



Simulation optimisation to inform economic evaluations of sequential therapies for chronic conditions: a case study in Rheumatoid Arthritis

Jonathan Cameron Tosh

BA (Econ), MSc.

Supervised by Professor Matt Stevenson, Professor Ron Akehurst and Dr Mark Strong

A thesis in partial fulfilment of the requirements for the degree of Doctor of Philosophy

School of Health and Related Research
Faculty of Medicine, Dentistry and Health
The University of Sheffield

Submitted September 2015

This work has been prepared as part of a Doctoral Research Fellowship Award funded by the
National Institute for Health Research

ACKNOWLEDGEMENTS

Thank you to my supervisors; Professor Matt Stevenson, Professor Ron Akehurst and Dr Mark Strong, for their support and encouragement. Thank you to my colleagues at SchARR, especially my line manager Allan Wailoo. Thanks to everyone working on the NICE RA Biologics Appraisal, and whose work contributed to Chapter 5. Thanks to Nick Latimer, Paul Tappenden, Jill Carlton and Brendan Mulhern for being available when I needed advice, a rant, or a distraction.

I would like to acknowledge the National Taught Course Centre in Operational Research (NATCOR). Thank you to the organisers of the residential courses, and thanks to the tutors and fellow students for putting up with an economist asking daft questions.

To my parents, Cheryl and Grahame, I am forever grateful and always inspired.

Finally, I must thank my wife Lauren and my son Rowan. Thank you so much for your love, support and tolerance. This PhD would not have happened without you.

DISSEMINATION

I have published parts of this research in the following peer reviewed journal article:

- Tosh J, Stevenson M, Akehurst RL. Health Economic Modelling of Treatment Sequences for Rheumatoid Arthritis: A Systematic Review. *Current Rheumatology Reports* (2014). 16:447

I have presented parts of this research at the following international conferences:

- Tosh J, Stevenson M, Akehurst RL. Health Economic Modelling of Treatment Sequences for Rheumatoid Arthritis: A Systematic Review. ISPOR 16th Annual European Congress. Dublin, Ireland. November 2013
- Tosh J, Stevenson M, Strong M, Akehurst RL. A Framework for the Economic Evaluation of Sequential Therapies for Chronic Conditions. ISPOR 17th Annual European Congress. Amsterdam, The Netherlands. November 2014

FUNDING

This PhD thesis is independent research arising from a Doctoral Research Fellowship supported by the National Institute for Health Research. The views expressed in this publication are those of the author and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

DEDICATION

For Lauren and Rowan

TABLE OF CONTENTS

List of tables, figures and boxes.....	7
List of abbreviations.....	11
Abstract.....	13
Chapter 1: Introduction	1
1.1 Motivation.....	1
1.2 Overview of thesis.....	2
1.3 Rheumatoid arthritis.....	3
Chapter 2: The economic framework for Health Technology Assessment (HTA)	5
2.1 Introduction	5
2.2 Objectives of a universal health system	6
2.3 Welfare economics and health care allocation	7
2.4 Economic evaluation to inform health care allocation.....	9
2.5 National level NHS health resource allocation	17
2.6 Operational Research.....	19
2.7 Conclusions	21
Chapter 3: Rationale for this thesis.....	23
3.1 Introduction	23
3.2 Treatment sequences	23
3.3 Treatment sequences and economic evaluation.....	29
3.4 Pharmaceutical markets and treatment sequences.....	36
3.5 Patient choice and treatment sequences.....	39
3.6 Conclusions	40
3.7 Aims and objectives	41
3.8 Structure of the thesis	41
Chapter 4: A systematic review of economic evaluations of Disease Modifying Anti-Rheumatic Drugs (DMARDs) for rheumatoid arthritis.....	43
4.1 Chapter overview.....	43
4.2 Introduction	43
4.3 Methods.....	45
4.4 Results.....	50
4.5 Discussion.....	71
4.6 Conclusions	73
Chapter 5: Developing a health economic model for the NICE technology appraisal of biologics for Rheumatoid Arthritis.....	75
5.1 Chapter overview.....	75
5.2 Introduction	76
5.3 Decision Problem	78

5.4 Model conceptualisation.....	81
5.5 Model development	88
5.6 Model results and discussion	101
5.7 Discussion	106
5.8 Conclusions.....	107
Chapter 6: A citation pearl growing systematic review of the methodological literature.....	109
6.1 Chapter overview	109
6.2 Introduction.....	109
6.3 Methods	111
6.4 Search strategy	114
6.5 Quality assessment.....	116
6.6 Data extraction and synthesis	116
6.7 Search results	117
6.8 Narrative synthesis of identified methods	118
6.9 Discussion and conclusions	147
Chapter 7: Application of simulation optimisation via simulated annealing for an economic evaluation of treatment sequences for rheumatoid arthritis	151
7.1 Chapter overview	151
7.2 Introduction.....	151
7.3 Methods	152
7.4 Results	204
7.5 Discussion and conclusions	219
Chapter 8: Discussions, recommendations and conclusions	223
8.1 Chapter overview	223
8.2 Contribution of this work in the context of other research	223
8.3 Strengths and limitations of this research.....	225
8.4 Recommendations for further research.....	229
8.5 Conclusions.....	231
References	233
Appendices	253
Appendix A: RA economic evaluation systematic review	253
Appendix B: Methodological review	274
Appendix C: Simul8 model.....	321
Appendix D: Simulation optimisation model.....	342

LIST OF TABLES, FIGURES AND BOXES

Table 2.1: Kleczkowski model of a health care system	6
Table 2.2: Results from a cost-effectiveness analysis (CEA)	12
Table 2.3: Types of NICE guidance	17
Table 3.1: Number of potential sequences given the perspective	35
Table 4.1: Keywords for systematic review (PICO format)	45
Table 4.2: Systematic review databases	46
Table 4.3: Eligibility criteria.....	47
Table 4.4: Data extraction template	49
Table 4.5: Summary of reviewed studies (Table also published in Tosh <i>et al.</i> (2014) ⁹⁴)	52
Table 4.6: Review patient population	58
Table 4.7: Treatments evaluated for recent-onset RA	59
Table 4.8: Treatments evaluated for established RA.....	65
Table 5.1: Disease modifying anti-rheumatic drugs (DMARDs).....	77
Table 5.2: Technology Appraisals being updated by NICE RA biologics appraisal	79
Table 5.3: Licenses for RA bDMARDs	81
Table 5.4: Populations and analyses	82
Table 5.5: Sequences (patients who could receive MTX)	84
Table 5.6: Sequences (patients who could not receive MTX).....	84
Table 5.7: Sequences evaluated for Populations 2 and 3 for those who can receive MTX.....	85
Table 5.8: Sequences evaluated for Populations 2 and 3 for those who cannot receive MTX ..	85
Table 5.9: Sequences evaluated for Population 1 for those who can receive MTX	85
Table 5.10: Sequences evaluated for Population 1 for those who cannot receive MTX.....	85
Table 5.11: Discrete vs Markovian patient level simulation.....	88
Table 5.12: Population parameters	90
Table 5.13: Simulation event list.....	91
Table 5.14: EULAR response criteria	93
Table 5.15: Mean HAQ improvement by EULAR response - BSRBR dataset	94
Table 5.16: Drug costs for all included treatments.....	98
Table 5.17: Monitoring costs	99
Table 5.18: HAQ related costs.....	99
Table 5.19: Summarised results: Median ICERS for all bDMARD strategies compared with MTX alone strategy. Populations 2 & 3 who can receive MTX	103

Table 5.20: Summarised results: Median ICERs for all bDMARD strategies compared with SSZ alone strategy. Populations 2 & 3 who cannot receive MTX	104
Table 5.21: Summarised results: Median ICERs for all bDMARD strategies compared with MTX alone strategy. Population 1 who can receive MTX	105
Table 5.22: Summarised results: Median ICERs for all bDMARD strategies compared with SSZ alone strategy. Population 1 who cannot receive MTX	105
Table 6.1: Framework for the methods review	113
Table 6.2: Search terms	115
Table 6.3: Inclusion and Exclusion criteria	115
Table 6.4: Data extraction form	116
Table 6.5: Methodological studies and categories.....	123
Table 6.6: Jacobson et al. algorithm formulations ²¹⁴	125
Table 6.7: Neighbourhood size and move operators	130
Table 6.8: Metaheuristics - memory structure and algorithm type	132
Table 7.1: TA198 Treatment decrement estimates.....	156
Table 7.2: NICE RA econometric models for patient level parameters.....	157
Table 7.3: Treatments included in analysis	161
Table 7.4: Maximum size of the decision space	162
Table 7.5: Validation tests	163
Table 7.6: No patient heterogeneity. No PSA (Deterministic)	169
Table 7.7: Patient heterogeneity. No PSA	169
Table 7.8: Patient Heterogeneity. PSA (Basecase model)	170
Table 7.9: Comparing certolizumab pegol to etanercept.....	171
Table 7.10: Tuning the neighbourhood function.....	181
Table 7.11: van Laarhoven formulae estimates of initial temperature parameter	183
Table 7.12: Rapid warm-up algorithm for initial temperature parameter.....	183
Table 7.13: Tuning the initial temperature parameter	185
Table 7.14: Static cooling schedules for simulated annealing.....	186
Table 7.15: Lundy parameter values	188
Table 7.16: Lundy cooling schedule.....	190
Table 7.17: Linear cooling schedule	192
Table 7.18: Geometric cooling parameters	193
Table 7.19: Geometric cooling schedule	195
Table 7.20: Final tuning - random starting solution for each round	197
Table 7.21: Final tuning - current best solution for each round	198
Table 7.22: Final test of Lundy cooling schedule.....	199

Table 7.23: Tuning the maximum number of consecutive failed attempts.....	200
Table 7.24: Maximum consecutive failed attempts (50 - 30)	201
Table 7.25: Maximum consecutive failed attempts (20 and 15)	202
Table 7.26: Maximum consecutive failed attempts (10 and 5)	203
Table 7.27: Basecase results	204
Table 7.28: Basecase results (top 20 ranked)	206
Table 7.29: Scenario analysis - Lambda values £0, £20,000, £50,000	208
Table 7.30: Scenario analysis - Lambda values £100,000, £1,000,000	209
Table 7.31: Scenario analysis - treatment decrement values.....	211
Table 7.32: Scenario analysis - alternative cooling schedules	212
Table 7.33: Optimisation of cDMARD treatment sequence	214
Table 7.34: Comparison to NICE MTA appraisal of RA biologics	215
Table 7.35: NICE guidance comparison.....	218
Figure 3.1: HRQL profile.....	30
Figure 3.2: QALY profile - relapsing/remitting chronic condition.....	30
Figure 3.3: QALY profiles for comparator sequences	31
Figure 3.4: Three treatment decision tree model.....	32
Figure 3.5: Three treatment sequential decision tree model.....	33
Figure 4.1: Quorum flow diagram.....	50
Figure 4.2: NICE Clinical Guideline (CG90) - Patient population and treatment algorithm (simplified)	56
Figure 5.1: Summary of biologics within NICE TA guidance	80
Figure 5.2: Simulation logic for RA model.....	92
Figure 5.3: Mean EULAR response probabilities for comparator treatments (Population 2&3, main analysis - all ACR trials) - reproduced from SchARR Technology Appraisal Report.....	94
Figure 5.4: Mean HAQ by EULAR response for bDMARD patients	96
Figure 5.5: Discounted ICER of bDMARD sequence compared to non bDMARD sequence in Population 1 (from Assessment Group report)	101
Figure 6.1: Classical search model (Levy 2012).....	112
Figure 6.2: QUORUM diagram for methods review.....	118
Figure 6.3: Local and global optima	120
Figure 7.1: Simulation Optimisation algorithm.....	154
Figure 7.2: HRUV meta model regression residuals	158
Figure 7.3: HRUV meta model regression fit	159

