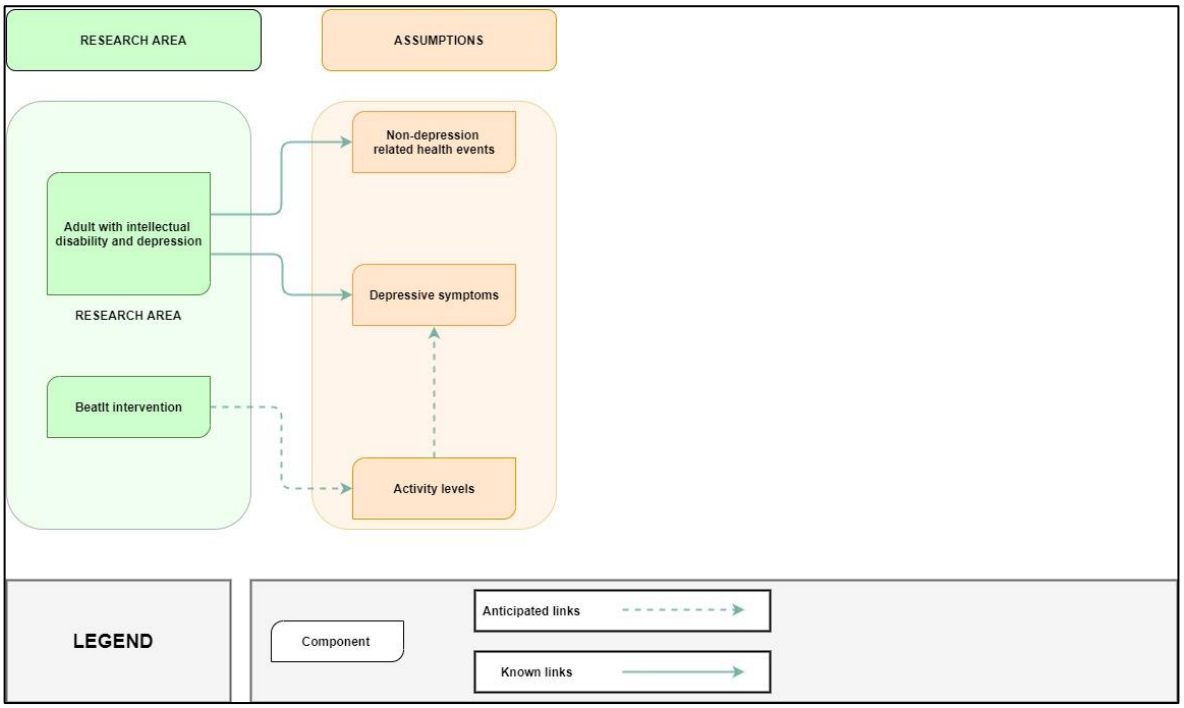


Figure A12.98: BeatIt v1

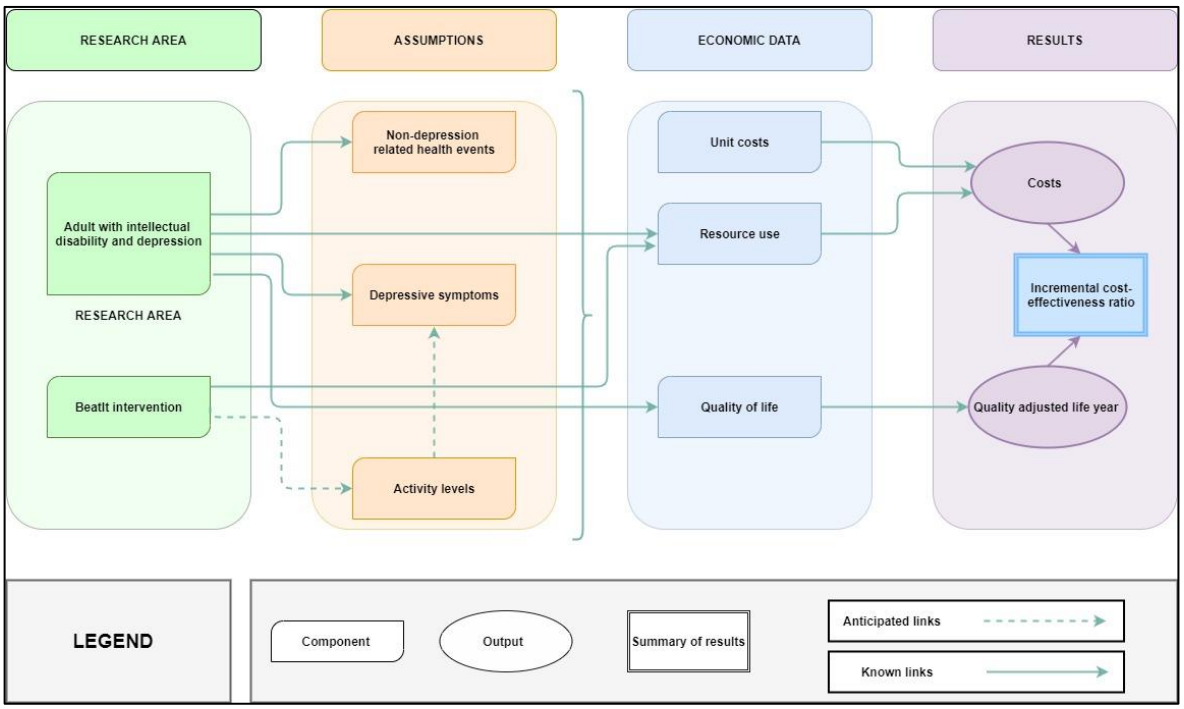


Figure A12.99: BeatIt v2

Appendix 12

The last figure shows an early draft of the conceptual model with StepUp therapy, it shows that StepUp was not expected to affect levels of activity (Figure A12.100).

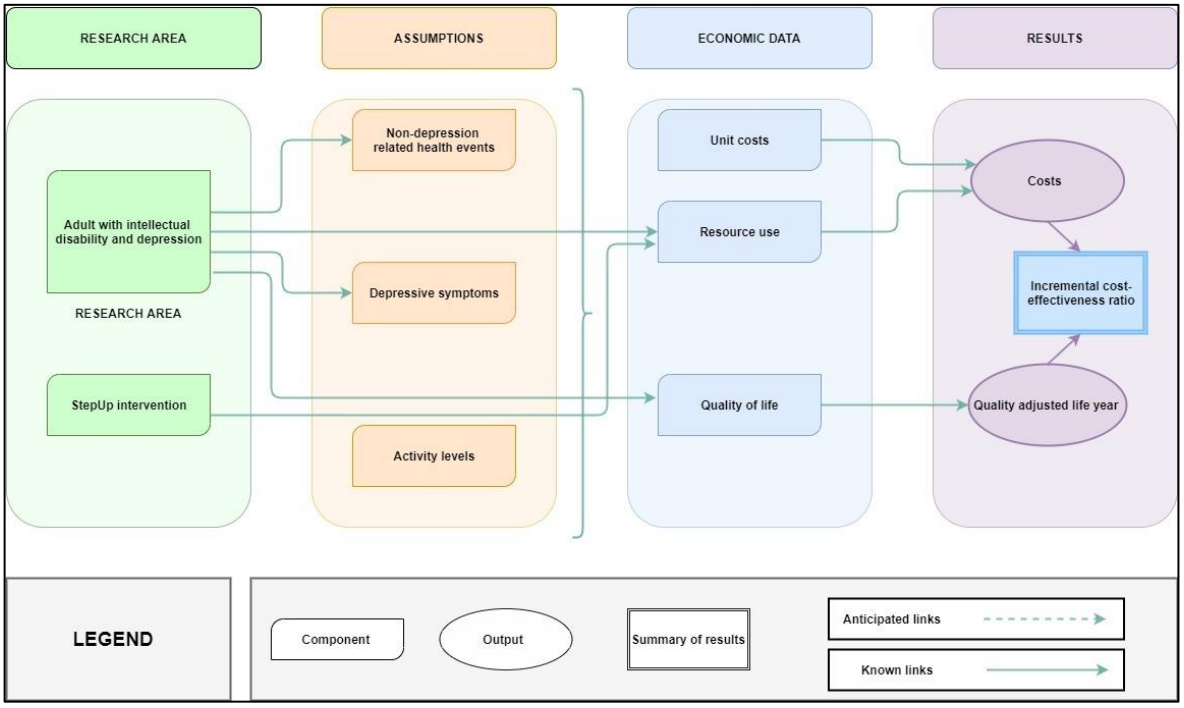


Figure A12.100: StepUp v1

Appendix 13: Stata syntax for TWICS and BeatIt conceptual model analysis case studies

TWICS syntax

```
*************************************************************************
*************************************************************************
***** TWICS ANALYSIS FOR CONCEPTUAL MODEL APPROACH *****
*************************************************************************
*************************************************************************

************************************************************
*** This Do file includes the TWICS conceptual model analysis
*** presented in Chapter 7 of PhD thesis
************************************************************

*** First upload dataset
use "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 6 TWICS\Further TWICS analysis novel approach\Analysis included in thesis\twics dataset for novel approach 030821.dta", clear

** This dataset includes the ITT population (n=1536)
** Missing data is replaced with treatment arm specific mean to enable use of a
** full dataset in the case study without complications.
** NB THE CASE STUDY IS FOR ILLUSTRATIVE PURPOSES ONLY

************************************************
*** Variables in dataset (n=15)
************************************************
** totalex - total number of exacerbations
** exac_tc - total cost of exacerbations
** exloc_tc - total cost of exacerbation treatment realting to location (home, service and hospital)
** extrt_tc - total cost of treating exacerbation (non location)
** exhome_tc - total cost of treating excerbations at home
** exserv_tc - total cost of treating excerbations with care by services to prevent hospitalisation
** exhosp_tc - total cost of treating excerbations in hospital
** treatmentno - treatment allocation (0 placebo, 1 theophylline)
** qaly_tot - total QALYs over trial period
** noofexacerbationstreatment - reported number of exac at baseline needing treatment for previous 12 months
** noofexacerbationshospital - reported number of exac at baseline needing hospital treatment in previous 12 months
** twics_tc - total TWICS costs
** nonintnonexac_tc - total costs excluding treatment and exac costs
** non_int_tc - total costs less treatment costs
** int_tc - total treatment costs

************************************************************
************************************************************
*** 1. REGRESSIONS TO TEST THE ACCURACY OF THE CONCEPTUAL MODEL
************************************************************
************************************************************

*******************************************************
** Regression 1) Does treatment arm predict treatment cost?
*******************************************************
reg int_tc treatmentno, nocon

************************************************************
```

Appendix 13

```
** Regression 2) Does treatment arm predict no. of exacerbations?
*****************************************************************

reg totalex treatmentno

*************************************************************************************************
** Regression 3) Do number of exacerbations predict QALYs, and possible association with trt?
*************************************************************************************************

reg qaly_tot totalex treatmentno

*************************************************************************************************
** Regression 4) Does number of exacerbations predict non-treatment costs, and possible
association with trt?
*************************************************************************************************

reg nonint_tc totalex treatmentno

************************************************
************************************************
*** Additional analysis on equation 4
************************************************
************************************************


** Treatment does predict non-treatment costs
** We think this is due to exacerbation costs for hospitalised exacerbations
** Split the exacerbation costs into location - hospital v non-hospital to explore the affect of
treatment on exacerbation costs

** generate a variable for non-hospital treatment costs
gen ex_nonhosp_tc = exhome_tc + exserv_tc

** Does treatment predict exacerbation costs - hospital and non-hospital?
** Plus does treatment predict the costs of treating exacerbations?
reg exac_tc treatmentno
reg exhosp_tc treatmentno
reg ex_nonhosp_tc treatmentno
reg extrt_tc treatmentno

** Only costs of exacerbations treated in hospital are predicted by treatment
** Try this with number of exacerbations included
reg exac_tc totalex treatmentno
reg exhosp_tc totalex treatmentno
reg ex_nonhosp_tc totalex treatmentno
reg extrt_tc totalex treatmentno

** Look at non-treatment, non-exacerbation costs
reg nonintnonexac_tc totalex treatmentno

*******************************************************
*******************************************************
****** 2. FINAL CONCEPTUAL MODEL ANALYSIS
*******************************************************
*******************************************************

*****************************************************************************
*****************************************************************************
** Cost-effectiveness plane 1 - trial-arm based - THIS IS FOR CHECKING PURPOSES ONLY
** Syntax to compute difference in costs and effects based on treatment arm only -
** should be similar to the CE plane included in chapter 3 (slight difference
** as CE plane in chapter 3 uses complete case data and this one uses dataset with
** missing data replaced with treatment arm specific mean)
*****************************************************************************
*****************************************************************************


** generate a temporary variable for treatment arms (predictor variables)
```