Figure 7.4: Time-on-treatment (Moderate EULAR response)	160
Figure 7.5: Time-on-treatment (Good EULAR response)	160
Figure 7.6: No patient heterogeneity. No PSA (Deterministic)	166
Figure 7.7: Patient heterogeneity. No PSA	167
Figure 7.8: Patient heterogeneity. PSA (Basecase model)	168
Figure 7.9: COM Interface	172
Figure 7.10: Temperature cooling steps and iterations	174
Figure 7.11: Lundy, Geometric and Linear Cooling Schedules (scaled for clarity)	187
Figure 7.12: Lundy cooling schedules.....	188
Figure 7.13: Linear cooling schedule	191
Figure 7.14: Geometric cooling schedule	193
Box 2.1: Cost Benefit Analysis	11
Box 2.2: Cost Effectiveness Analysis.....	12
Box 2.3: Process for undertaking a full incremental analysis.....	14
Box 2.4: Net Benefit.....	15
Box 2.5: Roles and applications of modelling within Health Technology Assessment ⁴²	16
Box 2.6: Linear Programming formulation of health care resource allocation ⁵⁶	20
Box 2.7: Integer Programming formulation of health care resource allocation ⁵⁹	21
Box 6.1: Metropolis algorithm.....	135
Box 6.2: Simulated Annealing acceptance criterion.....	136
Box 7.1: HRUV meta model regression results.....	158
Box 7.2: Metropolis acceptance criterion	174
Box 7.3: Simulated annealing algorithm	175
Box 7.4: Neighbourhood move operators.....	179

LIST OF ABBREVIATIONS

ABT	Abatacept
ABTS	Abatacept (subcutaneous)
ACR	American College of Rheumatology
ACR20/50/70	A clinical measure of rheumatoid arthritis treatment response. A 20/50/70% improvement in tender or swollen joint counts, as well as a 20/50/70% improvement in three of the five ACR response parameters (patient assessment, physician assessment, pain scale, functional questionnaire, acute phase reactant).
ADA	Adalimumab
ADA	Adalimumab
AM	Anti-malarial
ANA	Anakinra
AZA	Azathioprine
bDMARD	Biologic disease modifying anti rheumatic drug
BSC	Best supportive care
CBA	Cost-benefit analysis
cDMARD	Conventional disease modifying anti rheumatic drug
CEA	Cost-effectiveness analysis
CER	Cost-effectiveness ratio
CMA	Cost-minimisation analysis
CTZ	Certolizumab pegol
CUA	Cost-utility analysis
CYA	Cyclosporin A
DAM	Decision analytic modelling (A decision analytic model is simply referred to as a 'model')
DAS	Disease Activity Score
DES	Discrete event simulation
DMARD	Disease modifying anti rheumatic drug
ETN	Etanercept
EULAR	European league against rheumatism
GA	Genetic algorithm
GLD	Gold
GOL	Golimumab
HAQ	Health Assessment Questionnaire
HCQ	Hydroxychloroquine
HTA	Health technology assessment
HCQ	Hydroxychloroquine
HRQL	Health related quality of life
ICER	Incremental cost-effectiveness ratio
IFX	Infliximab
ILM	Individual level model
iNMB	Incremental net monetary benefit
IP	Integer programming

LEF	Leflunomide
LS	Local search
MTA	Multiple Technology Appraisal
MTX	Methotrexate
NMB	Net monetary benefit
NICE	National Institute of Health and Care Excellence
NP	Nested partitions
NPV	Net present value
OO	Ordinal optimisation
OR	Operational research
PC	Palliative care
PEN	D-Penicillamine
PNG	Pseudorandom number generator
PSA	Probabilistic sensitivity analysis
RA	Rheumatoid Arthritis
R&S	Ranking and selection
RS	Random search
RSM	Response surface methodology
RTX	Rituximab
SA	Simulated annealing
SO	Simulated optimisation
SOSA	Simulated optimisation via simulated annealing
SSZ	Sulfasalazine
TA	Technology Appraisal
TS	Tabu search
TCZ	Tocilizumab
TNF α	TNF- α inhibitor
UC	Usual care
WTP	Willingness to pay

ABSTRACT

This thesis investigates the problem of treatment sequencing within health economic evaluations. For some chronic conditions, sequences of treatments can be used. When there are a lot of alternative treatments, then the number of possible sequences becomes very large. When undertaking an economic evaluation, it may not be feasible to estimate the costs and benefits of every alternative treatment sequence. The objective of the thesis is to test the feasibility of simulation optimisation methods to find an optimal or set of near-optimal sequences of disease modifying treatments for rheumatoid arthritis in an economic evaluation framework.

A large number of economic evaluations have been undertaken to estimate the costs and benefits associated with different treatments for rheumatoid arthritis. Many of these have not considered the downstream sequence of treatments provided, and no published study has considered identifying the best, or optimal, treatment sequence. The published evidence is therefore of limited applicability if the objective is to maximise patient benefit while constrained by a finite budget. It is plausible that decision-makers have developed sub-optimal guidance for rheumatoid arthritis, and this could extend to other chronic conditions.

A simulation model can provide an expectation of the population mean costs and benefits for alternative treatment sequences. These models are routinely used to inform health economic evaluations. However, they can be computationally expensive to run, and therefore the evaluation of potentially millions of treatment sequences is not feasible. However, simulation optimisation methods exist to identify a good solution from a simulation model within a feasible period of time. Using these methods within an economic evaluation of treatment sequences has not previously been investigated.

In this thesis I highlight the importance of the treatment sequencing problem, review and assess relevant simulation optimisation methods, and implement a simulated annealing algorithm to explore its feasibility and appropriateness. From the implementation case study within rheumatoid arthritis, simulation optimisation via simulated annealing appears to be a feasible method to identify a set of good treatment sequences. However, the method requires a significant amount of time to implement and execute, which may limit its appropriateness for health resource allocation decision making. Further research is required to investigate the generalisability of the method, and further consideration regarding its use in a decision-making context is important.

(Blank page)

CHAPTER 1: INTRODUCTION

1.1 MOTIVATION

Economic evaluation is a framework to provide a formal quantification of the costs and benefits of alternative allocations of health care resources. The evidence provided by economic evaluations enables decision makers to make informed decisions regarding the allocation of scarce health resources. These decisions may involve the development of a hospital, the introduction of a screening programme, or the funding of a medical treatment.

The efficacy of many treatments can be uncertain, and even if they provide short term benefits, these may not persist. Therefore a switch may be made to an alternative treatment, and longer term and chronic conditions may be treated with a sequence of treatments. This sequence can provide disease control and symptomatic relief over a patient's lifetime. Therefore, for an economic evaluation to capture the full future costs and benefits, a comparison is required between alternative treatment sequences, rather than alternative individual treatments.

Evidence of the costs and benefits of competing alternatives or sequences are often unobserved, especially when accrued over a long time. Therefore decision analytic models are used to estimate expected future costs and benefits. These incorporate a range of evidence and assumptions, and for complex chronic conditions, a simulation model may be built to accurately reflect patient heterogeneity and capture future health events.

Simulation models may be more appropriate for chronic conditions, especially when patient heterogeneity has an impact on future costs and benefits. However, they are often computationally expensive.

In some chronic conditions, there can be a large number of treatments available. This is especially the case for rheumatoid arthritis (RA), where there are at least 13 unique treatments, and many treatments can be used in combination. There is a very large number of possible treatment sequences to be compared in an economic evaluation, so that the best sequence can be identified and the optimal use of health resources recommended. However, this large number of sequences, coupled with a simulation model which takes time to evaluate each sequence, means that it is not feasible to evaluate every possible sequence. This is the 'treatment sequencing problem'.

This problem can be framed as an optimisation problem, and simulation optimisation methods can be applied to determine a good enough solution within a feasible amount of time. However, applying a simulation optimisation method to address a treatment sequencing problem has not been conducted before.

The motivation for this thesis is to review simulation optimisation methods which can be applied to this treatment sequencing problem. The implementation of a method in this thesis for the RA treatment sequencing problem will enable a consideration about whether it is feasible for use in other treatment sequencing problems, and other large health economic evaluation problems more generally.

RESEARCH QUESTION

The primary research question for this PhD thesis is ‘How can economic evaluations of sequential therapies for chronic conditions improve health resource allocation decision making?’ The PhD uses RA as a case study.

The formal aims and objectives of the thesis are stated in Chapter 3, section 3.7.

1.2 OVERVIEW OF THESIS

Chapter 2 provides an introduction to economic evaluation. The theoretical basis to economic evaluation is explored and the decision analytic methods used to inform economic evaluations for decision makers are introduced. Chapter 3 provides a rationale for this thesis. It defines treatment sequences, considers a taxonomy for where treatment sequences may be used, and explains why treatment sequences represent a unique challenge when developing decision analytic models to inform an economic evaluation. Chapter 4 supports the rationale of this thesis by reporting a systematic review of economic evaluations of disease modifying anti-rheumatic therapies (DMARDs) for RA. It highlights that no previously conducted economic evaluation has attempted to estimate the optimal treatment sequence for RA.

In Chapter 5, a health economic model is developed to inform the currently ongoing NICE Technology Appraisal for biologic DMARDs. This allows a model to be utilised for the application of simulation-optimisation methods, which are reviewed in Chapter 6. Chapter 7 reports the application of simulation optimisation via simulated annealing (SOSA) for the RA treatment sequencing problem. The thesis ends with Chapter 8, a discussion about the strengths and limitations of the work undertaken, recommendations for policy makers and further research, and overall conclusions.

1.3 RHEUMATOID ARTHRITIS

EPIDEMIOLOGY

Rheumatoid arthritis (RA) is a chronic inflammatory disease. It is characterised by progressive and irreversible joint damage, as well as impaired joint function, pain, and tenderness. The condition leads to disability and reduced quality of life.¹ RA is associated with significant direct costs, as well as indirect costs due to reduced productivity.^{2,3} Evidence strongly suggests that patients with RA have a reduced life expectancy.^{4,5}

An estimated 400,000 people in England and Wales have RA,⁶ and it is more prevalent in females (1.16%) than males (0.44%).⁷ The majority of cases of RA are diagnosed when patients are between 40 and 70 years old.⁸

ASSESSMENT

In 1987, classification criteria for RA were produced by the American College of Rheumatology (ACR).⁹ In summary, for a diagnosis of RA, a patient must have at least four of the seven criteria: morning stiffness lasting for at least one hour; swelling in three or more joints; swelling in hand joints; symmetric joint swelling; x-ray imaging showing joint erosion or decalcification; rheumatoid nodules; and abnormal serum rheumatoid factor. The European League Against Rheumatism (EULAR) have also developed classification criteria, but these focus more on the identification of persistent synovitis, rather than satisfying the ACR criteria.¹⁰

The EULAR and ACR classification systems have led to the development of two measures of improvement in RA symptoms: ACR responses,¹¹ and EULAR responses.¹²

An ACR20 response requires: a 20% improvement in swollen joint counts; and a 20% improvement in at least three of the following five 'core set items': Physician global assessment; Patient global assessment; patient pain; self-reported disability (using a validated instrument); and Erythrocyte sedimentation rate / C-reactive protein. ACR50 and ACR70 are also routine measure of improvement, with 50% and 70% improvements required, rather than 20%. ACR response measures are routinely used in randomised controlled trials (RCTs).

In the UK and across Europe, the disease activity score of 28 joints (DAS28) is a routinely used measure of RA. The DAS28 can be used to classify both disease activity, and the level of improvement. The EULAR response criteria combine baseline DAS28 level, and the size of the DAS28 change, to classify response into 'good', 'moderate' and 'none'. The method for determining the EULAR response classification is provided later in the thesis, in Table 5.14.

A commonly used measure of patient functional capacity is the health assessment questionnaire (HAQ). The HAQ instrument has established reliability and validity and is routinely used in RCTs and observational registries. It is commonly used to provide a profile of functional worsening over time due to RA. HAQ scores range from 0 to 3, best to worst. The scale is discrete, with step values of 0.125, which results in 25 HAQ values.