Appendix 13

```
** define program and variables
gen temp=.
capture program drop booticer
program define booticer, rclass
tempvar cost0 cost1 qaly0 qaly1

** replicate original treatment arm based analysis by only using treatment arm in regression
replace temp=treatmentno
regress twics_tc temp
replace temp=0
predict `cost0'
replace temp=1
predict `cost1'
sum `cost0', meanonly
local mcost0=r(mean)
sum `cost1', meanonly
local mcost1=r(mean)

replace temp = treatmentno
regress qaly_tot temp
replace temp=0
predict `qaly0'
replace temp=1
predict `qaly1'
sum `qaly0',meanonly
local mqaly0=r(mean)
sum `qaly1',meanonly
local mqaly1=r(mean)

return scalar cost1 = `mcost1'
return scalar cost0 = `mcost0'
return scalar costDiff=`mcost1' - `mcost0'
return scalar qaly1 = `mqaly1'
return scalar qaly0 = `mqaly0'
return scalar qalyDiff = `mqaly1' - `mqaly0'
end
** END OF PROGRAM 'booticer'

*Bootstrap the difference in costs and effects & save the output file
bootstrap       cost1=r(cost1) cost0=r(cost0) costDiff=r(costDiff) ///
                        qaly1=r(qaly1) qaly0=r(qaly0) qalyDiff=r(qalyDiff) ///
                        icer=(r(costDiff)/r(qalyDiff)), ///
                        reps(1000) saving(bstwicsphd_ceplane1_thesis.dta, replace)
seed(12345):booticer

*Summarize the differences in costs and effects
use bstwicsphd_ceplane1_thesis.dta, replace
label var costDiff "Incremental cost"
label var qalyDiff "Incremental QALY"
log using icer_qaly, replace
sum costDiff qalyDiff icer
matrix list e(b)
quietly mean costDiff
matrix b = e(b)
local Cost = b[1,1]
quietly mean qalyDiff
matrix b=e(b)
local qaly = b[1,1]
display "ICER: " `Cost'/`qaly'

*plot the cost/effects on the cost-effectiveness plane
twoway (scatter costDiff qalyDiff, msize(tiny)), yline(0) xline(0) ///
        xlabel(-0.08(0.02)0.04) ylabel(-1500(500)500)
```

```
graph save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 6
TWICS\Further TWICS analysis novel approach\Analysis included in
thesis\twics_ceplane1_thesis.gph", replace

*** exporting .dta file to excel
export excel using "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old
layout\Chapter 6 TWICS\Further TWICS analysis novel approach\Analysis included in
thesis\bootstrap_twics_ceplane1_thesis.xls", replace

**Save dataset
save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 6
TWICS\Further TWICS analysis novel approach\Analysis included in
thesis\twicsceplane1_thesis.dta", replace


*************************************************************************
*************************************************************************
** CONCEPTUAL MODEL 2 (Figure 55)
** Conceptual model based analysis with all hypothesised equations
** Syntax to compute difference in costs and effects based on full conceptual
** model (including all regressions identified)
*************************************************************************
*************************************************************************

** Import same data as above
use "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 6
TWICS\Further TWICS analysis novel approach\Analysis included in thesis\twics dataset for novel
approach 030821.dta", clear

** Generate two temporary variables for predictor variables
gen temp=.
gen exac_temp=.
** Delete the program if it already exists
capture program drop cm_ceplane
** Create program
program define cm_ceplane, rclass
** Define variables for program
tempvar cost0t cost1t cost0nt cost1nt exac0 exac1 qaly0 qaly1

*************************************************************************
*** Equation #1 these are treatment costs only (predictor - treatment arm)
*************************************************************************
** Take a copy of treatment indicator
replace temp=treatmentno
** Conceptual model regression
regress int_tc temp
** Run as if everyone is in placebo arm and store prediction in new variable
replace temp=0
predict `cost0t'
** Run as if everyone is in theophylline arm and store prediction in new variable
replace temp=1
predict `cost1t'
** Calculate the mean for the bootstrap run for both arms
sum `cost0t', meanonly
local mcost0t=r(mean)
sum `cost1t', meanonly
local mcost1t=r(mean)

*******************************************************
*** Equation #2 exacerbations (predictor - treatment arm)
*******************************************************
** Take a copy of treatment indicator
replace temp = treatmentno
```

Appendix 13

```
** Conceptual model regression
regress totalex temp
** Run as if everyone is in placebo arm and store prediction in new variable
replace temp=0
predict `exac0'
** Run as if everyone is in theophylline arm and store prediction in new variable
replace temp=1
predict `exac1'

*******************************************************************************
*** Equation #3 QALYs (predictors - treatment arm and number of exacerbations)
*******************************************************************************
** Take a copy of treatment indicator
replace temp = treatmentno
** Take a copy of exacerbations variable
replace exac_temp = totalex
** Conceptual model regression
regress qaly_tot exac_temp temp
** Run as if everyone is in placebo arm and store prediction in new variable
replace temp=0
** Run with predicted exacerbations from equation #2
replace exac_temp = `exac0'
predict `qaly0'
** Run as if everyone is in theophylline arm and store prediction in new variable
replace temp=1
** Run with predicted exacerbations from equation #2
replace exac_temp = `exac1'
predict `qaly1'
** Calculate the mean for the bootstrap run for both arms
sum `qaly0',meanonly
local mqaly0=r(mean)
sum `qaly1',meanonly
local mqaly1=r(mean)

*******************************************************************************
*** Equation #4 these are non-treatment costs only (predictors treatment arm
** and number exacerbations)
*******************************************************************************
** Take a copy of treatment indicator
replace temp=treatmentno
** Take a copy of exacerbations variable
replace exac_temp = totalex
** Conceptual model regression
regress nonint_tc exac_temp temp
** Run as if everyone is in placebo arm and store prediction in new variable
replace temp=0
** Run with predicted exacerbations from equation #2
replace exac_temp = `exac0'
predict `cost0nt'
** Run as if everyone is in theophylline arm and store prediction in new variable
replace temp=1
** Run with predicted exacerbations from equation #2
replace exac_temp = `exac1'
predict `cost1nt'
** Calculate the mean for the bootstrap run for both arms
sum `cost0nt', meanonly
local mcost0nt=r(mean)
sum `cost1nt', meanonly
local mcost1nt=r(mean)

** Instruction to return the predictions and to combine treatment and non-treatment costs into
one total cost. Also to calculate cost and QALY differences
return scalar cost1 = `mcost1t' + `mcost1nt'
```

Appendix 13

```
return scalar cost0 = `mcost0t' + `mcost0nt'
return scalar cost1t = `mcost1t'
return scalar cost0t = `mcost0t'
return scalar cost1nt = `mcost1nt'
return scalar cost0nt = `mcost0nt'

return scalar costDiff= (`mcost1t' - `mcost0t')+(`mcost1nt' - `mcost0nt')
return scalar qaly1 = `mqaly1'
return scalar qaly0 = `mqaly0'
return scalar qalyDiff = `mqaly1' - `mqaly0'
end

*** END OF PROGRAM 'cm_ceplane'

**** Bootstrap the difference in costs and effects  & save the output file
bootstrap        cost1=r(cost1) cost0=r(cost0) costDiff=r(costDiff) ///
                        cost1t=r(cost1t) cost0t=r(cost0t) ///
                        cost1nt=r(cost1nt) cost0nt=r(cost0nt) ///
                        qaly1=r(qaly1) qaly0=r(qaly0) qalyDiff=r(qalyDiff) ///
                        icer=(r(costDiff)/r(qalyDiff)), ///
                        reps(1000) saving(bstwicsphd_ceplane2_thesis, replace)
seed(12345):cm_ceplane

** Summarise the differences in costs and QALYs
label var costDiff "Incremental cost"
label var qalyDiff "Incremental QALY"
sum costDiff qalyDiff icer
quietly mean costDiff
matrix b = e(b)
local Cost = b[1,1]
quietly mean qalyDiff
matrix b=e(b)
local qaly = b[1,1]
display "ICER: " `Cost'/`qaly'

*** Use difference in cost and effects to produce cost-effectiveness plane
ellip costDiff qalyDiff , c(f) level(95) plot(scatter costDiff qalyDiff, xlabel(-0.08(0.02)0.04)
ylabel(-1500(500)500) xline(0) yline(0) msize(tiny) mcolor(teal) xtitle("Difference in QALYs")
ytitle("Difference in costs (£)") title("Cost-effectiveness plane") subtitle("Conceptual model
driven analysis (all regressions)") legend(label(1 "95% confidence ellipse") label(2 "Bootstrap
samples")))
graph save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 6
TWICS\Further TWICS analysis novel approach\Analysis included in
thesis\twics_ceplane2v2_thesis.gph", replace

*** Export .dta file to excel
export excel using "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old
layout\Chapter 6 TWICS\Further TWICS analysis novel approach\Analysis included in
thesis\bootstrap_twics_ceplane2_thesis.xls", replace

**Save dataset
save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 6
TWICS\Further TWICS analysis novel approach\Analysis included in
thesis\twicsceplane2_thesis.dta", replace


*************************************************************************
*************************************************************************
** CONCEPTUAL MODEL 3 (Figure 56 & 57)
** Using strongly associated predictors identified in regressions
** Syntax to compute difference in costs and effects based on conceptual
** model identified using regressions, including strongly predictors
*************************************************************************
```

Appendix 13

```
*************************************************************************

* Import same data as above
use "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 6
TWICS\Further TWICS analysis novel approach\Analysis included in thesis\twics dataset for novel
approach 030821.dta", clear

** Generate two temporary variables for predictor variables
gen temp=.
gen exac_temp=.
** Delete the program if it already exists
capture program drop cm_ceplane
** Create program
program define cm_ceplane, rclass
** Define variables for program
tempvar cost0t cost1t cost0nt cost1nt exac0 exac1 qaly0 qaly1

*****************************************************************************
*** Equation #1 these are treatment costs only (predicted by treatment arm only)
*****************************************************************************
** Take a copy of treatment indicator
replace temp=treatmentno
** Conceptual model regression
regress int_tc temp
** Run as if everyone is in placebo arm and store prediction in new variable
replace temp=0
predict `cost0t'
** Run as if everyone is in theophylline arm and store prediction in new variable
replace temp=1
predict `cost1t'
** Calculate the mean treatment cost for the bootstrap run for both arms
sum `cost0t', meanonly
local mcost0t=r(mean)
sum `cost1t', meanonly
local mcost1t=r(mean)

*************************************************************
*** Equation #2 exacerbations (predicted by treatment arm only)
*************************************************************
** Take a copy of treatment indicator
replace temp = treatmentno
** Conceptual model regression
regress totalex temp
** Run as if everyone is in placebo arm and store prediction in new variable
replace temp=0
predict `exac0'
** Run as if everyone is in theophylline arm and store prediction in new variable
replace temp=1
predict `exac1'

*****************************************************************
*** Equation #3 QALYs (predicted by number of exacerbations only)
*****************************************************************
** Take a copy of exacerbations variable
replace exac_temp = totalex
** Conceptual model regression
regress qaly_tot exac_temp
** Run as if everyone is in placebo arm and store prediction in new variable
replace temp=0
** Run with predicted exacerbations from equation #2
replace exac_temp = `exac0'
predict `qaly0'
** Run as if everyone is in theophylline arm and store prediction in new variable
```

Appendix 13

```
replace temp=1
** Run with predicted exacerbations from equation #2
replace exac_temp = `exac1'
predict `qaly1'
** Calculate the mean QALYs for the bootstrap run for both arms
sum `qaly0',meanonly
local mqaly0=r(mean)
sum `qaly1',meanonly
local mqaly1=r(mean)

*****************************************************************************************
*** Equation #4 Non-treatment costs (predicted by number of exacerbations only)
*****************************************************************************************
** Take a copy of exacerbations variable
replace exac_temp = totalex
** Conceptual model regression
regress nonint_tc exac_temp
** Run as if everyone is in placebo arm and store prediction in new variable
replace temp=0
** Run with predicted exacerbations from equation #2
replace exac_temp = `exac0'
predict `cost0nt'
** Run as if everyone is in theophylline arm and store prediction in new variable
replace temp=1
** Run with predicted exacerbations from equation #2
replace exac_temp = `exac1'
predict `cost1nt'
** Calculate the mean non-treatment costs for the bootstrap run for both arms
sum `cost0nt', meanonly
local mcost0nt=r(mean)
sum `cost1nt', meanonly
local mcost1nt=r(mean)

** Instruction to return the predictions and to combine treatment and non-treatment costs into
one total cost. Also to calculate cost and QALY differences
return scalar cost1 = `mcost1t' + `mcost1nt'
return scalar cost0 = `mcost0t' + `mcost0nt'
return scalar cost1t = `mcost1t'
return scalar cost0t = `mcost0t'
return scalar cost1nt = `mcost1nt'
return scalar cost0nt = `mcost0nt'

return scalar costDiff= (`mcost1t' - `mcost0t')+(`mcost1nt' - `mcost0nt')
return scalar qaly1 = `mqaly1'
return scalar qaly0 = `mqaly0'
return scalar qalyDiff = `mqaly1' - `mqaly0'
end

*Bootstrap the difference in costs and effects & save the output file
bootstrap       cost1=r(cost1) cost0=r(cost0) costDiff=r(costDiff) ///
                        cost1t=r(cost1t) cost0t=r(cost0t) ///
                        cost1nt=r(cost1nt) cost0nt=r(cost0nt) ///
                        qaly1=r(qaly1) qaly0=r(qaly0) qalyDiff=r(qalyDiff) ///
                        icer=(r(costDiff)/r(qalyDiff)), ///
                        reps(1000) saving(bstwicsphd_ceplane3_thesis, replace)
seed(12345):cm_ceplane

** Summarize the differences in costs and QALYs
use bstwicsphd_ceplane3_thesis.dta, clear
label var costDiff "Incremental cost"
label var qalyDiff "Incremental QALY"
sum costDiff qalyDiff icer
quietly mean costDiff
```