TREATMENT

The traditional medical treatments for RA involve conventional disease-modifying anti-rheumatic drugs (cDMARDs), which include methotrexate (MTX), sulfasalazine (SSZ), hydroxychloroquine (HCQ), leflunomide (LEF) and gold injections. Alongside these disease modifying treatments, analgesics, steroids, and non-steroidal anti-inflammatories were (and still are) commonly prescribed.

However, more recently, a group of genetically-engineered biological therapies have been developed. Such drugs have been labelled as biologic DMARDs (bDMARDs). These treatments have amassed a significant amount of evidence to support their superiority over monotherapy cDMARDs, and their introduction to clinical practice has significantly improved the prognosis for patients with severe RA. However, the benefits provided by bDMARDs come with a significant price, and the introduction of bDMARDs has significantly increased spending on medical care for RA patients. The area of application within this thesis considers the cost-effectiveness of these bDMARDs, and their optimal use in clinical practice.

CHAPTER 2: THE ECONOMIC FRAMEWORK FOR HEALTH TECHNOLOGY ASSESSMENT (HTA)

2.1 INTRODUCTION

Economics is the study of how choices are made and how the resources of society are used.¹³ Resources are scarce, meaning there are not enough resources to satisfy the desires of all people. The scarcity of resources is the fundamental problem which economic theory and analysis look to address. Often, the definition of the social science of economics is posed as three related questions: “what is to be produced?”; “how is it to be produced?”; “who gets the output?”.¹⁴

Health itself is an economic good, which allows the analysis of its demand, production and consumption in similar ways to the analysis of other goods and services.¹⁵ However, health has particular characteristics, and is often viewed as a ‘basic pleasure’¹⁶ or a ‘fundamental commodity’.^{17,18} Health has a significant impact on people’s welfare or utility, both directly as a consumption good, but also indirectly as an investment good, because it provides more healthy days which can lead to additional earnings and utility.¹⁸ Because health itself cannot be purchased or traded, economic analysis has to focus on the production and allocation of health care resources, which have demand derived from the demand for health. People’s desire for health in turn sees desire for more tangible goods and services which are considered a means to create health. This heightens the need for analytical insights into the allocation of health resources, because of the significant impact they have on people’s welfare.

With scarcity of health care resources, an economic problem is observed in the field of health. The scarcity of resources must be considered alongside the objectives of a particular health system, such as maximising societal health. The production of health care is constrained by finite resources (factors of production) and technical possibilities, the two together represented by a production-possibility frontier. With a constrained supply of health care resources, decisions have to be made about how these scarce resources are allocated. The best allocation is that which satisfies the objective of the system, subject to the constrained supply of resources and the technologies available.

This chapter will look to discuss the economic theory and methods that have been used to inform the allocation of scarce health resources. The objective is not to provide a comprehensive discussion about every aspect of health economic theory, but to instead provide context and background for the aims and objectives of this thesis.

2.2 OBJECTIVES OF A UNIVERSAL HEALTH SYSTEM

A health care system is regularly defined as “the sum total of all the organisations, institutions and resources whose primary purpose is to improve health.”* This definition has been argued as being reductionist, by ignoring interrelations between components of the system, by removing the role of the population in a health care system, and by limiting its goal to just improving health.¹⁹ Kleczkowski (1984) provides a model of a health system, with five contributory components (Table 2.1).²⁰

Table 2.1: Kleczkowski model of a health care system

Component of a health care system	
1.	Development of health resources
2.	Organised arrangement of resources
3.	Delivery of health care
4.	Economic support
5.	Management

Kleczkowski’s model defines a health care system in terms of the production of health care resources, and also the delivery of health care and the overarching management of the system. The model also reveals the broad objective of a health care system; the efficient use of health resources at a population level. In economics, efficiency is concerned with both the production of goods and services (technical efficiency), as well as the distribution of goods and services (allocative efficiency). An allocatively efficient distribution of goods is where the social surplus is maximised, and therefore the marginal benefits equal the marginal costs.

The concept of allocative efficiency is aligned with the moral and political philosophy of distributive justice – what is just, or right, in allocating goods within a society.²¹ Two prominent moral theories of distributive justice are utilitarianism and egalitarianism, and they have a particular place in current health care resource allocation theory.²²

In utilitarianism, efficiency can broadly be defined as “greatest happiness of the greatest number”. Therefore the objective, to a utilitarian, is to maximise total social happiness, and a sacrifice from a minority that promotes the happiness of a majority is a worthy endeavour because the total social happiness has increased. Utilitarianism therefore makes no distributional judgement with regard to who ‘wins’ and who ‘loses’, and so equity of health

* http://www.who.int/topics/health_systems/ga/en/index.html - Accessed June 2015

care resource allocation would not be of concern. Egalitarianism is concerned with a distribution of resources so each member of society receives an equal share.²²

However, equity, along with efficiency, is an important policy objective in health care, as well as social policy more generally.²³ Equity incorporates a particular goal for which equality is desired. Equity in health care can be concerned with both the financing of health care, as well as the distribution of health care resources. With respect to the distribution of health care, equity can be considered in two dimensions. Vertical equity is the consideration that unequal people should be treated unequally, with horizontal equity the equal treatment of equal people. Defining how people are equal or unequal is specified by the equity goal. The quantity of literature related to equity in health care is enormous, however three equity goals are commonly discussed:²⁴

- Access
- Utilisation
- Outcomes

An objective of 'equal access for equal need' requires the conditions where the opportunities to access health care are equal for those with equal need – horizontal equity. The corollary, for vertical equity, is that those with unequal needs have unequal opportunities to access health care. It is the case that equal access does not necessarily translate to equal utilisation, perhaps due to varying individual preferences. The acceptability of these reasons for differential utilisation should be considered. Equity of access, equity of utilisation and equity of outcomes as equity goals have different implications in terms of the delivery and allocation of health resources.

Equity and efficiency do not align as objectives. In practice, maximising health outcomes may be at the expense of an equity goal (known as the 'equity-efficiency trade-off'²⁵). The corollary is that the achievement of an equity goal may be at the expense of health maximisation.

2.3 WELFARE ECONOMICS AND HEALTH CARE ALLOCATION

Welfare economics is the systematic analysis of the social desirability of any set of arrangements.¹⁵ With respect to resource allocation,¹⁵ it is the development of value judgements which allow a logical and consistent ranking of all alternative social states. With a scarce pool of health care resources, welfare economics allows informed and rational decision-making for resource allocation. Welfare economics recognises that value judgements, and

normative economics, cannot be avoided. A fundamental question for health economics is how to determine the relative desirability of alternative ways of providing health care.

Welfare economics is distinguished by four key underlying assumptions. Firstly, welfarism is a central axiom. Welfarism is the requirement that *"...the evaluation of any social state be based exclusively on the utilities generated in that state."*²⁶ Secondly, it is underpinned by individualism, where social choices are constructed with only the views of those individuals affected considered. This assumes that individuals are the best (and only) judge of their own welfare, with perfect information and rational behaviour always motivating them towards achieving utility-maximisation. The third assumption is consequentialism, where the judgement of an action is based on its impact. Finally, welfare economics is defined by aggregation, which estimates social welfare as the aggregate of individual welfare. Welfare, or utility, is treated as ordinal in modern economics. This requires a value judgement to rank different states of the world, or the observation of preference relations in the real world – revealed preference theory.

If utility is considered a cardinal concept as defined in the early neoclassical approach,²⁷ then the state with maximum utility could be identified via the aggregation of numeric utility, within a defined unit of measurement. With cardinal utility, equation [2.1] provides a social welfare function for state of the world (X), which is a function of the utility obtained by each individual (U) in that state.

$$W(X) = f(U_1(X), U_2(X), \dots, U_n(X)) \quad [2.1]$$

However, economists moved away from the belief that by observing the purchasing of goods by consumers, that you can attach a numerical value in terms of the utility derived. Many economists now assume that utility is only measurable in an ordinal sense. You may observe a preference for a particular good by consumers, but not the relative strength of that preference.

Vilfredo Pareto was one such economist, and his influential work at the turn of the 20th century highlighted that an improvement (a Pareto improvement) in social welfare is possible via a reallocation of resources that makes one person better off without making anyone else worse off. A Pareto optimum is achieved when all Pareto improvements have been made, and therefore the only possible way to make one individual better off is to make another worse off. Adam Smith's 'invisible hand' is confirmed by the First Welfare Theorem – that a competitive market finds equilibrium at a Pareto optimum.

Pareto efficiency can identify between optimal and non-optimal states of the world, however it cannot rank between multiple points of optimality. Also, achieving Pareto optimality in practice is likely to be impossible. The difficulty of achieving the criterion of Pareto optimality stems from the fact that a person's utility can only be inferred from revealed preferences, and so an evaluation of a distribution of resource that impacts on a large population becomes impossible. Instead, Kaldor-Hicks optimality is a less stringent criterion for an efficient allocation of resources. A Kaldor-Hicks allocation is superior to the status quo if the gains could be theoretically be used to compensate all those made worse off by the new distribution. A Kaldor-Hicks optimum is where no more Kaldor-Hicks improvements can be achieved. The difference between Pareto and Kaldor-Hicks is that the compensation does not actually have to be paid. If it is, then the losers do not lose and Pareto holds. The Kaldor-Hicks criterion is an attempt at being less restrictive; however flaws of the method have been identified. These include the fact that individuals require the same marginal utility of income,²⁸ and also the possibility that after an improvement to increase social welfare, a move back to the original allocation could again increase welfare - the Scitovsky paradox.²⁹ It is important to remember that Pareto and Kaldor-Hicks are neutral to any equity or distributional concerns,²⁸ in particular when considering how compensation will have an impact on the distribution of income.

The theoretical concepts of economic efficiency, Pareto and Kaldor-Hicks compensation, have been transferred to applied economics and policy evaluation. If the benefits of a policy are greater than the costs, then in principle the losers could be fully compensated with a net benefit remaining.³⁰ There are numerous market failures associated with the provision of health care, meaning perfectly competitive markets in health are unlikely to exist.³¹ Therefore, the evaluation of policy and government provision and regulation of healthcare is required to ensure efficiency and welfare maximisation.

2.4 ECONOMIC EVALUATION TO INFORM HEALTH CARE ALLOCATION

When considering resource-allocation processes, a useful classification of 'economic processes' and 'non-economic processes' has been determined by several authors.^{32,33} The fundamental difference is that economic methods explicitly account for the scarcity of resources. By deploying a particular resource, it has been exhausted, and can be valued in terms of its opportunity cost, or the benefit foregone from the next best alternative. Each decision to deploy a resource is therefore explicitly valued by the next best alternative that could have been used in its place. Non-economic processes may be undertaken by determining a core set of services, a minimum requirement of need for a population, or via the 'decibel

approach' and political processes.³² However, these non-economic processes do not account for the opportunity cost of the deployment of resources and the differential preferences and demand for health care resources across a population.

The economic processes of health care resource allocation are underpinned by the welfare economic theory detailed in 2.3. Brouwer and Koopmanschap (2000) discuss two competing views regarding economic evaluation in health care; the welfarist approach, and the pragmatist approach.³⁴ The key distinction being that welfarists may try to ground economic evaluation on individualistic models of welfare, compared to pragmatists (or decision-makers) basing their recommendations for economic evaluation on societal values and on more pragmatic assumptions. The limitation of welfare economics in estimating cardinal measures of welfare, in particular to allow interpersonal comparisons of utility has been discussed and debated.³⁵ This saw the development of Paretian theories and the Kaldor-Hicks criterion, as explained in Section 2.3.

These theories have been operationalised as cost-benefit analysis (CBA), with the objective to identify Pareto improvements, or potential-Pareto improvements. CBA requires a consistent unit of measurement for both costs and benefits (usually monetary). The aim is to therefore identify those competing alternatives with a positive net benefit. The link to welfare economics is enhanced when considering Kaldor-Hicks, because the quantification of an individual utility change by compensation can be aggregated across all individuals who are affected. If the sum of compensation across individuals is positive, then this satisfies the Kaldor-Hicks criterion, and is equivalent to the net benefit estimated in the CBA. The primary decision rule for a CBA is to undertake activities with a positive net benefit when compared to the current *status quo*, and with a constrained budget the appropriate rule is to prioritise in order of the activity with the largest net benefit. Undertaking a CBA requires the identification of potential Pareto improvements, and also the identification of all costs and benefits which are relevant to a decision maker. Analyses that are explicitly limited in their capture of costs and benefits are known as partial CBA. Box 2.1 presents the usual summary measures from a CBA, the net present value (NPV) and the cost-benefit ratio. The NPV is the sum of the present value (PV) of both benefits and costs.