Appendix 13

```
matrix b = e(b)
local Cost = b[1,1]
quietly mean qalyDiff
matrix b=e(b)
local qaly = b[1,1]
display "ICER: " `Cost'/`qaly'

*** Use difference in costs and effects to produce cost-effectiveness plane
ellip costDiff qalyDiff , c(f) level(95) plot(scatter costDiff qalyDiff, xlabel(-0.08(0.02)0.04)
ylabel(-1500(500)500) xline(0) yline(0) msize(tiny) mcolor(teal) xtitle("Difference in QALYs")
ytitle("Difference in costs (£)") title("Cost-effectiveness plane") subtitle("Conceptual model
driven analysis (strongly associated predictors)") legend(label(1 "95% confidence ellipse") label(2
"Bootstrap samples")))
graph save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 6
TWICS\Further TWICS analysis novel approach\Analysis included in
thesis\twics_ceplane3v2_thesis.gph", replace

*** Plot the cost/qalys on the cost-effectiveness plane with alternative large axis
ellip costDiff qalyDiff , c(f) level(95) plot(scatter costDiff qalyDiff, xlabel(-0.02(0.01)0.02)
ylabel(-200(100)300) xline(0) yline(0) msize(tiny) mcolor(teal) xtitle("Difference in QALYs")
ytitle("Difference in costs (£)") title("Cost-effectiveness plane") subtitle("Conceptual model
driven analysis (strongly associated predictors)") legend(label(1 "95% confidence ellipse") label(2
"Bootstrap samples")))
graph save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 6
TWICS\Further TWICS analysis novel approach\Analysis included in thesis\twics_ceplane3_large
axisv2_thesis.gph", replace

****************************************
**** Calculate net monetary benefit
****************************************

gen nmb1 = (qaly1 * 20000) - cost1
gen nmb0 = (qaly0 * 20000) - cost0
gen incnmb = (qalyDiff * 20000) - costDiff

*****************************************
**** Calculate point estimates for Chapter 7
*****************************************
tabstat costDiff, stat(mean n)
tabstat qalyDiff, stat(mean n)
tabstat icer, stat(mean n)
tabstat incnmb, stat(mean n)

*** Export .dta file to excel
export excel using "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old
layout\Chapter 6 TWICS\Further TWICS analysis novel approach\Analysis included in
thesis\bootstrap_twics_ceplane3_thesis.xls", replace

**Save dataset
save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 6
TWICS\Further TWICS analysis novel approach\Analysis included in
thesis\twicsceplane3_thesis.dta", replace

*** Work is done externally on CEAC imports in Excel sheet which is saved
*** as 'bootstrap_twics_ceplane3_workings_thesis.xls'

*** CEAC
import excel "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter
6 TWICS\Further TWICS analysis novel approach\Analysis included in
thesis\bootstrap_twics_ceplane3_working_thesis.xls", sheet("stata data") firstrow clear
```

Appendix 13

```
twoway (line Probabilitycosteffective Willingnesstopaythreshold, lcolor(teal)
xlabel(0(20000)120000) ylabel(0(0.2)1.0)title("Cost-effectiveness acceptability curve")
subtitle("Conceptual model driven analysis (strongly associated predictors)"))
graph save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 6
TWICS\Further TWICS analysis novel approach\Analysis included in
thesis\phdtwics_cm_ceac_thesis.gph", replace


**********
* END OF THE FILE*
**********
```

BeatIt syntax

```
**************************************************
**************************************************
**** BEATIT ANALYSIS FOR CONCEPTUAL MODEL APPROACH
**** REGRESSIONS AND ANALYSIS
**************************************************
**************************************************

*********************************************************************
*** This Do file includes the BeatIt conceptual model analysis
*** presented in Chapter 7 of the PhD thesis
*********************************************************************

*** Upload dataset
use "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 5
BeatIt\Further BI analysis\Included in thesis\bi_maindataset_for_cmanalysis_thesis.dta", replace

*** This dataset contains the ITT population (n=141)
*** It takes the BeatIt complete case trial data and replaces missing data with treatment arm
specific mean to enable the use of a full dataset in the case study without complications
*** NB THE CASE STUDY IS FOR ILLUSTRATIVE PURPOSES ONLY

**********************************************
**** Variables in the dataset (n=7)
**********************************************

** trt - treatment arm: StepUp (0) and BeatIt (1)
** activity_total - reported activity
** gds_v3 - GDS-LD scores at 12 months follow up.  NB A high GDS-LD score indicates more
depressive symptoms than a low score ie a lower score indicates improvement in depression.
** qaly_totCC - QALYs for follow-up period
** int_tc - treatment cost
** beatit_tc - total costs
** beatit_tc_nonint - non-treatment costs

*****************************************************************
*****************************************************************
*** 1. REGRESSIONS TO TEST THE ACCURACY OF THE CONCEPTUAL MODEL
*****************************************************************
```

Appendix 13

```
*************************************************************

*******************************************************
** Regression 1) Is activity predicted by treatment arm?
*******************************************************
reg activity_total trt

*****************************************************************************************
** Regression 2) Is depression predicted by treatment arm and is there a possible association with activity?
*****************************************************************************************
reg gds_v3 activity_total trt

***********************************************************
** Regression 3) Is treatment cost predicted by treatment arm?
***********************************************************
reg int_tc trt

*****************************************************************************************
** Regression 4) Is quality of life predicted by depression, and possible associations with activity and treatment arm?
*****************************************************************************************
reg qaly_totCC gds_v3 activity_total trt

*********************************************************************************
** 5) Is non-treatment cost predicted by depression and/or activity, with possible associations with treatment arm?
*********************************************************************************
reg beatit_tc_nonint gds_v3 activity_total trt


*************************************************************************
*************************************************************************
** 2. FINAL CONCEPTUAL MODEL ANALYSIS
*************************************************************************
*************************************************************************

*****************************************************************************************
*****************************************************************************************
*** Cost-effectiveness plane 1 - replicate original cost-effectiveness plane from trial-arm based analysis.  THIS IS FOR CHECKING PURPOSES ONLY
*** Syntax to compute the difference in costs and effects between treatment arms, should be similar to the cost-effectiveness plane in Chapter 3 - slight difference due to different datasets - original is complete case, conceptual model has replaced missing data
*****************************************************************************************
*****************************************************************************************

** Generate a temporary variable for predictor variable
gen temp=.
*** Program to implement set of equations
capture program drop booticer
program define booticer, rclass
tempvar cost0 cost1 qaly0 qaly1

** Replicate original treatment arm analysis
replace temp=trt
regress beatit_tc temp
replace temp=0
predict `cost0'
replace temp=1
predict `cost1'
sum `cost0', meanonly
local mcost0=r(mean)
```

Appendix 13

```
sum `cost1', meanonly
local mcost1=r(mean)

replace temp = trt
regress qaly_totCC temp
replace temp=0
predict `qaly0'
replace temp=1
predict `qaly1'
sum `qaly0',meanonly
local mqaly0=r(mean)
sum `qaly1',meanonly
local mqaly1=r(mean)

return scalar cost1 = `mcost1'
return scalar cost0 = `mcost0'
return scalar costDiff=`mcost1' - `mcost0'
return scalar qaly1 = `mqaly1'
return scalar qaly0 = `mqaly0'
return scalar qalyDiff = `mqaly1' - `mqaly0'
end

*** END OF PROGRAM 'booticer'

*Bootstrap the difference in costs and effects & save the output file
bootstrap       cost1=r(cost1) cost0=r(cost0) costDiff=r(costDiff) ///
                        qaly1=r(qaly1) qaly0=r(qaly0) qalyDiff=r(qalyDiff) ///
                        icer=(r(costDiff)/r(qalyDiff)), ///
                        reps(1000) saving(icer_bootsrep_BI, replace) seed(12345):booticer


** Summarize the differences in costs and effects
use icer_bootsrep_BI.dta, clear
label var costDiff "Incremental cost"
label var qalyDiff "Incremental QALY"
sum costDiff qalyDiff icer
matrix list e(b)
quietly mean costDiff
matrix b = e(b)
local Cost = b[1,1]
quietly mean qalyDiff
matrix b=e(b)
local qaly = b[1,1]
display "ICER: " `Cost'/`qaly'

** Plot the cost/effects on the cost-effectiveness plane
twoway (scatter costDiff qalyDiff, msize(tiny)), yline(0) xline(0) ///
        xlabel(-0.4(0.1)0.2) ylabel(-20000(10000)20000)
graph save "Graph" "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old
layout\Chapter 5 BeatIt\Further BI analysis\Included in thesis\ceplane_step1_thesis.gph"
, replace

*** Export .dta file to excel
export excel using "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old
layout\Chapter 5 BeatIt\Further BI analysis\Included in thesis\ceplane_step1_thesis.xls", replace

** Save dataset
save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 5
BeatIt\Further BI analysis\Included in thesis\ceplane_step1_thesis.dta", replace

*******************************************************************************
*******************************************************************************
*** CONCEPTUAL MODEL 2 (Figure 63)
```

Appendix 13

```
*** Conceptual model analysis with all hypothesised equations
*** Syntax to compute difference in costs and effects based on full conceptual model (including
all regressions identified)
*******************************************************************************
*******************************************************************************

*** Upload same dataset as above
use "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 5
BeatIt\Further BI analysis\Included in thesis\bi_maindataset_for_cmanalysis_thesis.dta", replace

** Generate three temporary variables for predictor variables
gen temp=.
gen act_temp = .
gen dep_temp = .
** Delete the program if it already exists
capture program drop booticer
** Create program
program define booticer, rclass
** Define variables for program
tempvar cost0t cost1t cost0nt cost1nt act0 act1 dep0 dep1 qaly0 qaly1

*****************************************************
*** Equation #1 Treatment arm only predicting activity
*****************************************************
** Take a copy of treatment indicator
replace temp = trt
** Conceptual model regression
regress activity_total temp
** Run as if everyone is in Step-up arm and store prediction in new variable
replace temp=0
predict `act0'
** Run as if everyone is in BeatIt arm and store prediction in new variable
replace temp=1
predict `act1'

*************************************************************
*** Equation #2 Activity and treatment arm predicting depression
*************************************************************
** Take a copy of predictors
replace temp = trt
replace act_temp = activity_total
** Conceptual model regression
regress gds_v3 act_temp temp
** Run as if everyone is in Step-up arm and store prediction in new variable
replace temp=0
** Run with predicted activity from equation #1
replace act_temp = `act0'
predict `dep0'
** Run as if everyone is in BeatIt arm and store prediction in new variable
replace temp=1
** Run with predicted activity from equation #1
replace act_temp = `act1'
predict `dep1'

*****************************************************
*** Equation #3 Treatment arm predicting treatment costs
*****************************************************
** Take a copy of the treatment indicator
replace temp=trt
** Conceptual model regression
regress int_tc temp
** Run as if everyone is in Step-up arm and store prediction in new variable
replace temp=0
```

Appendix 13

```
predict `cost0t'
** Run as if everyone is in BeatIt arm and store prediction in new variable
replace temp=1
predict `cost1t'
** Calculate the mean treatment cost for the bootstrap in each arm
sum `cost0t', meanonly
local mcost0t=r(mean)
sum `cost1t', meanonly
local mcost1t=r(mean)

*************************************************************************
*** Equation #4 QALYs predicted by treatment arm, depression and activity
*************************************************************************
** Take a copy of predictor variables
replace temp = trt
replace act_temp = activity_total
replace dep_temp = gds_v3
** Conceptual model regression
regress qaly_totCC dep_temp act_temp trt
** Run as if everyone is in the Step-up arm and store prediction in new variable
replace temp=0
** Run with predicted depression from equation #2
replace dep_temp = `dep0'
** Run with predicted activity from equation #1
replace act_temp = `act0'
predict `qaly0'
** Run as if everyone is in the BeatIt arm and store prediction in new variable
replace temp=1
** Run with predicted depression from equation #2
replace dep_temp = `dep1'
** Run with predicted activity from equation #1
replace act_temp = `act1'
predict `qaly1'
** Calculate the mean QALYs for the bootstrap run for both arms
sum `qaly0',meanonly
local mqaly0=r(mean)
sum `qaly1',meanonly
local mqaly1=r(mean)

*****************************************************************************************
*** Equation #5 non-treatment costs predicted by treatment arm, activity and depression
*****************************************************************************************
*** Take a copy of predictor variables
replace temp=trt
replace act_temp = activity_total
replace dep_temp = gds_v3
** Conceptual model regression
regress beatit_tc_nonint dep_temp act_temp trt
** Run as if everyone is in the Step-up arm and store prediction in new variable
replace temp=0
** Run with predicted depression from equation #2
replace dep_temp = `dep0'
** Run with predicted activity from equation #1
replace act_temp = `act0'
predict `cost0nt'
** Run as if everyone is in the BeatIt arm and store prediction in new variable
replace temp=1
** Run with predicted depression from equation #2
replace dep_temp = `dep1'
** Run with predicted activity from equation #1
replace act_temp = `act1'
predict `cost1nt'
** Calculate the mean non-treatment cost for the bootstrap run for both arms
```

Appendix 13

```
sum `cost0nt', meanonly
local mcost0nt=r(mean)
sum `cost1nt', meanonly
local mcost1nt=r(mean)

** Instruction to return the predictions and to combine treatment and non-treatment costs into
one total cost. Also to calculate cost and QALY differences
return scalar cost1 = (`mcost1t' + `mcost1nt')
return scalar cost0 = (`mcost0t' + `mcost0nt')
return scalar cost1t = `mcost1t'
return scalar cost0t = `mcost0t'
return scalar cost1nt = `mcost1nt'
return scalar cost0nt = `mcost0nt'
return scalar costDiff = (`mcost1t' - `mcost0t') + (`mcost1nt' - `mcost0nt')
return scalar qaly1 = `mqaly1'
return scalar qaly0 = `mqaly0'
return scalar qalyDiff = `mqaly1' - `mqaly0'
end

*** END OF PROGRAM 'booticer'

** Bootstrap the difference in costs and effects & save the output file
bootstrap       cost1=r(cost1) cost0=r(cost0) costDiff=r(costDiff) ///
                        cost1t=r(cost1t) cost0t=r(cost0t) ///
                        cost1nt=r(cost1nt) cost0nt=r(cost0nt) ///
                        qaly1=r(qaly1) qaly0=r(qaly0) qalyDiff=r(qalyDiff) ///
                        icer=(r(costDiff)/r(qalyDiff)), ///
                        reps(1000) saving(icer_bootsrep_BI2_thesis, replace)
seed(12345):booticer

** Summarize the differences in costs and effects
use icer_bootsrep_BI2_thesis.dta, clear
label var costDiff "Incremental cost"
label var qalyDiff "Incremental QALY"
sum costDiff qalyDiff icer
matrix list e(b)
quietly mean costDiff
matrix b = e(b)
local Cost = b[1,1]
quietly mean qalyDiff
matrix b=e(b)
local qaly = b[1,1]
display "ICER: " `Cost'/`qaly'

*** Plot cost-effectiveness plane
ellip costDiff qalyDiff , c(f) level(95) plot(scatter costDiff qalyDiff, xlabel(-0.2(0.1)0.2) ylabel(-
10000(10000)20000) xline(0) yline(0) msize(tiny) mcolor(teal) xtitle("Difference in QALYs")
ytitle("Difference in costs (£)") title("Cost-effectiveness plane") subtitle("Conceptual model
driven analysis (all regressions)") legend(label(1 "95% confidence ellipse") label(2 "Bootstrap
samples")))
graph save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 5
BeatIt\Further BI analysis\Included in thesis\bi_ceplane2v3_thesis.gph", replace

*** Export .dta file to excel
export excel using "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old
layout\Chapter 5 BeatIt\Further BI analysis\Included in thesis\ceplane_step2v3_thesis.xls",
replace

** Save dataset
save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 5
BeatIt\Further BI analysis\Included in thesis\ceplane_step2v3_thesis.dta", replace
```

Appendix 13

```
*******************************************************************************
*******************************************************************************
*** CONCEPTUAL MODEL 3 (Figure 64)
*** Using strongly associated predictors identified in regressions
*** Syntax to compute difference in costs and effects based on conceptual model identified after
regressions to test accuracy
*******************************************************************************
*******************************************************************************
*** Upload same dataset as above
use "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 5
BeatIt\Further BI analysis\Included in thesis\bi_maindataset_for_cmanalysis_thesis.dta", replace

** Generate temporary variables for predictor variables
gen temp=.
gen act_temp = .
gen dep_temp = .
** Delete program if it already exists
capture program drop booticer
** Create program
program define booticer, rclass
** Define variables for program
tempvar cost0t cost1t cost0nt cost1nt act0 act1 dep0 dep1 qaly0 qaly1

************************************************
** Equation #1 activity predicted by treatment arm
************************************************
** Take a copy of treatment indicator
replace temp = trt
** Conceptual model regression
regress activity_total temp
** Run as if everyone is in Step-up arm and store prediction in new variable
replace temp=0
predict `act0'
** Run as if everyone is in BeatIt arm and store prediction in new variable
replace temp=1
predict `act1'

*********************************************
** Equation #2 depression predicted by activity
*********************************************
** Take a copy of activity indicator
replace act_temp = activity_total
** Conceptual model regression
regress gds_v3 act_temp
** Run as if everyone is in Step-up arm and store prediction in new variable
** Using activity predicted in equation #1
replace act_temp = `act0'
predict `dep0'
** Run as if everyone is in BeatIt arm and store prediction in new variable
** Using activity predicted in equation #1
replace act_temp = `act1'
predict `dep1'

*******************************************************
** Equation #3 treatment costs predicted by treatment arm
*******************************************************
** Take a copy of treatment indicator
replace temp=trt
** Conceptual model regression
regress int_tc temp
** Run as if everyone is in Step-up arm and store prediction in new variable
replace temp=0
predict `cost0t'
```

Appendix 13

```
** Run as if everyone is in BeatIt arm and store prediction in new variable
replace temp=1
predict `cost1t'
** Calculate the mean treatment cost for both arms for bootstrapping exercise
sum `cost0t', meanonly
local mcost0t=r(mean)
sum `cost1t', meanonly
local mcost1t=r(mean)

*****************************************
** Equation #4 QALYs predicted by depression
*****************************************
** Take a copy of depression indicator
replace dep_temp = gds_v3
** Conceptual model regression
regress qaly_totCC dep_temp
** Run as if everyone is in Step-up arm and store prediction in new variable
** Using depression predicted in equation #2
replace dep_temp = `dep0'
predict `qaly0'
** Run as if everyone is in BeatIt arm and store prediction in new variable
** Using depression predicted in equation #2
replace dep_temp = `dep1'
predict `qaly1'
** Calculate the mean QALYs for both arms for bootstrapping exercise
sum `qaly0',meanonly
local mqaly0=r(mean)
sum `qaly1',meanonly
local mqaly1=r(mean)

*********************************************************
** Equation #5 non-treatment costs predicted by activity
*********************************************************
** Take a copy of activity indicator
replace act_temp = activity_total
** Conceptual model regression
regress beatit_tc_nonint act_temp
** Run as if everyone is in Step-up arm and store prediction in new variable
** Using activity predicted in equation #1
replace act_temp = `act0'
predict `cost0nt'
** Run as if everyone is in BeatIt arm and store prediction in new variable
** Using activity predicted in equation #1
replace act_temp = `act1'
predict `cost1nt'
** Calculate mean non-treatment costs for both arms for bootstrapping exercise
sum `cost0nt', meanonly
local mcost0nt=r(mean)
sum `cost1nt', meanonly
local mcost1nt=r(mean)

** Instructions to combine treatment and non-treatment costs into total costs and to calculate
differences in costs and QALYs
return scalar cost1 = (`mcost1t' + `mcost1nt')
return scalar cost0 = (`mcost0t' + `mcost0nt')
return scalar cost1t = `mcost1t'
return scalar cost0t = `mcost0t'
return scalar cost1nt = `mcost1nt'
return scalar cost0nt = `mcost0nt'
return scalar costDiff = (`mcost1t' - `mcost0t') + (`mcost1nt' - `mcost0nt')
return scalar qaly1 = `mqaly1'
return scalar qaly0 = `mqaly0'
return scalar qalyDiff = `mqaly1' - `mqaly0'
```

Appendix 13

```
end

** END OF PROGRAM 'booticer'

*Bootstrap the difference in costs and effects & save the output file
bootstrap       cost1=r(cost1) cost0=r(cost0) costDiff=r(costDiff) ///
                        cost1t=r(cost1t) cost0t=r(cost0t) ///
                        cost1nt=r(cost1nt) cost0nt=r(cost0nt) ///
                        qaly1=r(qaly1) qaly0=r(qaly0) qalyDiff=r(qalyDiff) ///
                        icer=(r(costDiff)/r(qalyDiff)), ///
                        reps(1000) saving(icer_bootsrep_BI3_thesis, replace)
seed(12345):booticer

*summarize the differences in costs and effects
use icer_bootsrep_BI3_thesis.dta, clear
label var costDiff "Incremental cost"
label var qalyDiff "Incremental QALY"
sum costDiff qalyDiff icer
matrix list e(b)
quietly mean costDiff
matrix b = e(b)
local Cost = b[1,1]
quietly mean qalyDiff
matrix b=e(b)
local qaly = b[1,1]
display "ICER: " `Cost'/`qaly'

*** Plot cost-effectiveness plane
ellip costDiff qalyDiff , c(f) level(95) plot(scatter costDiff qalyDiff, xlabel(-0.2(0.1)0.2) ylabel(-
10000(10000)20000) xline(0) yline(0) msize(tiny) mcolor(teal) xtitle("Difference in QALYs")
ytitle("Difference in costs (£)") title("Cost-effectiveness plane") subtitle("Conceptual model
driven analysis (strongly associated predictors)") legend(label(1 "95% confidence ellipse") label(2
"Bootstrap samples")))
graph save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 5
BeatIt\Further BI analysis\Included in thesis\bi_ceplane3v2_thesis.gph", replace

****************************************
**** Calculate net monetary benefit
****************************************
gen nmb1 = (qaly1 * 20000) - cost1
gen nmb0 = (qaly0 * 20000) - cost0
gen incnmb = (qalyDiff * 20000) - costDiff

*********************************************
**** Calculate point estimates for Chapter 7
*********************************************
tabstat costDiff, stat(mean n)
tabstat qalyDiff, stat(mean n)
tabstat icer, stat(mean n)
tabstat incnmb, stat(mean n)

*** Export .dta file to excel
export excel using "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old
layout\Chapter 5 BeatIt\Further BI analysis\Included in thesis\ceplane_step3v2thesis.xls", replace

** Save dataset
save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 5
BeatIt\Further BI analysis\Included in thesis\ceplane_step3v2_thesis.dta", replace

*** To produce the cost-effectiveness acceptability curve the above output is worked on to
produce data points for CEAC
** asn saved as 'ceplane_step3v2thesis_ceacworking'
```

Appendix 13

```
*** Load data for CEAC
import excel "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter
5 BeatIt\Further BI analysis\Included in thesis\ceplane_step3v2thesis_ceacworking.xls",
sheet("stata data") firstrow clear

twoway (line Probabilitycosteffective Willingnesstopaythreshold, lcolor(teal)
xlabel(0(20000)120000) ylabel(0(0.2)1.0)title("Cost-effectiveness acceptability curve")
subtitle("Conceptual model driven analysis (strongly associated predictors)"))
graph save "C:\Users\nmm13h\OneDrive - University of Glasgow\PhD\Writing\Old layout\Chapter 5
BeatIt\Further BI analysis\Included in thesis\phdbi_cm_ceac_thesis.gph", replace

******************
*** END OF FILE ***
******************
```

List of References

1. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. Third ed. New York, US: Oxford University Press; 2005.
2. Fenwick E, Byford S. A guide to cost-effectiveness acceptability curves. British Journal of Psychiatry. 2005 Aug;187:106-8.
3. Briggs A, Claxton K, Sculpher M. Decision Modelling for Health Economic Evaluation. 2011 ed. Gray A, Briggs A, editors. New York: Oxford University Press; 2006.
4. Glick H, Doshi J, Sonnad S, Polsky D. Economic Evaluation in Clinical Trials. Second ed: Oxford University Press; 2015.
5. Ramsey SD, Willke RJ, Glick H, Reed SD, Augustovski F, Jonsson B, et al. Cost-Effectiveness Analysis Alongside Clinical Trials II-An ISPOR Good Research Practices Task Force Report. Value in Health. 2015 Mar;18(2):161-72.
6. Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. Bmj-British Medical Journal. 2011 Apr 7;342.
7. Raftery J, Williams HC, Clarke A, Thornton J, Norrie J, Snooks H, et al. 'Not clinically effective but cost-effective'-paradoxical conclusions in randomised controlled trials with 'doubly null' results: a cross-sectional study. Bmj Open. 2020 Jan;10(1).
8. Hollingworth W, McKell-Redwood D, Hampson L, Metcalfe C. Cost-utility analysis conducted alongside randomized controlled trials: Are economic end points considered in sample size calculations and does it matter? Clinical Trials. 2013 Feb;10(1):43-53.
9. Backhouse ME. Use of randomised controlled trials for producing cost-effectivehess evidence - Potential impact of design choices on sample size and study duration. Pharmacoeconomics. 2002 2002;20(15):1061-77.
10. Anderson R, Hardwick R. Realism and resources: Towards more explanatory economic evaluation. Evaluation. 2016 Jul;22(3):323-41.
11. Zeigler BP. Theory of Modeling and Simulation. New York: Wiley; 1976.
12. Abdelmegid MA, Gonzalez VA, Naraghi AM, O'Sullivan M, Walker CG, Poshdar M, editors. Towards a conceptual modeling framework for construction simulation. Winter Simulation Conference (WSC); 2017 2017 Dec 03-06; Las Vegas, NV; 2017.
13. Montevechi JAB, Friend JD. Using A Soft Systems Methodology Framework To Guide The Conceptual Modeling Process In Discrete Event Simulation. In: Laroque C, Himmelspach J, Pasupathy R, editors. 2012 Winter Simulation Conference2012.
14. Morgan A. Basic Guidance for Cross-Cutting Tools: Conceptual Models, WWF Programme Standards. 2016 [Accessed 1 August 2016]; Available from: http://wwf.panda.org/what_we_do/how_we_work/programme_standards/index.cfm.
15. Brassington FC, Younger PL. A proposed framework for hydrogeological conceptual modelling. Water and Environment Journal. 2010 Dec;24(4):261-73.
16. Gray K, Sockolow P. Conceptual Models in Health Informatics Research: A Literature Review and Suggestions for Development. Jmir Medical Informatics. 2016 Jan-Mar;4(1).
17. Robinson S. Conceptual modelling for simulation Part I: definition and requirements. Journal of the Operational Research Society. 2008 Mar;59(3):278-90.