The limitations of Paretian theory have seen a movement of welfare economics towards identifying independent arguments in the welfare function (the extra-welfarist approach), which would allow the methods of welfare analysis to survive. Health has been proposed as an important independent argument in the welfare function,²⁸ as it would be an obvious measure to allow interpersonal comparisons within the health system. Therefore health itself, rather

than welfare, can be the objective when looking to evaluate the distribution of health care resources. The challenge in recent times has involved the development of a quantifiable and commensurate measure of health benefit. The quality adjusted life year (QALY) has evolved as a regularly used measure of health benefit, particularly in the UK. It has two primary dimensions, quantity and quality of life. A QALY requires quality of life to be anchored on a scale of one equal to perfect health, and zero equivalent to death. The scores for quality of life are more formally called health related quality of life (HRQL) states, or utility values. One QALY is therefore equivalent to one year spent in full health.*

<p>Where: r = discount rate t = year n = time horizon (years)</p>	$NPV = \sum_{t=0}^n \frac{(Benefits - Costs)_t}{(1 + r)^t} \quad [2.2]$
<p>Where: $PV_{benefits}$ = present value of benefits PV_{costs} = present value of costs</p>	$Cost\ Benefit\ Ratio = \frac{PV_{benefits}}{PV_{costs}} \quad [2.3]$

Box 2.1: Cost Benefit Analysis

The valuation of QALYs, or more formally the valuation of defined HRQL states, does not commit the evaluation to a narrow concept of utility or welfare. Instead, the use of QALY analysis allows health states to be valued, but also allows these values to be determined by peoples values and feelings for a particular health state, or by some objective principles.³⁶

COST EFFECTIVENESS ANALYSIS

The QALY has increased the use of cost-effectiveness analyses (CEA), where benefits are measured in units other than money.¹⁵ Costs and benefits are evaluated across competing alternatives, to inform the allocation of health care resources. Cost-effectiveness analyses are related to cost-minimisation analyses (CMA), where the benefits are assumed equal between two or more alternatives, and therefore the solution is to pick the alternative with the lowest cost. CEA can incorporate benefits of interest such as life years gained (LYG), disability adjusted life years (DALYs), as well as clinical outcomes such as hip fractures avoided. When QALYs are used to quantify health benefits, the evaluation is formally called a Cost-Utility Analysis (CUA). This label recognises that overall health benefits, like utility, are of value.

* 1 year of life x 1 perfect health HRQL state = 1 QALY
 3 years of life x 0.4 HRQL state = 1.2 QALYs

The major advantage of the QALY and CUA is that it allows a comparison of benefits across different treatments and conditions due to the common metric to value health benefits. The consistent application of CUA allows interpersonal comparison within a health system. For consistency and clarity, from now on only the term cost-effectiveness analysis (CEA) will be used, although the assumption is that this is equivalent to cost-utility analysis (CUA), and unless specified QALYs will be the metric of health benefits.

Because a CEA does not provide a direct comparison of the value of the effects and the costs, decision rules are required. If comparing two treatments, one may cost more but also provide more QALYs. The problem for decision-makers is determining how much extra benefit is required to justify the extra expenditure, because there is an opportunity cost associated with allocating resources. A cost-effective treatment is one where, given limited resources, its use will contribute to the maximisation of health benefits. Traditionally, the output of a CEA is reported as a ratio of the difference in costs and effects between two alternatives, the incremental cost-effectiveness ratio (ICER), as demonstrated in Box 2.2.

<p>Where: <i>a</i> = Treatment A <i>b</i> = Treatment B <i>C</i> = Costs <i>E</i> = Effects</p>	$ICER = \frac{(C_a - C_b)}{(E_a - E_b)} \quad [2.4]$
	$ICER = \frac{\Delta C}{\Delta E} \quad [2.5]$

Box 2.2: Cost Effectiveness Analysis

The ICER can be interpreted as the cost per unit of effect, as represented in equation [2.4]. For decision makers, there are four possible situations when using an ICER, which are presented in Table 2.2.

Table 2.2: Results from a cost-effectiveness analysis (CEA)

CEA result	ICER	Interpretation	Fund new treatment?
1. + Inc. Cost + Inc. QALYs	Positive	More costly and more effective	?
2. - Inc. Cost - Inc. QALYs	Positive	Less costly and less effective	?
3. + Inc. Cost - Inc. QALYs	Negative	More costly and less effective	✘
4. - Inc. Cost + Inc. QALYs	Negative	Less costly and more effective	✔

With situations 3 and 4, the solution for decision-makers is very straight forward. However, it is clear that a definitive answer cannot be provided based on just an ICER value. The estimates of costs and effects are required to ensure dominant (less costly and more effective) treatments are adopted and dominated (more costly and less effective) treatments are refused. With situations 1 and 2, the magnitude of the ICER will have an impact on whether it is viewed as cost-effective. The interpretation of an ICER as the cost per additional unit of effect means that the higher the ICER, the more you will have to be willing to pay for that additional unit of effectiveness. Formally, a cost-effectiveness threshold (λ) is a ceiling for which an ICER must be less than for it to be considered cost-effective. The λ threshold can be used as a tool to resolve situations 1 and 2, however 2 is generally rare and often other decision rules are considered, such as not accepting treatments which result in a health loss, which may also be unethical. In situation 1, an ICER for an option below λ is considered cost-effective, or the additional (reduced) effect is at an acceptable cost (saving).

The true definition of the λ threshold value has been widely debated in health economics. Some have proposed that the λ threshold is a societal estimate of the willingness to pay for a QALY.³⁷ This definition allows an empirical estimate of λ to be found through either revealed preferences or studies with hypothetical games. The major issue with this approach is that a societal willingness to pay value for a QALY does not consider the fact that the NHS budget is fixed, and decision-makers must consider the opportunity cost of the deployment of resources. An exhaustive consideration of the costs and QALYs of all NHS services would allow a ranking of all by their cost-effectiveness and a budget to be allocated via a cost-effectiveness league table.³⁸ A threshold λ can be derived from the league table approach; it is the ICER of the least cost-effective intervention that is currently funded. This intervention would ideally be decommissioned by the NHS if a more cost-effective intervention was developed. This approach may be feasible for a small budget-holder with only a small set of activities to prioritise, however for the NHS the task of identifying these opportunity costs is likely to be impossible. Also, the use of a league table could clearly highlight equity issues, such as populations of people having no NHS care.

To allow economic evaluation to inform health care resource allocation, a pragmatic approach has been proposed.^{39,40} The λ threshold can be informed empirically by past decisions, or by estimates from other public sectors. By using past decisions, a decision-maker can be a 'threshold seeker',⁴⁰ and analyses can highlight services which should be displaced, with clear recommendations for disinvestment. The search for a λ threshold does highlight a limitation of economic evaluation and the identification of optimal allocations of health care resources. The

current approaches for the economic evaluation of health resources continue to be contentious, and the ongoing research to inform the λ threshold is of great importance.⁴¹

When an evaluation is of only two competing alternatives, then application of the decision rule is usually straight forward. The ICER can be calculated between the two alternatives and the standard λ threshold applied. However, when an economic evaluation considers more than two alternatives, the decision rules become more complicated, and the quick interpretation of results requires a set procedure for the presentation of results. This procedure, known as an incremental analysis, consists of ruling out options which are simply dominated (where a less costly and more effective alternative exists), and extendedly dominated (where a combination of two alternative options is less costly and more effective). The process for undertaking a full incremental analysis is presented in Box 2.3.

1. Rank options by increasing cost
2. Eliminate options that are simply dominated – there is a less costly and more effective comparator
3. Eliminate options that are extendedly dominated – there is a combination of two other options that are less costly and more effective
4. Calculate the incremental costs and incremental benefits of each remaining option
5. Calculate the ICER compared to the next best (non-dominated) alternative

Box 2.3: Process for undertaking a full incremental analysis

While this process is routinely undertaken for economic evaluations, and ICERs are the standard output for a CEA, the ratio properties of an ICER means that often their interpretation can be a challenge, especially when there are many comparators. Therefore to avoid the limitations of an ICER, it is common to internalise the λ threshold decision-rule which essentially returns us to a CBA, but does not require the imposition of a welfarist framework. The aim is to allow cost-effectiveness to be interpreted by a single figure which is not a ratio. The net benefit approach is used, with λ being used to convert either costs into units of effect, or benefits into monetary units.

The net benefit approach is demonstrated in Box 2.4, with the cost effectiveness threshold λ used to convert either costs or benefits into a consistent unit. Using the net benefit approach for more than two options is simple; the decision rule is to select the option with the greatest net benefit.

	$Net\ Benefit = \Delta E - \Delta C$	[2.6]
Where: ΔE and ΔC are in the same units Cost Effectiveness Threshold λ		
	$Net\ Monetary\ Benefits = \lambda \cdot \Delta E - \Delta C$	[2.7]
If $NMB > 0$, then cost-effective		
	$Net\ Health\ Benefits = \Delta E - \frac{\Delta C}{\lambda}$	[2.8]
If $NHB > 0$, then cost-effective		

Box 2.4: Net Benefit

DECISION ANALYTIC MODELLING

Health technology assessment (HTA) is “..a multidisciplinary process to evaluate the social, economic, organisational and ethical issues of a health intervention or technology.”* It is the framework which underpins health resource allocation decision making. For a CEA to inform HTA, the costs and effects of competing alternatives are required. However, the estimation of costs and effects requires evidence which in general is not observable. Even where costs and effects have been observed, for instance within a clinical trial, the evidence may be limited in terms of its generalisability and appropriateness for decision-making. In particular, the evidence may be short-term, where long-term implications are important, or a trial may not include all relevant comparators (placebo-controlled trials are common). For decision-making within a HTA process, trial and observational evidence may form a subset of the evidence required; however HTA often requires the synthesis of evidence to estimate the costs and effects of a technology in circumstances which often have not been observed. For example, a novel cancer therapy may have a short placebo-controlled trial to prove efficacy, however the long term implications in clinical practice may have never been observed.

Decision analytic modelling (DAM) “represents the real world with a series of numbers and mathematical and statistical relationships.”⁴² Decision analysis and DAM in the context of economic evaluation uses mathematical models to determine the possible consequences that would emerge from the competing alternatives being evaluated. A decision analytic model (from now on referred to as a ‘model’) allows the consideration of the costs and effects of a range of future consequences, with the likelihood of those consequences also being estimated. This allows the calculation of the expected costs and expected effects of each option. The expected cost (effect) is the sum of all costs (effect) of each consequence weighted by the

* http://www.who.int/medical_devices/assessment/en/ - Accessed June 2015

probability of that consequence.⁴³ Brennan and Akehurst (2000) detail five roles and applications which modelling has within HTA, which are provided in Box 2.5:⁴²

1. Extending results (from a single trial)
2. Combining multiple sources of evidence to answer policy questions
3. Generalising results from one specific context to others
4. Modelling to inform research strategy and design
5. Modelling uncertainties in the knowledge base.

Box 2.5: Roles and applications of modelling within Health Technology Assessment⁴²

Modelling has become a key component in HTA and the consideration of the cost-effectiveness of new treatments. Briggs, Claxton and Sculpher (2007) discuss that the increased role of modelling is due to the requirements of economic evaluation. Namely, that all relevant evidence is synthesised, all relevant comparators are considered, that an evaluation time horizon is appropriately long, and that uncertainty in the evidence is captured as decision uncertainty.⁴³

Many models use a cohort approach, where expected costs and effects are estimated for an average person, or a cohort of average persons. The outcomes of the model are therefore population estimates. This includes methods such as decision trees, and cohort state-transition/Markov models.⁴⁴ These methods are often appropriate for the decision context and available evidence, however their shortcomings mean that more advanced individual level methods have emerged. The decision-maker is concerned with the expected cost and effect per patient, to allow an estimation of the cost-effectiveness of a new treatment across a population. However, cohort modelling approaches in general do not allow for variability in patients outcomes according to particular characteristics. However, these characteristics may contribute to costs and effects. Cohort models are often built which allow cohort sub-groups to be tracked through a model, and parameters applied which relate to the particular characteristics of the subgroup cohort. However these approaches only partially capture patient variability, and may result in very complex models, or many models with different subgroup populations evaluated.

An alternative to cohort model approaches are individual level models (ILMs), where individuals with their own characteristics can be simulated in a model and the impact of these characteristics on costs and effects can be captured. Also, the complexity of treatment pathways and the fact that patient history could have an impact on the future can be more easily represented.

ILMs such as patient-level simulation, discrete event simulation (DES), and agent based models estimate individual patient output (cost and effects) which is contingent on individual patient

covariates. When the simulation is run across a large number of individual patients, the expected cost and effects across the population can be estimated. However, the benefits of ILMs are at least partially offset due to the requirement of usually thousands of individual patient simulations, meaning that ILMs are usually more computationally expensive when compared to cohort decision analytic models.

2.5 NATIONAL LEVEL NHS HEALTH RESOURCE ALLOCATION

In England, the National Institute for Health and Care Excellence (NICE) was established in 1999.* It is a HTA decision making organisation and has a mandate from the Department of Health to evaluate (appraise) the health benefits and costs of new and established health technologies and clinical practice. NICE publishes guidance in six areas, as detailed in Table 2.3.

Table 2.3: Types of NICE guidance

NICE Programmes	Guidance published*	Remit	Cost-effectiveness analyses
Clinical Guidelines	158	Advisory	Existing CEA's reviewed Potential CEA's prioritised New analysis – focussed CUA
Public Health	57	Advisory	Existing CEA's reviewed Potential CEA's prioritised New analysis – comprehensive CUA
Technology Appraisals (TA)	270	Mandatory	Existing CEA's reviewed New analysis – comprehensive CUA
Interventional procedures	458	Advisory	None
Medical technologies	25	Advisory	Existing CEA's reviewed New analysis – cost-consequence
Highly specialised technologies	1	Advisory	Existing CEA's reviewed New analysis – comprehensive CUA
Diagnostic technologies	16	Advisory	Existing CEA's reviewed New analysis – comprehensive CUA

* as of September 2015 – includes replaced guidance.

CEA = Cost effectiveness analysis. CUA = Cost utility analysis

* Then called the National Institute for Clinical Excellence, with a remit for England and Wales

Only NICE Technology Appraisals have a mandatory status, which means that if a positive conclusion is reached about the use of a particular health technology in the NHS, then there is legal requirement for that technology to be available if considered clinically appropriate by a patient's physician.⁴⁵ Although NICE Clinical Guidelines are not legally mandated, there is an expectation by the Care Quality Commission that NICE Clinical Guidelines provide the basis for routine clinical practice.⁴⁶

The majority of NICE programmes require a review of relevant economic evaluation literature to be undertaken, and where appropriate a development of a new economic evaluation to ensure that any guidance developed promotes a cost-effective use of NHS resources. Therefore the work of NICE has seen a large increase in economic evaluations for health treatments and services, and has also driven the discussion and development of new methodologies for economic evaluation. The methods, assumptions and evidence used to undertake an economic evaluation for NICE, and similar organisations worldwide, will therefore have a direct impact on health resource allocation and patient outcomes.

For their Technology Appraisals programme, NICE have a "Guide to the Methods of Technology Appraisal."⁴⁷ This document describes the methods that should be used when submitting evidence to NICE, as well as their decision making process. This enables consistency across Technology Appraisals, as well as ensuring the methods used are robust and transparent. For an economic evaluation being submitted as part of a NICE Technology Appraisal, the methods guide prescribes which methods should be used within a 'reference case'. Important components of the reference case include that health effects should be measured using QALYs, and that a time horizon is long enough to ensure all future health and effects are captured.

The methods guide also states that if the most plausible ICER for a technology is below £20,000 per QALY gained, then it is likely to be recommended for NHS use. If the ICER falls between £20,000 to £30,000 per QALY gained, then additional factors will be considered by the appraisal committee. These include: the degree of certainty around the ICER, whether changes to HRQL have been appropriately captured, if the technology is particularly innovative, if the technology is a life-extending treatment at the end of life, and if there are aspects of the technology related to non-health objectives of the NHS. Above £30,000 per QALY gained, then even stronger arguments with respect to the factors listed above will need to be made for a positive recommendation to be passed.

Therefore the decision rule used by NICE is predominantly based on cost-effectiveness, although there is the ability for the decision-makers (one of four appraisal committees) to incorporate their judgement regarding those other specific factors.

The theoretical construct behind NICE's decision making approach is that recommending technologies under £20,000 - £30,000 per QALY will require the disinvestment of technologies that have an ICER above £20,000 - £30,000 per QALY gained. This enables total QALYs to increase but with no net cost to the total NHS budget. However, while the appraisal process clearly explores the costs and benefits provided by a new technology, the interventions that are displaced are unknown. This is a source of criticism for NICE, and some have commented that the approach is un-economic, with disinvestment and opportunity cost not explicitly accounted for.⁴⁸⁻⁵⁰ However, proponents of NICE highlight that the methodology is grounded in extra-welfarism theory, while maintaining a process which is pragmatic and feasible.

2.6 OPERATIONAL RESEARCH

Operational research (OR) is the discipline of applying advanced analytical methods to help make better decisions.* OR is also known as 'management science'.⁵¹ It developed as a discipline during the wartime period of the early 20th century to inform decision-making regarding military strategy and resourcing, and in the post-war period operational researchers moved out into industries such as steel and engineering. OR expanded in academia, as well as private and governmental organisations, during the second half of the 20th century.⁵² The increase in computational power during the 1980s and 1990s saw practical applications of OR methods in wider fields, and interdisciplinary working across academic fields and organisations has been a key characteristic of OR.

Modern OR is often dichotomised as 'hard' OR, which involves quantitative analysis, and 'soft' OR, which incorporates non-mathematical techniques and problem-structuring methods (PSMs).⁵³ Soft OR includes methods for structuring and exploring (often complex) problems, and the facilitation of engagement in problem-solving, and is often best considered as methods to tackle problems in their own right, as well as being complements to hard OR methods.⁵³ The hard OR methods draw upon mathematics, economics and computing for theoretical underpinning. This is to be expected, due to OR being developed as a relatively new interdisciplinary field that included people with these backgrounds. In particular, areas of common interest include optimisation, game theory, production, finance and forecasting.

* www.theorsociety.com – Accessed June 2015

However, it should be highlighted that often economists and operational researchers have different interests, and therefore the fields are distinct. Health economics and health resource allocation is an area where the fields of economics and OR have collaborated and combined.

OR methods have had an important role to play in HTA and economic evaluations. Soft OR methods, in particular PSMs and model conceptualisation methods, are used to support the development of DAMs for HTA.⁵⁴ Hard OR methods, such as simulation, have been used to provide estimates of long term costs and QALYs for economic evaluations via the development of ILMs.⁴⁴

Optimisation is the task of making the best decision among various alternatives.⁵⁵ Optimisation is a prominent area in OR, involving analytical and heuristic methods to identify an optimal (best) or near-optimal solution for a particular maximisation or minimisation problem. Optimisation allows the representation of health resource allocation as a maximisation problem. That is, how to maximise population health given a fixed monetary budget and competing healthcare programs. Health care resource allocation can be represented in a linear programming (LP) formulation, as demonstrated by Stinnett and Paltiel (Box 2.6).⁵⁶ Linear Programming is a form of constrained optimisation where the objective function (the function that is to be maximised) is linear. The constraints are that the proportion of each program i can only be between 0 and 1, and the total cost must not exceed the budget C .

Maximise:	$\sum x_i e_i$	[2.9]
Subject to:	$0 \leq x_i \leq 1$ (for all i)	[2.10]
	$\sum x_i c_i \leq C$	[2.11]
Where:		
e_i and c_i are the effectiveness and cost of program i if fully implemented		
C is the total available budget		
x_i is the proportion implemented of program i		

Box 2.6: Linear Programming formulation of health care resource allocation⁵⁶

A contested aspect of the linear programming formulation, along with CEA more generally, is that implementing only proportions of programs (their divisibility) is not likely to be appropriate. Also, the divisibility of programs is related to the assumption of constant returns to scale for programs, which is contested most notably by Birch and Gafni.^{48,57} As a program increases in size, it is expected that the output will increase by more than a proportional increase in inputs (at least to a certain point), but the linear programming formulation does not account for this. Instead, a solution is to formulate the problem as an integer programme (IP). This has been presented by Torrance (1972)⁵⁸ and Birch & Gafni (1992), as illustrated in

Box 2.7.⁵⁹ Constraint [2.13] has been altered from the linear programming formulation so that programs can only be fully implemented, and therefore programs are not divisible.

Maximise:	$\sum x_i e_i$	[2.12]
Subject to:	$x_i \in \{0,1\}$	[2.13]
	$\sum x_i c_i \leq C$	[2.14]

Box 2.7: Integer Programming formulation of health care resource allocation⁵⁹

The IP formulation above is defined in OR as the knapsack problem (which is an IP with just one constraint). Given a set of items with a given weight and value, the problem is to determine which items to place in a knapsack with a given weight limit as to maximise the total value. For health, the constrained optimisation problem is to allocate a fixed budget to competing programs, each with their own cost and effectiveness. The problem cannot be solved analytically, and therefore mathematical algorithms are developed to solve the problem. When mathematical algorithms are required, the assumptions and the size of the problem have a direct impact on the ability to reach an optimal answer within a given period of time.

The algorithms used to solve or approximately solve constrained optimisation problems like the knapsack problem may also have a role when looking to estimate the appropriate sequence of therapies for patients with chronic conditions. The objective is to maximise net benefits by selecting treatments from an available set, each with their own cost and health benefit.

There is therefore the potential to apply optimisation methods from OR to address economic evaluations which have a very large set of competing alternatives. In the following chapter, the treatment sequencing problem for economic evaluations of chronic conditions is introduced. In Chapter 6, a systematic review of OR literature to identify optimisation methods which may be applicable to the treatment sequencing problem is undertaken.

2.7 CONCLUSIONS

The use of economic evaluations and cost-effectiveness analyses to inform the allocation of scarce health resources has increased in recent years, in particular with the development of organisations like NICE. However, these methods are likely to be limited in their ability to achieve allocative efficiency in a true economic sense, due to the pragmatic approaches required to allow a feasible estimate of the impact of alternative states of the world. Instead,

economic evaluation represents a transparent and formally defined set of methods which allow an informed attempt at meeting social objectives where government provision of health care is required. The use of DAMs and OR methods has increased, to allow the synthesis of evidence and the estimation of lifetime costs and QALYs for complex conditions and treatment pathways.

Chapter 1 provided an introduction and brief motivation for this thesis. This current chapter provides a theoretical underpinning for the remaining thesis. In particular, that models are simplifications of reality, and therefore 'wrong', however to allocate health care resources fairly and efficiently they are required to better inform decision-makers.

In the next chapter (Chapter 3), the treatment sequencing problem is formally introduced, and the rationale for this thesis is provided. The chapter explores what treatment sequences are, when they are used, and why they present a challenge for both developing models and undertaking economic evaluations. After this, the following chapter reveals how the treatment sequencing problem has not been formally address in rheumatoid arthritis (RA), via a systematic review (Chapter 4). A model is then developed for an economic evaluation to inform a NICE appraisal of biologic treatments for RA (Chapter 5). A review of OR optimisation methods is conducted to identify suitable methods to solve the treatment sequencing problem (Chapter 6). A method is selected, implemented and evaluated in Chapter 7, before discussions and conclusions are drawn in the final chapter.