References

18. Furian N, O'Sullivan M, Walker C, Vossner S, Neubacher D. A conceptual modeling framework for discrete event simulation using hierarchical control structures. Simulation Modelling Practice and Theory. 2015 Aug;56:82-96.
19. Hesch W. Conceptual Model Development for Modflow or Feflow Models. 2017.
20. Gross J. National Park Service Inventory and Monitoring Program: Developing Conceptual Models for Monitoring Programs. 2003 [Accessed 29 July 2016]; Available from: http://science.nature.nps.gov/im/monitor/docs/Conceptual_modelling.pdf.
21. Chwif L, Banks J, de Moura Filho JP, Santini B. A framework for specifying a discrete-event simulation conceptual model. Journal of Simulation. 2013 Feb;7(1):50-60.
22. Squires H, Chilcott J, Akehurst R, Burr J, Kelly MP. A Framework for Developing the Structure of Public Health Economic Models. Value in Health. 2016 Jul-Aug;19(5):588-601.
23. Tappenden P. Conceptual modelling for health economic model development. 2012 2012.
24. Brennan A, Akehurst R. Modelling in health economic evaluation - What is its place? What is its value? Pharmacoeconomics. 2000 May;17(5):445-59.
25. Chilcott J, Tappenden P, Rawdin A, Johnson M, Kaltenthaler E, Paisley S, et al. Avoiding and identifying errors in health technology assessment models: qualitative study and methodological review. Health Technology Assessment. 2010 May;14(25):1-+.
26. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M, et al. Conceptualizing a Model: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. Medical Decision Making. 2012 Sep-Oct;32(5):678-89.
27. National Institute for Health and Care Excellence (NICE). Interim methods guide for developing service guidance 2014. NICE; 2016; Available from: https://www.nice.org.uk/process/pmg8/chapter/1-introduction.
28. Reed MS, Kenter J, Bonn A, Broad K, Burt TP, Fazey IR, et al. Participatory scenario development for environmental management: A methodological framework illustrated with experience from the UK uplands. Journal of Environmental Management. 2013 Oct 15;128:345-62.
29. Reidsma P, Konig H, Feng S, Bezlepkina I, Keulen Hv, Ittersum MKv, et al. A methodological framework for sustainability impact assessment of land use policies in developing countries: re-using and complementing approaches. Proceedings of the Conference on integrated assessment of agriculture and sustainable development: Setting the Agenda for Science and Policy (AgSAP 2009), Hotel Zuiderduin, Egmond aan Zee, The Netherlands, 10-12 March 2009. 2009 2009:138-9.
30. Lee J, Jang S. A methodological framework for instructional design model development: Critical dimensions and synthesized procedures. Etr&D-Educational Technology Research and Development. 2014 Dec;62(6):743-65.
31. George H, Bosc PM, Even MA, Belieres JF, Bessou C. WAW proposed methodological framework to monitor agricultural structural transformations and their contributions to sustainable development. Producing and reproducing farming systems New modes of organisation for sustainable food systems of tomorrow 10th European IFSA Symposium, Aarhus, Denmark, 1-4 July 2012. 2012 2012.
32. Halbe J, Pahl-Wostl C, Adamowski J. A methodological framework to support the initiation, design and institutionalization of participatory modeling

References

processes in water resources management. Journal of Hydrology. 2018 Jan;556:701-16.
33. Battini D, Faccio M, Persona A, Sgarbossa F. New methodological framework to improve productivity and ergonomics in assembly system design. International Journal of Industrial Ergonomics. 2011 Jan;41(1):30-42.
34. Anagnostou A, Taylor SJE. A distributed simulation methodological framework for OR/MS applications. Simulation Modelling Practice and Theory. 2017 Jan;70:101-19.
35. Achillas C, Aidonis D, Iakovou E, Thymianidis M, Tzetzis D. A methodological framework for the inclusion of modern additive manufacturing into the production portfolio of a focused factory. Journal of Manufacturing Systems. 2015 Oct;37:328-39.
36. Procházka J, Melichar J. Methodological Framework for Operational Risk Assessment. Vojenské rozhledy. 2017;26:19-34.
37. Pahl-Wostl C, Holtz G, Kastens B, Knieper C. Analyzing complex water governance regimes: the Management and Transition Framework. Environmental Science & Policy. 2010 Nov;13(7):571-81.
38. Kallio H, Pietila A-M, Johnson M, Kangasniemi M. Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. Journal of Advanced Nursing. 2016 Dec;72(12):2954-65.
39. Rivera SC, Kyte DG, Aiyegbusi OL, Keeley TJ, Calvert MJ. Assessing the impact of healthcare research: A systematic review of methodological frameworks. Plos Medicine. 2017 Aug;14(8).
40. Global I. What is Methodological Framework | IGI Global. 2019; Available from: https://www.igi-global.com/dictionary/methodological-framework/18485.
41. JRC For-learn,. 2019.
42. Rodgers M, Thomas S, Harden M, Parker G, Street A, Eastwood A. Developing a methodological framework for organisational case studies: a rapid review and consensus development process. Health Serv Deliv Res. 2016;4(1).
43. Robinson S. Conceptual modelling for simulation Part II: a framework for conceptual modelling. Journal of the Operational Research Society. 2008 Mar;59(3):291-304.
44. van der Zee DJ. An Integrated Conceptual Modeling Framework For Simulation Linking Simulation - Modeling To The Systems Engineering Process. In: Laroque C, Himmelspach J, Pasupathy R, editors. 2012 Winter Simulation Conference2012.
45. McMeekin N, Wu O, Germeni E, Briggs A. How methodological frameworks are being developed: evidence from a scoping review. BMC medical research methodology. 2020 2020 Jun;20(1):173-.
46. Kluge E-HW. Resource allocation in healthcare: implications of models of medicine as a profession. MedGenMed : Medscape general medicine. 2007 2007 Mar;9(1):57-.
47. Mitton C, Donaldson C. Health care priority setting: principles, practice and challenges. Cost-effectiveness and resource setting. 2004;2(3).
48. Goodman CS. HTA 101: Introduction to Health Technology Assessment Bethesda, MD: National Library of Medicine (US) 2014
49. McIntosh E, Clarke P, Frew E, Louviere J. Applied methods of cost-benefit analysis in health care. Oxford: Oxford University Press Incorporated; 2010.

References

50. Coast J, Smith RD, Lorgelly P. Welfarism, extra-welfarism and capability: The spread of ideas in health economics. Social Science & Medicine. 2008 Oct;67(7):1190-8.
51. Culyer A. The Normative Economics of Health Care Finance and Provision. Oxford Review of Economic Policy. 1989;5(1):34-58.
52. Karimi M, Brazier J, Basarir H. The Capability Approach: A Critical Review of Its Application in Health Economics. Value in Health. 2016 Sep-Oct;19(6):795-9.
53. Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. BMJ 336:1281. 2008.
54. Vreman RA, Naci H, Goettsch WG, Mantel-Teeuwisse AK, Schneeweiss SG, Leufkens HGM, et al. Decision Making Under Uncertainty: Comparing Regulatory and Health Technology Assessment Reviews of Medicines in the United States and Europe. Clinical Pharmacology & Therapeutics. 2020 Aug;108(2):350-7.
55. National Institute for Health and Care Excellence. Guide to the Methods of Technology Appraisal: 2013. [Accessed 29 October 2015]; Available from: https://www.nice.org.uk/article/pmg9/chapter/foreword.
56. Parody-Rua E, Rubio-Valera M, Guevara-Cuellar C, Gomez-Lumbreras A, Casajuana-Closas M, Carbonell-Duacastella C, et al. Economic Evaluations Informed Exclusively by Real World Data: A Systematic Review. International Journal of Environmental Research and Public Health. 2020 Feb;17(4).
57. UK. NHS. Clinical trials. 2019; Available from: https://www.nhs.uk/conditions/clinical-trials/.
58. Li YYR, Mai H, Trudeau ME, Mittmann N, Chiasson K, Chan KKW, et al. Reimbursement recommendations for cancer drugs supported by phase II evidence in Canada. Current Oncology. 2020 Oct;27(5):E495-E500.
59. Hariton E, Locascio JJ. Randomised controlled trials - the gold standard for effectiveness research Study design: randomised controlled trials. Bjog-an International Journal of Obstetrics and Gynaecology. 2018 Dec;125(13):1716-.
60. Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? Health Economics. 2006 Jul;15(7):677-87.
61. Ramsey S, Willke R, Briggs A, Brown R, Buxton M, Chawla A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: The ISPOR RCT-CEA task force report. Value in Health. 2005 Sep-Oct;8(5):521-33.
62. O'Sullivan AK, Thompson D, Drummond MF. Collection of health-economic data alongside clinical trials: Is there a future for piggyback evaluations? Value in Health. 2005 Jan-Feb;8(1):67-79.
63. Excellence NIfHaC. Developing NICE guidelines: the manual | 7-Incorporating-economic-evaluation | Guidance and guidelines | NICE. NICE; 2015 [1 May 2015]; Available from: http://www.nice.org.uk/article/PMG20/chapter/7-incorporating-economic-evaluation.
64. Ciani O, Buyse M, Drummond M, Rasi G, Saad ED, Taylor RS. Time to Review the Role of Surrogate End Points in Health Policy: State of the Art and the Way Forward. Value in Health. 2017 Mar;20(3):487-95.
65. Prentice RL. Surrogate Endpoints In Clinical-Trials - Definition And Operational Criteria. Statistics in Medicine. 1989 Apr;8(4):431-40.
66. Hemming C, Constable L, Goulao B, Kilonzo M, Boyers D, Elders A, et al. Surgical interventions for uterine prolapse and for vault prolapse: the two VUE RCTs. Health Technology Assessment. 2020 Mar;24(13):1-+.

References

67. Bellary S, Krishnankutty B, Latha MS. Basics of case report form designing in clinical research. Perspectives in clinical research. 2014 2014-Oct;5(4):159-66.
68. Tang K. Estimating Productivity Costs in Health Economic Evaluations: A Review of Instruments and Psychometric Evidence. Pharmacoeconomics. 2015 Jan;33(1):31-48.
69. Ridyard CH, Hughes DA, Team D. Development of a Database of Instruments for Resource-Use Measurement: Purpose, Feasibility, and Design. Value in Health. 2012 Jul-Aug;15(5):650-5.
70. Reilly MC, Zbrozek AS, Dukes EM. The Validity And Reproducibility Of A Work Productivity And Activity Impairment Instrument. Pharmacoeconomics. 1993 Nov;4(5):353-65.
71. Lerner D, Amick BC, Rogers WH, Malspeis S, Bungay K, Cynn D. The work limitations questionnaire. Medical Care. 2001 Jan;39(1):72-85.
72. Prasad M, Wahlqvist P, Shikiar R, Shih YCT. A review of self-report instruments measuring health-related work productivity - A patient-reported outcomes perspective. Pharmacoeconomics. 2004 2004;22(4):225-44.
73. EuroQol G. EuroQol--a new facility for the measurement of health-related quality of life. Health policy (Amsterdam, Netherlands). 1990 1990-Dec;16(3):199-208.
74. EuroQol. EuroQol - About EQ-5D. 2017; Available from: http://www.euroqol.org/about-eq-5d.html.
75. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI): concepts, measurement properties and applications. Health and quality of life outcomes. 2003 2003 Oct;1:54-.
76. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of Economics. 2002;21(2):271-92.
77. Wille N, Badia X, Bonsel G, Burstrom K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Quality of Life Research. 2010 Aug;19(6):875-86.
78. Ravens-Sieberer U, Karow A, Barthel D, Klasen F. How to assess quality of life in child and adolescent psychiatry. Dialogues in Clinical Neuroscience. 2014 Jun;16(2):147-58.
79. Brazier J, Ratcliffe J, Salomon J, Tsuchiya A. Measuring and valuing health benefits for economic evaluation: Oxford University Press; 2017.
80. Kaplan RM, Feeny D, Revicki DA. Methods For Assessing Relative Importance In Preference Based Outcome Measures. Quality of Life Research. 1993 Dec;2(6):467-75.
81. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. European Journal of Health Economics. 2010 Apr;11(2):215-25.
82. Thorn JC, Davies CF, Brookes ST, Noble SM, Dritsaki M, Gray E, et al. Content of Health Economics Analysis Plans (HEAPs) for Trial-Based Economic Evaluations: Expert Delphi Consensus Survey. Value in Health. 2020 Apr;24(4):539-47.
83. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Economics. 2005 May;14(5):487-96.
84. Faria R, Gomes M, Epstein D, White IR. A Guide to Handling Missing Data in Cost-Effectiveness Analysis Conducted Within Randomised Controlled Trials. Pharmacoeconomics. 2014 Dec;32(12):1157-70.