CHAPTER 3: RATIONALE FOR THIS THESIS

3.1 INTRODUCTION

The aim of this thesis is to identify a simulation optimisation method which can be applied to an economic evaluation of sequential therapies for a chronic condition. The rationale for undertaking this research will be described in this chapter.

In Section 3.2, treatment sequences are defined and the reasons for their use are explored. In particular, the conditions in which a treatment sequence may be used are highlighted. In Section 3.3, the reasons why treatment sequences represent a unique challenge for economic evaluation are introduced. It is demonstrated why a long sequence or large number of potential treatments can result in an infeasible number of sequences to compare explicitly.

In Section 3.4, a consideration of the influence of pharmaceutical markets and the development of new technologies is reported. These often compound the difficulties in evaluating treatment sequences.

In Section 3.5, there is a discussion regarding how patient choice may impact on the development of guidance for treatment sequences. Section 3.6 draws conclusions. Section 3.7 and Section 3.8 report the aims and objectives, and the structure of the thesis, respectively.

3.2 TREATMENT SEQUENCES

A literature search using Google Scholar and PubMed was undertaken, but no agreed definition of ‘sequential treatments’ was identified. However, it is frequently used to describe medical treatment which may follow or precede other treatments.^{60–62} For this thesis, sequential therapies are defined as *“the purposeful use of treatments administered one at a time to manage a condition over time”*. However, to improve the definition, it should be stated that *“the sequence is determined a priori”*, and so the subsequent treatment or remaining sequence are known before a prior therapy is delivered. This is because in clinical practice the pool of available treatments will be considered before selecting a first therapy, and potentially the order of subsequent therapies. In general, a treatment refers to one particular medical drug, however many drugs can be used concomitantly, and so combination drug therapy also represents a particular treatment within the set of available treatments.

The definition highlights that at a decision point* a physician and patient may have multiple treatments available to them, and therefore have multiple potential sequences of treatment to consider. The selection of a sequence is based on considering the set of available treatments and all possible sequences.

An example of the issues which arise in considering sequences of treatments is given by the NICE Technology Appraisal of tocilizumab for the treatment of RA.† In this evaluation, treatment sequences were widely discussed and considered, because tocilizumab could be approved in multiple positions within a treatment sequence. The appraisal focussed on the assumptions made by the manufacturer, Roche, regarding the sequence of treatments for patients with RA. Alternative assumptions about what may occur both before and after tocilizumab treatment had a significant impact on the potential cost-effectiveness of tocilizumab treatment compared to conventional DMARDs.

During the course of the appraisal, the NICE Decision Support Unit (NICE-DSU) was commissioned to provide additional evidence for the appraisal committee. In their report to the committee, the authors Palmer and Sculpher comment that *“...the appropriate use of new biologics therapies in RA inevitably involves a consideration of the appropriate sequence of therapies given the chronic nature of the condition, the fact that therapies do not typically remain efficacious and tolerable on a permanent basis and the availability of a number of biologic therapies which are licensed for RA.”*⁶³ This statement is found in a small report within a complex technology appraisal but contains three primary reasons for why treatment sequences are used in clinical practice:

- Chronicity – a chronic condition is more likely to require treatment with a treatment sequence
- Uncertainty – when the effectiveness of a treatment is limited, and response to a treatment cannot be predicted, then a sequence may be used to identify a treatment which is effective
- Competition – if there are many possible treatments available, then sequential therapy may be used.

In the remainder of this section, these reasons for why treatments are used in clinical practice will be examined further (Competition is discussed within Section 3.4). Also, some further

* A decision point could be a new diagnosis, or the failure (loss of efficacy or an adverse effect) of a treatment and a consultation to determine the next therapy

† <http://www.nice.org.uk/guidance/ta247> - Accessed June 2015

reasons will be considered to determine the key characteristics which are required for treatment sequences to be clinically appropriate.

EFFECTIVENESS

The effectiveness of a particular treatment is defined as the extent to which “...they achieve health improvements in real practice settings.”⁶⁴ In clinical practice, the effectiveness of a treatment may be determined by a measure of improvement, which may be a measure of survival, HRQL, or a surrogate outcome such as a laboratory test for a biomarker. Determining effectiveness may be complex because a treatment may have multiple impacts on health. For example, a treatment might immediately reduce disease activity, but increase the risk of a future undesired health effect. These multiple effects, both benefits and risks, can be aggregated to provide a benefit-risk profile for a particular treatment.⁶⁵ Formally, a benefit-risk profile is a “reflection of the overall balance of a treatment’s potential benefits with its identified risks as revealed through the safety and efficacy evidence.”* A benefit-risk profile could be a formally weighted and quantitative assessment of a treatment, or a qualitative judgement. Quantitative benefit-risk assessments are included in the regulatory processes undertaken by licensing bodies such as the Food and Drug Administration (FDA) in the USA and the European Medicines Agency (EMA). New medical products have to provide evidence to ensure that the potential benefits clearly outweigh any risks. Qualitative judgements regarding the benefit-risk profile of a treatment are routinely undertaken by a clinician who draws upon evidence, experience and clinical expertise, along with considering a patient’s views and particular circumstances.

Many treatments will only have a limited effectiveness, both for chronic and acute conditions. In some cases it may be immediately obvious that a treatment has not been effective in achieving patient benefit. For some treatments, a consideration of the long-run outcomes is required, where any benefit may diminish over time, or negative effects such as adverse events may occur. In either case, the clinical decision to switch to another treatment to try to achieve health benefits may be made. The limited effectiveness is dependent on both the treatment, and also the condition. A person’s condition may become refractory over time, and control be lost, or a treatment may provide short term benefit but is not sustained. For chronic conditions, where a treatment cannot provide cure, it may be the case that a treatment will only be effective for a proportion of a patient’s lifetime.

The effectiveness of available treatments is one determinant of the likelihood of treatment sequences being required. If a treatment is curative, or highly effective, then there is less

* <http://www.hc-sc.gc.ca/dhp-mps/homologation-licensing/gloss/index-eng.php#b> - Accessed June 2015

reason to require a sequence of therapies. However, if the treatments available for a particular condition are less effective, or a condition is chronic, then a sequence of treatments is more likely to be observed. Modern medicine, in particular for chronic conditions, has seen the development of long-term treatment plans.⁶⁶ These plans frequently incorporate a strategy for sequence selection to ensure effective long term care.

UNCERTAINTY

Often in clinical practice, a clinician is uncertain about the best treatment for a particular patient. Treatment sequences recognise the limitations of modern medicines and treatments, that there is uncertainty with respect to their effectiveness, and also to their adverse event profile. If physicians knew with certainty that a treatment would be effective, or that a treatment would definitely cause a serious adverse event, then clinical decision-making would be rather easier than it actually is. Sequences emerge because a treatment may work in one patient, but in a clinically identical patient the same treatment may not work. One patient may not tolerate a treatment, while another can. There is uncertainty about how a patient will respond to a particular treatment, and treatment sequences emerge through a process of trial and error - persisting when a treatment works and switching if it does not.

With this uncertainty, sequences are required to achieve clinical goals. These goals may be the remission or cure of a condition, the control of a chronic condition, or ensuring that a patient remains alive. In all cases, a purposeful decision is made to employ a sequence of treatments to account for treatment failure.

An example of this uncertainty is the use of methotrexate, a disease-modifying anti rheumatic drug (DMARD) for the treatment of people with rheumatoid arthritis (RA). For many patients with RA, methotrexate is effective at reducing disease activity and improving HRQL.⁶⁷⁻⁷⁰ However, the mechanism of action for methotrexate is still not fully understood, and for some patients methotrexate is not effective.^{71,72} The uncertainty is a challenge for clinicians who are initiating DMARD therapy, because their effectiveness is unpredictable, and the mechanism is still unclear. This uncertainty leads to variation in treatment selection, with alternative first line treatments being prescribed by clinicians.⁷³ It also leads to variation in the time spent on a treatment, with therapies tried and rapidly changed if not immediately successful. This rapid switching through alternative therapies naturally leads to list of alternatives to try if a first line DMARD fails; a treatment sequence.

CHRONICITY

The chronicity of a particular condition increases the likelihood of sequential therapies being utilised. If a condition is acute, either the treatment may be curative, or there may only be a

short period of time to intervene and so a sequence of treatments will not be used. Acute conditions may require follow-up care or support, however the future downstream treatments are unrelated to the initial decision-point and so these downstream treatments are not formally part of a treatment sequence. It should be noted that this has implications for how economic evaluations are conducted and will be discussed further in Section 3.3.

Where a condition is lifelong and chronic, sequences of treatments are very likely to be observed. The emphasis of treatment is on long-term control of the primary condition and its impact on a patient's HRQL. Control (the impact of a treatment on a person's condition) may be specific to symptoms (short term control) or disease activity (long term control). The objective of a treatment may be to control acute episodes of symptomatic and active disease which may be a characteristic of a chronic condition (e.g. 'flares' in RA⁷⁴ and Crohn's disease⁷⁵, relapses in Multiple Sclerosis⁷⁶). The natural history of chronic conditions may be of chronic progression, where disease activity increases over time and worsens a patient's HRQL. Alternatively, it may be of a relapsing and remitting nature, where an episode of active disease begins and a patient is provided with active treatment. Relapsing-remitting conditions include multiple sclerosis and some forms of chronic depression, where there are periods of active and inactive disease. Treatments for relapsing-remitting conditions may only have short term benefits for the current episode of disease, or they may have long term benefits where they reduce the risk or increase the time to a future relapse. Relapsing-remitting conditions may lend themselves to sequential treatments more naturally, because a treatment may only provide short term disease control and repeated control may not be possible or plausible, and so a switch to a different treatment is made when a patient experiences a future episode of active disease. These 'rhythms' are important when considering how a decision analytic model may be constructed to evaluate the cost-effectiveness of treatments and/or treatment sequences.⁷⁷

The characteristics and the impact a chronic condition will have on a patient's HRQL define the treatment plan developed for a patient, and also the use of sequential therapies. Also, a condition may not be acute or chronic, but instead have an element of chronicity or long-term duration which may lead to the use of sequential therapies to maintain long term disease control.

SWITCHING RULE

As has been discussed in the previous section, factors including chronicity, effectiveness and uncertainty will lead to a decision to purposefully employ a sequence of treatments to manage a condition over the long term. The reason for change from one treatment to another is often

called the 'switching rule'. In clinical practice, this may be a clinically observable and measurable signal that a patient will receive greater benefit from switching to another treatment, as opposed to continuing with the incumbent treatment. Alternatively, it may be because a treatment is toxic and the benefit/risk profile has changed.

In particular, the switching rule is the result of a sequence of treatments being possible. If there are no alternative treatments, then the only switch possible is to no treatment, and an active treatment will be used for as long as health benefits are realised. If many treatments are available, the switching rule may allow a change of treatment after a short period of time or if only small benefits occur. This allows alternative treatments to be attempted and rapidly changed until one is found to provide significant benefits.

In recent years, there has been the development of biologic therapies. In contrast to standard molecular pharmaceuticals, these treatments are an extraction or semisynthesis of a biological source. They often look to copy the effect of substances that are produced by a body's immune system. They are highly effective in treating many conditions, but their unique design and process for manufacture means they are costly to produce, as well as very profitable for manufacturers because generics (formally biosimilars) are harder to produce after patent expiry.

For patients with RA, treatment with biologics is limited to people with severe active RA (Disease Activity Score 28 (DAS28) > 5.1). Some biologics have been shown to be effective in moderately active RA ($3.2 < \text{DAS28} < 5.1$), however there is less scope for improvement (i.e. fewer QALYs are generated by moving from moderate RA to remission, compared to severe RA to remission) and so the use of biologics in moderate RA is less likely to be cost-effective when the dose (and therefore treatment cost) is the same. Therefore a switch to a subsequent treatment may only occur when disease severity has increased and the patient is back in a severe active RA state. Determining an appropriate switching rule is often where cost-effectiveness analyses and decision-analytic modelling can prove to be very informative. It also raises the distinction between individual clinical decision making and decision-making based on population health. Clinical decision making is concerned with health improvements for each patient who is face to face with the clinician, but population health maximisation requires trading off the health gains between individuals to ensure that population health is maximised. If a clinician is a budget holder or is very budget aware, an individual clinician may have to consider both perspectives, to their considerable discomfort in some cases.