References

85. Thorn J, Coast J, Andronis L. Interpretation of the Expected Value of Perfect Information and Research Recommendations: A Systematic Review and Empirical Investigation. Medical Decision Making. 2016 Apr;36(3):285-95.
86. Fenwick E, Steuten L, Knies S, Ghabri S, Basu A, Murray JF, et al. Value of Information Analysis for Research Decisions-An Introduction: Report 1 of the ISPOR Value of Information Analysis Emerging Good Practices Task Force. Value in Health. 2020 Feb;23(2):139-50.
87. Rothery C, Strong M, Koffijberg H, Basu A, Ghabri S, Knies S, et al. Value of Information Analytical Methods: Report 2 of the ISPOR Value of Information Analysis Emerging Good Practices Task Force. Value in Health. 2020 Mar;23(3):277-86.
88. Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analyses. Jama-Journal of the American Medical Association. 1996 Oct 23;276(16):1339-41.
89. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. British Medical Journal. 1996 Aug 3;313(7052):275-83.
90. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. European Journal of Health Economics. 2013 Jun;14(3):367-72.
91. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Deutsche Medizinische Wochenschrift. 2011 Apr;136(15):E26-E.
92. Sculpher M. Clinical Trials Provide Essential Evidence, but Rarely Offer a Vehicle for Cost-Effectiveness Analysis. Value in Health. 2015 Mar;18(2):141-2.
93. Hughes D, Charles J, Dawoud D, Edwards RT, Holmes E, Jones C, et al. Conducting Economic Evaluations Alongside Randomised Trials: Current Methodological Issues and Novel Approaches. Pharmacoeconomics. 2016 May;34(5):447-61.
94. Marshall DA, Hux M. Design and Analysis Issues for Economic Analysis Alongside Clinical Trials. Medical Care. 2009 Jul;47(7):S14-S20.
95. Zhong B. How to calculate sample size in randomized controlled trial? Journal of thoracic disease. 2009 2009-Dec;1(1):51-4.
96. Bell ML. New guidance to improve sample size calculations for trials: eliciting the target difference. Trials. 2018 Nov 5;19.
97. Briggs A. Economic evaluation and clinical trials: size matters - The need for greater power in cost analyses poses an ethical dilemma. British Medical Journal. 2000 Dec 2;321(7273):1362-3.
98. Bodrogi J, Kalo Z. Principles of pharmacoeconomics and their impact on strategic imperatives of pharmaceutical research and development. British Journal of Pharmacology. 2010 Apr;159(7):1367-73.
99. Briggs AH, O'Brien BJ, Blackhouse G. Thinking outside the box: Recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. Annual Review of Public Health. 2002 2002;23:377-401.
100. Briggs AH, O'Brien BJ. The death of cost-minimization analysis? Health Economics. 2001 Mar;10(2):179-84.
101. Salter KL, Kothari A. Using realist evaluation to open the black box of knowledge translation: a state-of-the-art review. Implementation Science. 2014 Sep 5;9.
102. KPMG. Impact and Value of the NIHR Clinical Research Network 2019 (infographic summarising key findings). [cited 2021 16 February 2021]; Available

References

from: https://www.nihr.ac.uk/documents/impact-and-value-of-the-nihr-clinical-research-network-2019-infographic-summarising-key-findings/22486.
103. Brennan VK, Dixon S. Incorporating Process Utility into Quality Adjusted Life Years: A Systematic Review of Empirical Studies. Pharmacoeconomics. 2013 Aug;31(8):677-91.
104. Brennan VK, Jones G, Radley S, Dixon S. Incorporating Process Utility into Cost-Effectiveness Analysis via a Bolt-On Domain to the SF-6D: An Exploratory Study. Applied Health Economics and Health Policy. 2021.
105. Dalkin SM, Greenhalgh J, Jones D, Cunningham B, Lhussier M. What's in a mechanism? Development of a key concept in realist evaluation. Implementation Science. 2015 Apr 16;10.
106. Wright S. Correlation and causation Part I. Method of path coefficients. Journal of Agricultural Research. 1920 Oct-Mar;20:0557-85.
107. Olkin I, Sampson AR. Multivariate analysis: overview. International Encyclopedia of the Social & Behavioral Sciences2001. p. 10240-7.
108. Lewis M, Kuerbis A. An Overview of Causal Directed Acyclic Graphs for Substance Abuse
Researchers. Journal of Drug and Alcohol Research. 2016;5.
109. Devereux G, Cotton S, Fielding S, McMeekin N, Barnes PJ, Briggs A, et al. Low-dose oral theophylline combined with inhaled corticosteroids for people with chronic obstructive pulmonary disease and high risk of exacerbations: a RCT. . *Health Technol Assess*. 2019;23(37).
110. Devereux G, Cotton S, Fielding S, McMeekin N, Bames PJ, Briggs A, et al. Effect of Theophylline as Adjunct to Inhaled Corticosteroids on Exacerbations in Patients With COPD A Randomized Clinical Trial. Jama-Journal of the American Medical Association. 2018 Oct 16;320(15):1548-59.
111. Devereux G, Cotton S, Barnes P, Briggs A, Burns G, Chaudhuri R, et al. Use of low-dose oral theophylline as an adjunct to inhaled corticosteroids in preventing exacerbations of chronic obstructive pulmonary disease: study protocol for a randomised controlled trial. Trials. 2015 Jun 10;16.
112. World Health Organization (WHO). Burden of COPD. Available from: https://www.who.int/respiratory/copd/burden/en/.
113. Health and Safety Executive (HSE). Work-related Chronic Obstructive Pulmonary Disease (COPD) statistics in
Great Britain 2019. 2019 [1 October 2020]; Available from: https://www.hse.gov.uk/statistics/causdis/copd.pdf.
114. UK COPD treatment: failing to progress. Lancet. 2018 Apr 21;391(10130):1550-.
115. Burge S, Wedzicha JA. COPD exacerbations: definitions and classifications. European Respiratory Journal. 2003 Jun;21:46S-53S.
116. Soler-Cataluna JJ, Martinez-Garcia MA, Sanchez PR, Salcedo E, Navarro M, Ochando R. Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. Thorax. 2005 Nov;60(11):925-31.
117. Barnes PJ, Ito K, Adcock IM. Corticosteroid resistance in chronic obstructive pulmonary disease: inactivation of histone deacetylase. Lancet. 2004 Feb 28;363(9410):731-3.
118. Chisholm D, Knapp MRJ, Knudsen HC, Amaddeo F, Gaite L, van Wijngaarden B, et al. Client Socio-Demographic and Service Receipt Inventory - European Version: Development of an instrument for international research - EPSILON Study 5. British Journal of Psychiatry. 2000 Jul;177:S28-S33.
119. British National Formulary Publications. BNF Publications. 2016 [Accessed 21 January 2016]; Available from: http://www.bnf.org/.

References

120. Department of Health and Social Care. NHS reference costs 2015 to 2016 - GOV.UK. 2018 [25.1.18]; Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016.
121. Information Services Division Scotland. Finance | Costs | Health Topics | ISD Scotland. 2017; Available from: http://www.isdscotland.org/Health-Topics/Finance/Costs/.
122. Curtis L. Unit Costs of Health and Social Care 2016 | PSSRU. 2018; Available from: http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2016/.
123. Oostenbrink JB, Rutten-van Molken M, Monz BU, FitzGerald JM. Probabilistic markov model to assess the cost-effectiveness of bronchodilator therapy in COPD patients in different countries. Value in Health. 2005 Jan-Feb;8(1):32-46.
124. Scott A, Simoens S, Heaney D, O'Donnell CA, Thomson H, Moffat KJ, et al. What does GP out of hours care cost? - An analysis of different models of out of hours care in Scotland. Scottish Medical Journal. 2004 May;49(2):61-6.
125. Choi JI, Han KD, Kim SJ, Kim MJ, Shin JE, Lee HN. Relationship between delivery history and health-related quality of life in menopausal South Korean women: The Korea National Health and Nutrition Examination Surveys. Maturitas. 2016;92:24-9.
126. Jahoda A, Hastings R, Hatton C, Cooper S-A, Dagnan D, Zhang R, et al. Comparison of behavioural activation with guided self-help for treatment of depression in adults with intellectual disabilities: a randomised controlled trial. The lancet Psychiatry. 2017 2017-Nov-13.
127. Jahoda A, Hastings R, Hatton C, Cooper S-A, McMeekin N, Dagnan D, et al. Behavioural activation versus guided self-help for depression in adults with learning disabilities: the Beatlt RCT. Health Technology Assessment. 2018 Sep;22(53):VII-+.
128. Jahoda A, Melville C, Cooper S-A, Hastings R, Briggs A, Dagnan D, et al. BEAT-IT: Comparing a behavioural activation treatment for depression in adults with intellectual disabilities with an attention control: study protocol for a randomised controlled trial. Trials. 2015 Dec 30;16.
129. Herrman H, Kieling C, McGorry P, Horton R, Sargent J, Patel V. Reducing the global burden of depression: a Lancet-World Psychiatric Association Commission. Lancet. 2019 Jun 15;393(10189):E42-E3.
130. McCrone P, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. PAYING THE PRICE
The cost of mental health care in England to 2026. King's Fund; 2008.
131. Information Services Division NHS National Services Scotland. Medicines used in Mental Health Years 2007/08 – 2017/18. 2018; Available from: https://www.isdscotland.org/Health-Topics/Prescribing-and-Medicines/Publications/2018-10-09/2018-10-09-PrescribingMentalHealth-Report.pdf.
132. Cooper S-A, Smiley E, Morrison J, Williamson A, Allan L. An epidemiological investigation of affective disorders with a population-based cohort of 1023 adults with intellectual disabilities. Psychological Medicine. 2007 Jun;37(6):873-82.
133. Collishaw S, Maughan B, Pickles A. Affective problems in adults with mild learning disability: the roles of social disadvantage and ill health. British Journal of Psychiatry. 2004 Oct;185:350-1.

References

134. Cooney P, Jackman C, Coyle D, O'Reilly G. Computerised cognitive behavioural therapy for adults with intellectual disability: randomised controlled trial. British Journal of Psychiatry. 2017 Aug;211(2):95-+.
135. Lejuez CW, Hopko DR, Acierno R, Daughters SB, Pagoto SL. Ten Year Revision of the Brief Behavioral Activation Treatment for Depression: Revised Treatment Manual. Behavior Modification. 2011 Mar;35(2):111-61.
136. Ekers D, Richards D, McMillan D, Bland JM, Gilbody S. Behavioural activation delivered by the non-specialist: phase II randomised controlled trial. British Journal of Psychiatry. 2011 Jan;198(1):66-72.
137. Richards DA, Ekers D, McMillan D, Taylor RS, Byford S, Warren FC, et al. Cost and Outcome of Behavioural Activation versus Cognitive Behavioural Therapy for Depression (COBRA): a randomised, controlled, non-inferiority trial. Lancet. 2016 Aug 27;388(10047):871-80.
138. Dobson KS, Dimidjian S, Kohlenberg RJ, Rizvi SL, Dunner DL, Jacobson NS, et al. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the prevention of relapse and recurrence in major depression. Journal of Consulting and Clinical Psychology. 2008 Jun;76(3):468-77.
139. National Institute for Health and Care Excellence. Chronic obstructive pulmonary disease in over 16s: diagnosis and management 2010 [Date accessed 10 November 2015]; Available from: http://www.nice.org.uk/guidance/cg101.
140. Jahoda A, Melville CA, Pert C, Cooper SA, Lynn H, Williams C, et al. A feasibility study of behavioural activation for depressive symptoms in adults with intellectual disabilities. Journal of Intellectual Disability Research. 2015 Nov;59(11):1010-21.
141. Cuthill FM, Espie CA, Cooper SA. Development and psychometric properties of the Glasgow Depression Scale for people with a Learning Disability - Individual and carer supplement versions. British Journal of Psychiatry. 2003 Apr;182:347-53.
142. Willner P, Rose J, Jahoda A, Kroese BS, Felce D, MacMahon P, et al. A cluster randomised controlled trial of a manualised cognitive-behavioural anger management intervention delivered by supervised lay therapists to people with intellectual disabilities. Health Technology Assessment. 2013 May;17(21):1-+.
143. Curtis L. PSSRU | Unit Costs of Health and Social Care 2015. 2016; Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2015/.
144. Health Do. NHS reference costs 2014 to 2015 - Publications - GOV.UK. 2016.
145. Dolan P. Modeling valuations for EuroQol health states. Medical Care. 1997 Nov;35(11):1095-108.
146. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. Bmc Medical Research Methodology. 2018 Nov 19;18.
147. Tricco AC, Lillie E, Zarin W, O'Brien K, Colquhoun H, Kastner M, et al. A scoping review on the conduct and reporting of scoping reviews. Bmc Medical Research Methodology. 2016 Feb 9;16.
148. Armstrong R, Hall BJ, Doyle J, Waters E. 'Scoping the scope' of a cochrane review. Journal of Public Health. 2011 Mar;33(1):147-50.
149. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. International Journal of Evidence-Based Healthcare. 2015 Sep;13(3):141-6.

References

150. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology. 2005;8(1):19-32.
151. Jahan N, Naveed S, Zeshan M, Tahir MA. How to Conduct a Systematic Review: A Narrative Literature Review. Cureus. 2016 Nov;8(11).
152. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implementation Science. 2010 Sep;5:9.
153. Colquhoun HL, Levac D, O'Brien KK, Straus S, Tricco AC, Perrier L, et al. Scoping reviews: time for clarity in definition, methods, and reporting. Journal of Clinical Epidemiology. 2014 Dec;67(12):1291-4.
154. Benzies KM, Premji S, Hayden KA, Serrett K. State-of-the-evidence reviews: Advantages and challenges of including grey literature. Worldviews on Evidence-Based Nursing. 2006 2006;3(2):55-61.
155. Adams J, Hillier-Brown FC, Moore HJ, Lake AA, Araujo-Soares V, White M, et al. Searching and synthesising 'grey literature' and 'grey information' in public health: critical reflections on three case studies. Systematic reviews. 2016 2016 Sep;5(1):164-.
156. Godin K, Stapleton J, Kirkpatrick SI, Hanning RM, Leatherdale ST. Applying systematic review search methods to the grey literature: a case study examining guidelines for school-based breakfast programs in Canada. Systematic reviews. 2015 2015 Oct;4:138.
157. Linek SB, Schwarz D, Bopp M, Albert D. When Playing Meets Learning: Methodological Framework for Designing Educational Games. Web Information Systems and Technologies. 2010 2010;45:73-85.
158. Stratigea A, Papadopoulou CA. Foresight Analysis at the Regional Level - A Participatory Methodological Framework Journal of Management and Strategy. 2013;4(2).
159. Tondel K, Niederer SA, Land S, Smith NP. Insight into model mechanisms through automatic parameter fitting: a new methodological framework for model development. Bmc Systems Biology. 2014 May 20;8.
160. Tappenden P, Chilcott J, Brennan A, Squires H, Stevenson M. Whole Disease Modeling to Inform Resource Allocation Decisions in Cancer: A Methodological Framework. Value in Health. 2012 Dec;15(8):1127-36.
161. Lopes AMB, Ruiz-Cecilia R. Designing Technology-Mediated Tasks for Language Teaching: A Methodological Framework. Hacettepe Universitesi Egitim Fakultesi Dergisi-Hacettepe University Journal of Education. 2017 Apr;32(2):265-79.
162. Stremke S, Van Kann F, Koh J. Integrated Visions (Part I): Methodological Framework for Long-term Regional Design. European Planning Studies. 2012 2012;20(2):305-19.
163. Dean E, Taylor MJ, Francis H, Lisboa P, Appleton D, Jones M. A Methodological Framework for Geographic Information Systems Development. Systems Research and Behavioral Science. 2017 Nov-Dec;34(6):759-72.
164. Sun Y, Strobel J. Elementary Engineering Education (EEE) Adoption and Expertise Development Framework: An Inductive and Deductive Study. Journal of Pre-College Engineering Education Research (J-PEER). 2013;3(1).
165. Nicod E, Kanavos P. Developing an evidence-based methodological framework to systematically compare HTA coverage decisions: A mixed methods study. Health Policy. 2016 Jan;120(1):35-45.
166. Legrand E, Flipo RM, Guggenbuhl P, Masson C, Maillefert JF, Soubrier M, et al. Management of nontuberculous infectious discitis. Treatments used in 110

References
patients admitted to 12 teaching hospitals in France. Joint Bone Spine. 2001 Dec;68(6):504-9.
167. Kumke M, Watschke H, Vietor T. A new methodological framework for design for additive manufacturing. Virtual and Physical Prototyping. 2016 2016;11(1):3-19.
168. Chesson LA, Howa JD, Lott MJ, Ehleringer JR. Development of a methodological framework for applying isotope ratio mass spectrometry to explosive components. Forensic Chemistry. 2016 Nov;2:9-14.
169. Pham MT, Rajic A, Greig JD, Sargeant JM, Papadopoulos A, McEwen SA. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Research Synthesis Methods. 2014 Dec;5(4):371-85.
170. Zamawe FC. The Implication of Using NVivo Software in Qualitative Data Analysis: Evidence-Based Reflections. Malawi Medical Journal. 2015 Mar;27(1):13-5.
171. Jamieson S. Analyse qualitative data. Education for Primary Care. 2016 2016;27(5):398-402.
172. Gentile JH, Harwell MA, Cropper W, Harwell CC, DeAngelis D, Davis S, et al. Ecological conceptual models: a framework and case study on ecosystem management for South Florida sustainability. Science of the Total Environment. 2001 Jul;274(1-3):231-53.
173. Barra Montevechi JA, Friend JD, editors. Using A Soft Systems Methodology Framework To Guide The Conceptual Modeling Process In Discrete Event Simulation. Winter Simulation Conference (WSC); 2012 2012 Dec 09-12; Berlin, GERMANY; 2012.
174. van der Zee D-J. An Integrated Conceptual Modeling Framework For Simulation Linking Simulation - Modeling To The Systems Engineering Process. In: Laroque C, Himmelspach J, Pasupathy R, editors. 2012 Winter Simulation Conference2012.
175. Tako AA, Kotiadis K, Vasilakis C. A Participative Modelling Framework For Developing Conceptual Models In Healthcare Simulation Studies. In: Johansson B, Jain S, MontoyaTorres J, Hugan J, Yucesan E, editors. Proceedings of the 2010 Winter Simulation Conference2010. p. 500-12.
176. Pace DK. Ideas about simulation conceptual model development. Johns Hopkins Apl Technical Digest. 2000 Jul-Sep;21(3):327-36.
177. Pereira TF, Barra Montevechi JA, Miranda RdC, Friend JD. Integrating soft systems methodology to aid simulation conceptual modeling. International Transactions in Operational Research. 2015 Mar;22(2):265-85.
178. Heritage NSWOoEa. Building a conceptual model | NSW Environment & Heritage. 2016; Available from: http://www.environment.nsw.gov.au/4cmas/buildconceptmodel.htm.
179. Pereira TF, Montevechi JAB, Miranda RD, Friend JD. Integrating soft systems methodology to aid simulation conceptual modeling. International Transactions in Operational Research. 2015 Mar;22(2):265-85.
180. Pace D, editor. Conceptual model development for C4ISR simulations. 5th International Command and Control Research and Technology Symposium, October 24-26 2000; 2000.
181. Winston R, editor. Managing the Development of Large Software Systems. Proceedings of IEEE WESCON; 1970.
182. Shah A, Baral NR, Manandhar A. Chapter Four - Technoeconomic Analysis and Life Cycle Assessment of Bioenergy Systems. In: Li Y, Ge X, editors. Advances in Bioenergy2016. p. 189 - 247.

References

183. Mills T, Lawton R, Sheard L. Advancing complexity science in healthcare research: the logic of logic models. Bmc Medical Research Methodology. 2019 Mar 12;19.
184. Langley T, Gillespie D, Lewis S, Eminson K, Brennan A, Docherty G, et al. Developing logic models to inform public health policy outcome evaluation: an example from tobacco control. Journal of public health (Oxford, England). 2020 2020-Mar-05.
185. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. Bmj-British Medical Journal. 2015 Mar 19;350.
186. Hoogendoorn M, Rutten-van Molken M, Hoogenveen RT, van Genugten MLL, Buist AS, Wouters EFM, et al. A dynamic population model of disease progression in COPD. European Respiratory Journal. 2005 Aug;26(2):223-33.
187. Borg S, Ericsson A, Wedzicha J, Gulsvik A, Lundback B, Donaldson GC, et al. A computer simulation model of the natural history and economic impact of chronic obstructive pulmonary disease. Value in Health. 2004 Mar-Apr;7(2):153-67.
188. Celli BR, Cote CG, Marin JM, Casanova C, de Oca MM, Mendez RA, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. New England Journal of Medicine. 2004 Mar 4;350(10):1005-12.
189. Patel AR, Patel AR, Singh S, Singh S, Khawaja I. Global Initiative for Chronic Obstructive Lung Disease: The Changes Made. Cureus. 2019 Jun 24;11(6).
190. Britton M. The burden of COPD in the UK: results from the Confronting COPD survey. Respiratory Medicine. 2003 Mar;97:S71-S9.
191. Pickard AS, Wilke C, Jung E, Patel S, Stavem K, Lee TA. Use of a preference-based measure of health (EQ-5D) in COPD and asthma. Respiratory Medicine. 2008 Apr;102(4):519-36.
192. Nolan CM, Longworth L, Lord J, Canavan JL, Jones SE, Kon SSC, et al. The EQ-5D-5L health status questionnaire in COPD: validity, responsiveness and minimum important difference. Thorax. 2016 Jun;71(6):493-500.
193. Darcet F, Gardier AM, Gaillard R, David DJ, Guilloux J-P. Cognitive Dysfunction in Major Depressive Disorder. A Translational Review in Animal Models of the Disease. Pharmaceuticals. 2016 Mar;9(1).
194. Preventing recurrent depression: long-term treatment for major depressive disorder. Primary care companion to the Journal of clinical psychiatry. 2007 2007;9(3):214-23.
195. Organisation. WH. Depression. 2020; Available from: https://www.who.int/news-room/fact-sheets/detail/depression.
196. Kennedy SH. Core symptoms of major depressive disorder: relevance to diagnosis and treatment. Dialogues in clinical neuroscience. 2008 2008;10(3):271-7.
197. Bachmann S. Epidemiology of Suicide and the Psychiatric Perspective. International Journal of Environmental Research and Public Health. 2018 Jul;15(7).
198. Tu Y-K. Commentary: Is structural equation modelling a step forward for epidemiologists. International Journal of Epidemiology. 2009 Apr;38(2):549-51.
199. Kozak M, Azevedo RA. Does using stepwise variable selection to build sequential path analysis models make sense? Physiologia Plantarum. 2011 Mar;141(3):197-200.
200. Halsey LG. The reign of the p-value is over: what alternative analyses could we employ to fill the power vacuum? Biology Letters. 2019 May 1;15(5).

References

201. Allotey J, Snell KIE, Smuk M, Hooper R, Chan CL, Ahmed A, et al. Validation and development of models using clinical, biochemical and ultrasound markers for predicting pre-eclampsia: an individual participant data meta-analysis. 2020;24:72.
202. Brealey S, Northgraves M, Kottam L, Keding A, Corbacho B, Goodchild L, et al. Surgical treatments compared with early structured physiotherapy in secondary care for adults with primary frozen shoulder: the UK FROST three-arm RCT. 2020;24:71.
203. Hagen S, Bugge C, Dean SG, Elders A, Hay-Smith J, Kilonzo M, et al. Basic versus biofeedback-mediated intensive pelvic floor muscle training for women with urinary incontinence: the OPAL RCT. 2020;24:70.
204. Shalhoub J, Lawton R, Hudson J, Baker C, Bradbury A, Dhillon K, et al. Compression stockings in addition to low-molecular-weight heparin to prevent venous thromboembolism in surgical inpatients requiring pharmacoprophylaxis: the GAPS non-inferiority RCT. 2020;24:69.
205. McRobbie HJ, Phillips-Waller A, El Zerbi C, McNeill A, Hajek P, Pesola F, et al. Nicotine replacement treatment, e-cigarettes and an online behavioural intervention to reduce relapse in recent ex-smokers: a multinational four-arm RCT. 2020;24:68.
206. Simpson CR, Lone NI, Kavanagh K, Englishby T, Robertson C, McMenamin J, et al. Vaccine effectiveness of live attenuated and trivalent inactivated influenza vaccination in 2010/11 to 2015/16: the SIVE II record linkage study. 2020;24:67.
207. Medina-Lara A, Grigore B, Lewis R, Peters J, Price S, Landa P, et al. Cancer diagnostic tools to aid decision-making in primary care: mixed-methods systematic reviews and cost-effectiveness analysis. 2020;24:66.
208. Barker KL, Room J, Knight R, Dutton SJ, Toye F, Leal J, et al. Outpatient physiotherapy versus home-based rehabilitation for patients at risk of poor outcomes after knee arthroplasty: CORKA RCT. 2020;24:65.
209. Batchelor JM, Thomas KS, Akram P, Azad J, Bewley A, Chalmers JR, et al. Home-based narrowband UVB, topical corticosteroid or combination for children and adults with vitiligo: HI-Light Vitiligo three-arm RCT. 2020;24:64.
210. Taylor AH, Taylor RS, Ingram WM, Anokye N, Dean S, Jolly K, et al. Adding web-based behavioural support to exercise referral schemes for inactive adults with chronic health conditions: the e-coachER RCT. 2020;24:63.
211. Crawford F, Chappell FM, Lewsey J, Riley R, Hawkins N, Nicolson D, et al. Risk assessments and structured care interventions for prevention of foot ulceration in diabetes: development and validation of a prognostic model. 2020;24:62.
212. Pickard R, Goulao B, Carnell S, Shen J, MacLennan G, Norrie J, et al. Open urethroplasty versus endoscopic urethrotomy for recurrent urethral stricture in men: the OPEN RCT. 2020;24:61.
213. Clarkson JE, Pitts NB, Goulao B, Boyers D, Ramsay CR, Floate R, et al. Risk-based, 6-monthly and 24-monthly dental check-ups for adults: the INTERVAL three-arm RCT. 2020;24:60.
214. Abel KM, Bee P, Gega L, Gellatly J, Kolade A, Hunter D, et al. An intervention to improve the quality of life in children of parents with serious mental illness: the Young SMILES feasibility RCT. 2020;24:59.
215. Appleton RE, Rainford NEA, Gamble C, Messahel S, Humphreys A, Hickey H, et al. Levetiracetam as an alternative to phenytoin for second-line emergency treatment of children with convulsive status epilepticus: the EcLiPSE RCT. 2020;24:58.