3.3 TREATMENT SEQUENCES AND ECONOMIC EVALUATION

The time horizon of a cost-effectiveness analysis should extend far enough into the future to capture the major health and economic outcomes – both intended effects and unintended side effects and costs.⁶⁴ By omitting any future consequences, the present value estimated for a particular intervention may be biased, and a sub-optimal decision could be made regarding the allocation of health resources. Future costs and consequences which do not occur at time zero, but at some time in the future, must be taken into account.⁷⁸ This is to ensure that all costs and consequences of choosing one action over another are captured. Future costs and consequences will include treatments which were not administered at the primary decision point, but are subsequent therapies as part of a sequence.

An economic evaluation contributes to a particular decision problem. Often for NICE, the decision problem is whether or not to recommend a new treatment in a particular patient population. For NICE, the decision problem is clearly specified. This decision problem determines the focus of the economic evaluation.

A cost-effectiveness analysis may compare treatment A vs. treatment B at a point in time; however it is more appropriate to consider it a comparison of treatment pathway with A vs. treatment pathway with B. If the treatment is likely to have long term costs and consequences, then a short time horizon is unlikely to be appropriate. More specifically, treatments within a sequence will have different effectiveness and treatment duration, meaning that a switch may happen sooner or later, and the patients at that point of switch will be different. This means the costs and benefits attributable to a downstream sequence of treatments do not just cancel out in the comparison. NICE in their Guide to the Methods of Technology Appraisal (2013) recommend that *“the time horizon should be sufficient to reflect important cost and benefit differences between the technologies being compared.”*⁴⁷

In Figure 3.1, the QALY profile for two treatments, A and B, are illustrated. Treatment A is both life enhancing and life extending. The curves represent HRQL over time, with the area under the curves the QALYs gained. The area between the two HRQL curves represents the QALYs gained by A compared to B. Only by estimating the QALYs up until death (assumed to be where HRQL = 0) will the full impact in terms of QALYs gained be estimated.

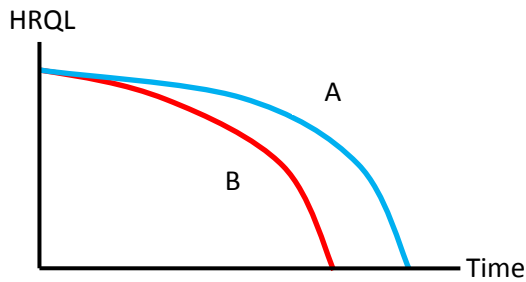


Figure 3.1: HRQL profile

For a chronic condition of a relapsing/remitting nature, the HRQL profile may look different. In Figure 3.2, a patient has a relapsing/remitting condition and is treated a sequence of two treatments:

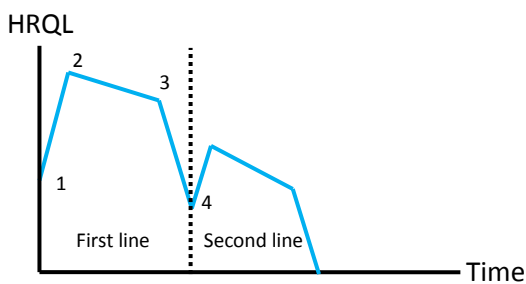


Figure 3.2: QALY profile - relapsing/remitting chronic condition

At point 1 a patient begins a first line treatment which significant and quickly improves their HRQL up to point 2. The disease slowly worsens over time while on treatment. At point 3, treatment effectiveness is lost and the patients HRQL declines during a relapse of severe disease until point 4. At this stage, a treatment switch occurs and an improvement (albeit smaller) in HRQL is observed. The condition continues to worsen. The overall trajectory of the condition and the overall QALYs that are aggregated across a patient’s lifetime are a function of the magnitude of initial response, what happens over the longer term, how long a treatment is effective for, and when a patient will die.

The effectiveness of a downstream treatment is contingent on what has happened as a result of previous treatments. The patient may be more or less likely to achieve a response if response was obtained on a prior therapy.* The patient may be older, and in turn have more or less potential for response. The condition may have caused irreversible changes to a person and therefore a subsequent treatment may have less potential for benefit.

* A patient may be more likely to respond if they responded to a previous treatment. Alternatively, a patient may be less likely to respond to a treatment if the previously responded to a treatment with a similar mechanism.

In Figure 3.3, three different comparator sequence HRQL profiles (S_B) are compared to a common sequence HRQL profile (S_A). S_A is identical in all three scenarios.

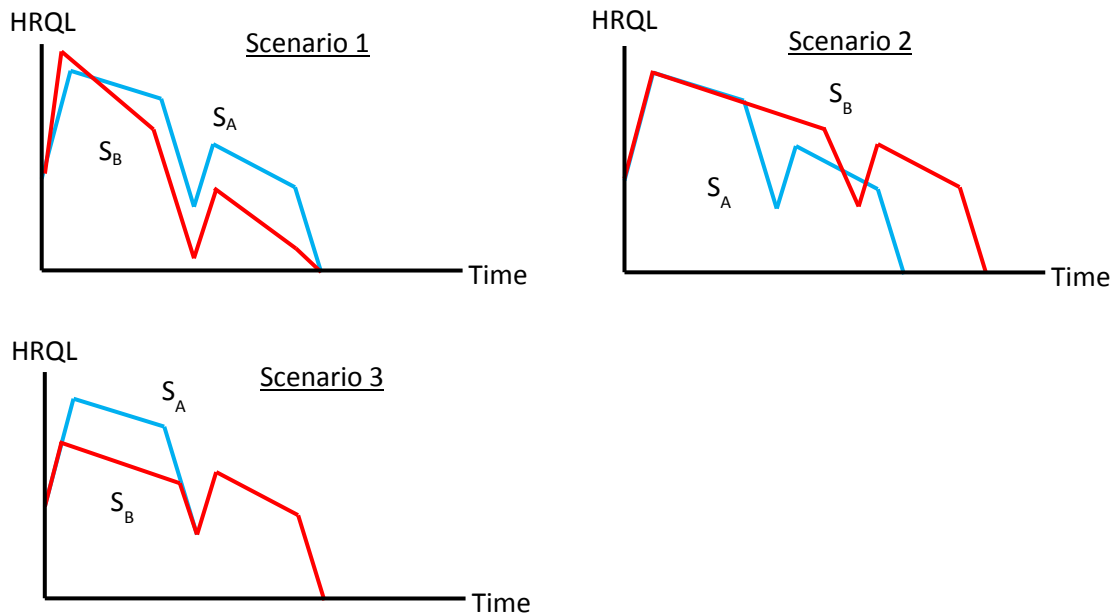


Figure 3.3: QALY profiles for comparator sequences

In Scenario 1, S_B has an improved initial response, but faster worsening of HRQL over time. The switch to a second treatment sees an improvement in HRQL, but HRQL is below where it would be on S_A . Although death occurs at the same time, there is a net QALY loss on S_B . A model will need to capture the full sequence until death to ensure the full impact on QALYs is captured.

In Scenario 2, S_B has an identical initial response, and HRQL declines at the same rate as S_A . However, the patient remains on treatment for longer before switching. The QALY contribution of the second line treatment is identical, but due to occurring at different points in the future, the discounting undertaken to account for time preferences and estimate present values will result in different estimates of this QALY contribution. Again, a model will need to capture the full sequence until death to ensure the full impact on QALYs is captured. Although it may be fair to assume a future treatment will provide the same benefits, if treatments are initiated at different times then discounting to obtain the present value will reduce the QALYs accrued.

Scenario 3 is the only illustration provided where the second line treatment in S_A and S_B results in identical QALYs, and in an economic evaluation these will cancel out between the two comparator sequences. In this case, it would be fair to develop a model which has a time horizon up until the end of the first treatment, because an accurate estimation of lifetime QALYs is not required to provide an accurate estimation of the incremental QALYs for the

purpose of an economic evaluation, there is no incremental difference after the first line treatment.

For any economic evaluation there is a decision point, time zero, where a comparison between alternatives is made. Some researchers have called this point the 'divergence point' of a decision analytic model, because the treatment sequence up until the divergence point is identical. An evaluation of sequential therapy requires the movement of the divergence point to the first line therapy, so that the consequences of all possible subsequent sequences can be estimated and compared. The evaluation of a sequence of treatments is important if there is a particular treatment which can be used at multiple points within a sequence. For a model to allow an estimation of costs and QALYs of all possible sequences, then a model is required which allows treatments to vary along the full length of the sequence. Also, the comparison is between alternative sequences of treatments, rather than a head to head comparison between one particular treatment against a comparator treatment.

Figure 3.4 provides a schematic for a decision tree model to evaluate three treatments (A, B and C). There are three treatment pathways for the model to enumerate the costs and QALYs.

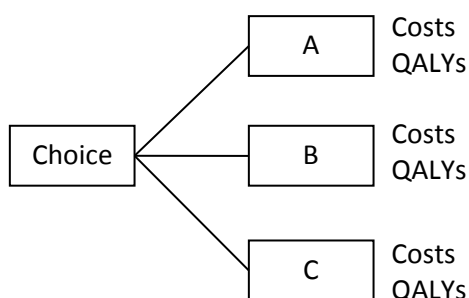


Figure 3.4: Three treatment decision tree model

However, if the three treatments (A, B and C) can be used sequentially, once only, and in any order, then there are six treatment pathways to be enumerated. This is shown diagrammatically in Figure 3.5.

The additional pathways (branches) add both complexity to the decision analytic model and evaluation, and require greater amounts of evidence about the costs and consequences of each pathway, and the position of treatments therein.

As the number of treatments included in the sequence increases, it becomes increasingly challenging to know what the true treatment effects are likely to be for every technology in every position in the sequence. For example, if treatments have been studied in clinical trials as first-line treatments, but then are placed second or third-line in a sequence, the efficacy of

these treatments in the sequence may be very different from that observed in the trial. The corollary is that, as seen frequently with modelling treatment sequences, there is a danger that the model has parameters that are difficult to estimate.

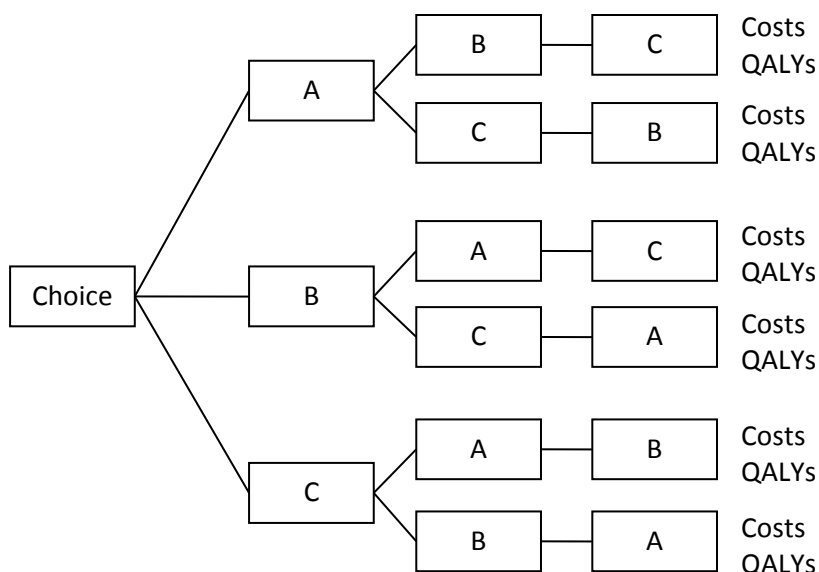


Figure 3.5: Three treatment sequential decision tree model

To resolve this issue, treatment effect decrements have been used in NICE appraisals.⁶³ These use the assumption that a treatment will be less effective, the further down the treatment sequence it is used. The effectiveness evidence (often a trial) may provide parameters for a particular point in the sequence (e.g. first line). This evidence is used to provide treatment effectiveness parameters for the model if used in other places in the sequence (where trial data may not exist) and a decrement applied to reduce the effectiveness parameters and account for diminished effectiveness. These decrements could potentially be informed by external data, potentially from registries, or from expert opinion (including elicitation).