References

216. Gilbert R, Brown M, Faria R, Fraser C, Donohue C, Rainford N, et al. Antimicrobial-impregnated central venous catheters for preventing neonatal bloodstream infection: the PREVAIL RCT. 2020;24:57.
217. Stephenson J, Bailey JV, Blandford A, Brima N, Copas A, D'Souza P, et al. An interactive website to aid young women's choice of contraception: feasibility and efficacy RCT. 2020;24:56.
218. Duffy S, Vulkan D, Cuckle H, Parmar D, Sheikh S, Smith R, et al. Annual mammographic screening to reduce breast cancer mortality in women from age 40 years: long-term follow-up of the UK Age RCT. 2020;24:55.
219. Rodgers H, Bosomworth H, Krebs HI, van Wijck F, Howel D, Wilson N, et al. Robot-assisted training compared with an enhanced upper limb therapy programme and with usual care for upper limb functional limitation after stroke: the RATULS three-group RCT. 2020;24:54.
220. Van den Bruel A, Verbakel J, Wang K, Fleming S, Holtman G, Glogowska M, et al. Non-contact infrared thermometers compared with current approaches in primary care for children aged 5 years and under: a method comparison study. 2020;24:53.
221. Dias J, Brealey S, Cook L, Fairhurst C, Hinde S, Leighton P, et al. Surgical fixation compared with cast immobilisation for adults with a bicortical fracture of the scaphoid waist: the SWIFFT RCT. 2020;24:52.
222. Armstrong N, Büyükkaramikli N, Penton H, Riemsma R, Wetzelaer P, Huertas Carrera V, et al. Avatrombopag and lusutrombopag for thrombocytopenia in people with chronic liver disease needing an elective procedure: a systematic review and cost-effectiveness analysis. 2020;24:51.
223. Bray N, Kolehmainen N, McAnuff J, Tanner L, Tuersley L, Beyer F, et al. Powered mobility interventions for very young children with mobility limitations to aid participation and positive development: the EMPoWER evidence synthesis. 2020;24:50.
224. Foster NE, Konstantinou K, Lewis M, Ogollah R, Saunders B, Kigozi J, et al. Stratified versus usual care for the management of primary care patients with sciatica: the SCOPiC RCT. 2020;24:49.
225. Walton M, Wade R, Claxton L, Sharif-Hurst S, Harden M, Patel J, et al. Selective internal radiation therapies for unresectable early-, intermediate- or advanced-stage hepatocellular carcinoma: systematic review, network meta-analysis and economic evaluation. 2020;24:48.
226. Gilson R, Nugent D, Bennett K, Doré CJ, Murray ML, Meadows J, et al. Imiquimod versus podophyllotoxin, with and without human papillomavirus vaccine, for anogenital warts: the HIPvac factorial RCT. 2020;24:47.
227. Leaviss J, Davis S, Ren S, Hamilton J, Scope A, Booth A, et al. Behavioural modification interventions for medically unexplained symptoms in primary care: systematic reviews and economic evaluation. 2020;24:46.
228. Baker P, Coole C, Drummond A, Khan S, McDaid C, Hewitt C, et al. Occupational advice to help people return to work following lower limb arthroplasty: the OPAL intervention mapping study. 2020;24:45.
229. Crawford MJ, Thana L, Evans R, Carne A, O'Connell L, Claringbold A, et al. Switching antipsychotic medication to reduce sexual dysfunction in people with psychosis: the REMEDY RCT. 2020;24:44.
230. Melton H, Meader N, Dale H, Wright K, Jones-Diette J, Temple M, et al. Interventions for adults with a history of complex traumatic events: the INCiTE mixed-methods systematic review. 2020;24:43.

References

231. Lewis AL, Young GJ, Selman LE, Rice C, Clement C, Ochieng CA, et al. Urodynamics tests for the diagnosis and management of bladder outlet obstruction in men: the UPSTREAM non-inferiority RCT. 2020;24:42.
232. Worthington J, Lane JA, Taylor H, Young G, Noble SM, Abrams P, et al. Thulium laser transurethral vaporesection versus transurethral resection of the prostate for benign prostatic obstruction: the UNBLOCS RCT. 2020;24:41.
233. Earl H, Hiller L, Vallier A-L, Loi S, McAdam K, Hughes-Davies L, et al. Six versus 12 months' adjuvant trastuzumab in patients with HER2-positive early breast cancer: the PERSEPHONE non-inferiority RCT. 2020;24:40.
234. Corbett M, Duarte A, Llewellyn A, Altunkaya J, Harden M, Harris M, et al. Point-of-care creatinine tests to assess kidney function for outpatients requiring contrast-enhanced CT imaging: systematic reviews and economic evaluation. 2020;24:39.
235. Costa ML, Achten J, Knight R, Png ME, Bruce J, Dutton S, et al. Negative-pressure wound therapy compared with standard dressings following surgical treatment of major trauma to the lower limb: the WHiST RCT. 2020;24:38.
236. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, et al. Active monitoring, radical prostatectomy and radical radiotherapy in PSA-detected clinically localised prostate cancer: the ProtecT three-arm RCT. 2020;24:37.
237. Jones AP, Clayton D, Nkhoma G, Sherratt FC, Peak M, Stones SR, et al. Different corticosteroid induction regimens in children and young people with juvenile idiopathic arthritis: the SIRJIA mixed-methods feasibility study. 2020;24:36.
238. Rodgers M, Marshall D, Simmonds M, Le Couteur A, Biswas M, Wright K, et al. Interventions based on early intensive applied behaviour analysis for autistic children: a systematic review and cost-effectiveness analysis. 2020;24:35.
239. Hernández Alava M, Wailoo A, Pudney S, Gray L, Manca A. Mapping clinical outcomes to generic preference-based outcome measures: development and comparison of methods. 2020;24:34.
240. Coomarasamy A, Harb HM, Devall AJ, Cheed V, Roberts TE, Goranitis I, et al. Progesterone to prevent miscarriage in women with early pregnancy bleeding: the PRISM RCT. 2020;24:33.
241. Lobban F, Akers N, Appelbe D, Iraci Capuccinello R, Chapman L, Collinge L, et al. A web-based, peer-supported self-management intervention to reduce distress in relatives of people with psychosis or bipolar disorder: the REACT RCT. 2020;24:32.
242. Fraser H, Gallacher D, Achana F, Court R, Taylor-Phillips S, Nduka C, et al. Rapid antigen detection and molecular tests for group A streptococcal infections for acute sore throat: systematic reviews and economic evaluation. 2020;24:31.
243. Hounsome J, Pilkington G, Mahon J, Boland A, Beale S, Kotas E, et al. Prophylactic removal of impacted mandibular third molars: a systematic review and economic evaluation. 2020;24:30.
244. Davis S, Simpson E, Hamilton J, James MM-S, Rawdin A, Wong R, et al. Denosumab, raloxifene, romosozumab and teriparatide to prevent osteoporotic fragility fractures: a systematic review and economic evaluation. 2020;24:29.
245. Winkley K, Upsher R, Stahl D, Pollard D, Kasera A, Brennan A, et al. Psychological interventions to improve self-management of type 1 and type 2 diabetes: a systematic review. 2020;24:28.

References

246. Witham MD, Band M, Chong H, Donnan PT, Hampson G, Hu MK, et al. Sodium bicarbonate to improve physical function in patients over 60 years with advanced chronic kidney disease: the BiCARB RCT. 2020;24:27.
247. Bhatnagar R, Luengo-Fernandez R, Kahan BC, Rahman NM, Miller RF, Maskell NA. Thoracoscopy and talc poudrage compared with intercostal drainage and talc slurry infusion to manage malignant pleural effusion: the TAPPS RCT. 2020;24:26.
248. Poolman M, Roberts J, Wright S, Hendry A, Goulden N, Holmes EAF, et al. Carer administration of as-needed subcutaneous medication for breakthrough symptoms in people dying at home: the CARiAD feasibility RCT. 2020;24:25.
249. Shaw L, Bhattarai N, Cant R, Drummond A, Ford GA, Forster A, et al. An extended stroke rehabilitation service for people who have had a stroke: the EXTRAS RCT. 2020;24:24.
250. Tume LN, Woolfall K, Arch B, Roper L, Deja E, Jones AP, et al. Routine gastric residual volume measurement to guide enteral feeding in mechanically ventilated infants and children: the GASTRIC feasibility study. 2020;24:23.
251. Dennis M, Forbes J, Graham C, Hackett M, Hankey GJ, House A, et al. Fluoxetine to improve functional outcomes in patients after acute stroke: the FOCUS RCT. 2020;24:22.
252. Davidson B, Gurusamy K, Corrigan N, Croft J, Ruddock S, Pullan A, et al. Liver resection surgery compared with thermal ablation in high surgical risk patients with colorectal liver metastases: the LAVA international RCT. 2020;24:21.
253. Beard DJ, Davies LJ, Cook JA, MacLennan G, Price A, Kent S, et al. Total versus partial knee replacement in patients with medial compartment knee osteoarthritis: the TOPKAT RCT. 2020;24:20.
254. Palmer R, Dimairo M, Latimer N, Cross E, Brady M, Enderby P, et al. Computerised speech and language therapy or attention control added to usual care for people with long-term post-stroke aphasia: the Big CACTUS three-arm RCT. 2020;24:19.
255. Dorling J, Hewer O, Hurd M, Bari V, Bosiak B, Bowler U, et al. Two speeds of increasing milk feeds for very preterm or very low-birthweight infants: the SIFT RCT. 2020;24:18.
256. Mallucci CL, Jenkinson MD, Conroy EJ, Hartley JC, Brown M, Moitt T, et al. Silver-impregnated, antibiotic-impregnated or non-impregnated ventriculoperitoneal shunts to prevent shunt infection: the BASICS three-arm RCT. 2020;24:17.
257. Surr CA, Holloway I, Walwyn REA, Griffiths AW, Meads D, Kelley R, et al. Dementia Care Mapping™ to reduce agitation in care home residents with dementia: the EPIC cluster RCT. 2020;24:16.
258. Francis NA, Gillespie D, White P, Bates J, Lowe R, Sewell B, et al. C-reactive protein point-of-care testing for safely reducing antibiotics for acute exacerbations of chronic obstructive pulmonary disease: the PACE RCT. 2020;24:15.
259. Day C, Briskman J, Crawford MJ, Foote L, Harris L, Boadu J, et al. An intervention for parents with severe personality difficulties whose children have mental health problems: a feasibility RCT. 2020;24:14.
260. Snowden C, Lynch E, Avery L, Haighton C, Howel D, Mamasoula V, et al. Preoperative behavioural intervention to reduce drinking before elective orthopaedic surgery: the PRE-OP BIRDS feasibility RCT. 2020;24:12.

References

261. Stevenson M, Uttley L, Oakley JE, Carroll C, Chick SE, Wong R. Interventions to reduce the risk of surgically transmitted Creutzfeldt–Jakob disease: a cost-effective modelling review. 2020;24:11.
262. Rake C, Gilham C, Bukasa L, Ostler R, Newton M, Peto Wild J, et al. High-dose oral vitamin D supplementation and mortality in people aged 65–84 years: the VIDAL cluster feasibility RCT of open versus double-blind individual randomisation. 2020;24:10.
263. Davies NM, Taylor AE, Taylor GMJ, Itani T, Jones T, Martin RM, et al. Varenicline versus nicotine replacement therapy for long-term smoking cessation: an observational study using the Clinical Practice Research Datalink. 2020;24:9.
264. Costa ML, Achten J, Wagland S, Marian IR, Maredza M, Schlüssel MM, et al. Plaster cast versus functional bracing for Achilles tendon rupture: the UKSTAR RCT. 2020;24:8.
265. Riley P, Glenny A-M, Worthington HV, Jacobsen E, Robertson C, Durham J, et al. Oral splints for patients with temporomandibular disorders or bruxism: a systematic review and economic evaluation. 2020;24:7.
266. Froggatt K, Best A, Bunn F, Burnside G, Coast J, Dunleavy L, et al. A group intervention to improve quality of life for people with advanced dementia living in care homes: the Namaste feasibility cluster RCT. 2020;24:6.
267. Edwards SJ, Wakefield V, Jhita T, Kew K, Cain P, Marceniuk G. Implantable cardiac monitors to detect atrial fibrillation after cryptogenic stroke: a systematic review and economic evaluation. 2020;24:5.
268. Lincoln NB, Bradshaw LE, Constantinescu CS, Day F, Drummond AER, Fitzsimmons D, et al. Group cognitive rehabilitation to reduce the psychological impact of multiple sclerosis on quality of life: the CRAMMS RCT. 2020;24:4.
269. Duarte R, Stainthorpe A, Greenhalgh J, Richardson M, Nevitt S, Mahon J, et al. Lead-I ECG for detecting atrial fibrillation in patients with an irregular pulse using single time point testing: a systematic review and economic evaluation. 2020;24:3.
270. Fleeman N, Houten R, Bagust A, Richardson M, Beale S, Boland A, et al. Lenvatinib and sorafenib for differentiated thyroid cancer after radioactive iodine: a systematic review and economic evaluation. 2020;24:2.
271. Maguire A, Clarkson JE, Douglas GVA, Ryan V, Homer T, Marshman Z, et al. Best-practice prevention alone or with conventional or biological caries management for 3- to 7-year-olds: the FiCTION three-arm RCT. 2020;24:1.
272. Brondizio ES, Vogt ND, Mansur AV, Anthony EJ, Costa S, Hetrick S. A conceptual framework for analyzing deltas as coupled social-ecological systems: an example from the Amazon River Delta. Sustainability Science. 2016 Jul;11(4):591-609.
273. Ianni M, de Leon MS. Applying Energy Performance-Based Design in Early Design Stages A methodological framework for integrating multiple BPS tools. Ecaade 2013: Computation and Performance, Vol 1. 2013 2013:31-40.
274. Kumar A, Singh AR, Deng Y, He X, Kumar P, Bansal RC. A Novel Methodological Framework for the Design of Sustainable Rural Microgrid for Developing Nations. Ieee Access. 2018 2018;6:24925-51.
275. Panagiotopoulou M, Stratigea A. A participatory methodological framework for paving alternative local tourist development paths-the case of Sterea Ellada Region. European Journal of Futures Research. 2014 Dec;2(1).
276. Rijke J, Brown R, Zevenbergen C, Ashley R, Farrelly M, Morison P, et al. Fit-for-purpose governance: A framework to make adaptive governance operational. Environmental Science & Policy. 2012 Oct;22:73-84.

References

277. Schmitt J, Apfelbacher C, Spuls PI, Thomas KS, Simpson EL, Furue M, et al. The Harmonizing Outcome Measures for Eczema (HOME) Roadmap: A Methodological Framework to Develop Core Sets of Outcome Measurements in Dermatology. Journal of Investigative Dermatology. 2015 Jan;135(1):24-30.