Observational studies could potentially be used to estimate treatment effects, although the limitations of this approach have been widely discussed.⁷⁹⁻⁸¹ In particular, estimating treatment effects from observational studies is likely to be biased because groups of patients (e.g. patients on different treatments) will vary systematically in ways related to the outcome of interest – unmeasured confounders. These confounders can result in mistaken causal inferences and biased estimates of treatment effect unless they are accounted for.

HOW MANY SEQUENCES?

The number of sequences to be enumerated in an economic evaluation is dependent on the number of comparator treatments, and assumptions about the perspective of the sequence. The sequence perspective determines how a new treatment will interact with the existing

treatment sequence. The perspective of the sequence affects the potential size of an economic evaluation and the number of comparators therein.

Firstly, are the existing treatments within a sequence in a fixed sequential order? An example is a treatment only licensed for use after a previous therapy. In this case, there is at least a partially fixed sequence of treatments.

Secondly, is the length of the sequence fixed? If so, the new treatment will replace a therapy in the sequence to ensure the length of treatment sequence is fixed. If the sequence is not fixed, then it is possible for a new treatment to be an addition into the treatment sequence and the resulting sequence is extended.

Thirdly, are truncated sequences of treatments possible? This means that instead of providing a full sequence of active treatments for a patient's lifetime, a sequence could instead be stopped. Therefore the evaluation requires the inclusion of all truncated sequences for comparison.

For hypothetical situations the number of sequences to be evaluated can be calculated, as shown in Table 3.1. However, in practice, it is likely that the sequences are more complex. Some treatments may have a fixed position, as determined by clinical guidance or their licensed indication, and some treatments may have a fixed position relative to other therapies (again, determined by clinical guidance or by their licensed indication).

In Table 3.1, A and B are the two existing treatments, and X is a new treatment. The number of possible sequences to evaluate in order to identify an optimal sequence is dependent on the perspective of the sequence. If the sequence is ordered, that means A and B must retain their order. If the sequence is variable, then all treatments can be in alternative orders. If a new treatment can be an addition, then the sequence length will extend. However if the treatment is a replacement, the length of the sequence is fixed and an existing treatment is removed. If truncated sequences are possible, then all sequences of all possible lengths must be estimated.

With an ordered sequence (either addition or replacement), the number of sequences to compare grows linearly with the number of treatments eligible. With a variable sequence, the growth is exponential. If truncated sequences are plausible, then every plausible truncated variable sequence must be included for evaluation. Suddenly for just eight unique treatments, the number of treatment sequences that are plausible could be over 100,000.

Table 3.1: Number of potential sequences given the perspective

Sequence perspective	Comparator sequences (2 existing treatments (A and B) and a new treatment (X))		Permutation formula $n =$ number of treatments	Number of sequences		
				$n =$ 3	4	8
Ordered - addition	{A, B}		$n + 1$	4	5	9
	{X, A, B}					
	{A, X, B}					
	{A, B, X}					
Ordered - replacement	{A, B}		n	3	4	8
	{X, A}					
	{A, X}					
Variable - addition	{A, B}	{B, A}	$n! + (n - 1)!$	8	30	45,360
	{X, A, B}	{X, B, A}				
	{A, X, B}	{B, X, A}				
	{A, B, X}	{B, A, X}				
Variable - replacement	{A, B}	{B, A}	$n!$	6	24	40,320
	{X, A}	X, B}				
	{A, X}	B, X}				
Truncated (subsets)	{A}	{A, B}	$\sum_{k=1}^n \frac{n!}{(n-k)!}$	15	64	109,600
	{B}	{B, A}				
	{X}	{X, A, B}				
		{A, X, B}				
		{A, B, X}				
		{X, A}				
		{A, X}				
		{B, X}				

TREATMENT SEQUENCE OR DOWNSTREAM TREATMENT?

It must also be noted that treatment sequences are not always the same as downstream treatments (often called ‘future costs’). Downstream treatments are the full range of possible treatments that may be employed after the use of a particular intervention. These treatments may be related to the primary condition, and therefore are conditional on the effectiveness of the primary intervention, or they may be unrelated to the primary condition. For example a patient with RA may be treated with a DMARD, and a downstream treatment may be a biologic DMARD. However some patients with RA may also have depression, and will receive unrelated (to RA) downstream treatments for their depression. With most modern health services structured around specialist treatment and care, unrelated downstream treatments are unlikely to be the primary concern for a physician, however several economists argue that all costs incurred, related or not, should be included.⁸²⁻⁸⁴

However, related downstream treatments are not always the same as a treatment sequence, because a sequence implies that the order of the therapies can alter. A downstream treatment for the purpose of this thesis is defined as one which cannot change position with another treatment. For example, a downstream treatment may be palliative care for a patient, and would not be provided prior to an active therapy. Treatment sequences therefore capture many of the instances where downstream treatments may occur, however reversibility represents a key difference.

Future related health costs still require inclusion within a model to provide an estimate of the cost-effectiveness of a particular treatment or treatment sequence. This is because the initiation of that particular treatment (sequence) reflects a decision about a course for the patients' condition, and therefore an evaluation of its cost-effectiveness should include health costs that are attributable to the primary treatment.

3.4 PHARMACEUTICAL MARKETS AND TREATMENT SEQUENCES

The pharmaceutical market continues to expand, in particular for chronic conditions. The population of many developing countries is ageing, and novel therapies are being developed with improved effectiveness and tolerability. The RA drugs market in 2010 generated an estimated \$12.7bn,^{*} and is expected to continue increasing.

In a potentially profitable market, pharmaceutical manufacturers will look to develop novel therapies for the improved treatment of a condition. Also, manufacturers will look to develop 'me-too' therapies, and generic equivalents when patent protection expires. These competitor products increase therapeutic options for clinicians and patients, and therefore increase the potential use of sequences to effectively manage chronic conditions. Also, pharmaceutical companies will compete for market share to generate a return for their investment in developing the new product. In a condition with many incumbent therapies, a company may promote sequential use of their product. This decision to build the therapeutic value proposition around a sequence is likely to be made early in the development process so that appropriate trials and evidence can be generated for both marketing authorisation and reimbursement.

New treatments with marketing authorisation are often constrained by the earliest they can be used (i.e. not before another treatment has been attempted), or by a level of disease

^{*} <http://www.visiongain.com/Report/679/Rheumatoid-Arthritis-World-Drug-Market-2011-2021> - Accessed June 2015

severity. Therefore the position of treatment within a sequence is not fully fixed, and there may be several potential positions that a drug can be used in clinical practice. If a license is not restrictive, a drug could potentially be used in any position in a sequence, and therefore allow alternative sequences of treatments to be used in clinical practice.

The biological DMARD (bDMARD) anti-tumor necrosis factor alpha therapies (TNFa's) for patients with RA are broadly similar. They are targeted at the same cytokine, although their mechanism may be different.* Also they may differ through mode of administration, and by their effectiveness and adverse event profile. The similarities of newer TNFa therapies to existing TNFa therapies means they have been categorised as 'me-too' therapies.^{85,86} The term 'me-too' in general refers to products which have many similarities to a competitor, but have enough of a difference to allow them to be marketable as a separate product along with its own patent. A me-too product on which research is started after the initial product in the class is on the market is likely to be less costly to develop, compared to the novel therapy. Much research and development may be conducted to develop an innovative therapy, and then only a relatively small amount of research and development is required to alter the original therapy and obtain a separate licence. The 'me-too' title may be unfair on products which come to market only a short time after the novel therapy, because the long process of drug development means that they were probably equally innovative during their development but just happened to enter the market slightly later.

Me-too therapies, along with all treatment developments in a particular condition, allow price competition between competitors.⁸⁷ This price competition means that often prices are very similar (or identical) for rival products. This adds complexity to evaluating cost-effectiveness, because similar cost and similar effects can mean that the results are extremely sensitive to evidence, assumptions and uncertainty. The availability of multiple therapies also supports physician and patient choice to select a particular drug based on favourable characteristics such as mode of administration, particular side-effect profiles, or patient co-morbidities. These attributes may not directly affect HRQL, or may only have a small impact, however patients may value particular characteristics or they may have an impact on wider wellbeing. Alternatively, these characteristics may not have been valued within a QALY framework, but revealed preferences have still been elicited via an alternative method.⁸⁸

Competition also exists between therapies for positioning in a sequence. Manufacturers will look to enhance their profitability, generally by offering long term treatment to as many

* Monoclonal antibodies for infliximab, adalimumab, certolizumab pegol and golimumab, or a circulating receptor fusion protein such as etanercept

patients as possible, which maximises the sales of their particular product. In general this means that the earlier the position in the sequence, the more profitable it is for a manufacturer, due to the length of time available for active therapy, the size of the patient population, and assumption that the patients are likely to have less refractory disease.

For example, if an incumbent treatment exists for a chronic condition, and a new therapy is granted marketing authorisation for use in patients with that condition (with no restriction on position), then the decision by that manufacturer is whether to attempt to compete at first line, or enter as a second (or later) line therapy. If the evidence generated for market authorisation and HTA purposes is focussed on first line use, then the manufacturer may be at risk of an unfavourable decision at second line due to less robust evidence. More likely, NICE would not make a recommendation either way for second line use if not in the original scope or manufacturer's submission, and it would be up to local level decision-making if commissioners wished to fund it.

The market for second line use may be less profitable (smaller and more refractory patient population) but if the entrant therapy is only equivalent (in terms of costs and benefits) compared to the incumbent therapy then it may not capture a significant market share at first line. This example ignores defensive strategies by the incumbent manufacturer. In this example, a me-too therapy may not be appropriate at second line if it has the same mechanism of action, because it may not be effective after the first line therapy. However, if a new therapy can be effective in second line, then it may prove to be a profitable position for that therapy and sequential treatment may expand.

Over the longer term, once patent protection expires for a branded therapy, generic and biosimilar alternatives may be developed and granted a licence. Like me-too therapies, these will increase competition in the market for a condition. Because of the reduced R&D costs for a generic product, costs and prices are generally lower, and competition is increased.⁸⁹

In all these instances, increased competition may lead to sequential therapies because only one therapy can generally be used at a time. There are instances where combination therapies are used, such as concomitant methotrexate in RA. Often combination therapies are used to combine multiple or complementary mechanism of actions,⁹⁰ or to minimise drug resistance.⁹¹

Whether a novel therapy, a me-too, or generic equivalent, new pharmaceuticals will increase the therapeutic options for patients, and are likely to increase the likelihood of sequential therapies being used.

3.5 PATIENT CHOICE AND TREATMENT SEQUENCES

In the UK, the National Health Service (NHS) highlights in the NHS Constitution that ‘choice’ is a clear ideology as a component of universal health care.* Patients have the choice of GP practice, as well as specific health care. In particular, the NHS supports informed choice, with support for patients when making choices, as well as providing information to help people participate in healthcare decisions.

Choice is integral in the treatments of chronic conditions. This is for a number of reasons:

1. Self-management – often the effective management of chronic conditions requires both medical treatment as well as self-management
2. Adherence – By allowing patients to be an active participant in decision making, rather than a passive recipient, adherence to treatment is improved
3. Long-term ownership – For long term chronic conditions, people have an increased risk of mental health issues. Psychological ownership is the term for people finding meaning when they are diagnosed with a chronic illness, and learning to make sense of living with a chronic illness.⁹² As with the above reasons, choice empowers patients, leading to engagement with health services, adherence to medication and the desire to improve their health.

Some of these reasons may have positive implications for cost-effectiveness. Self-management and adherence both ensure that the realised benefits of treatment are maximised. However, sometimes choice may be promoted but the implications for efficiency and cost-effectiveness are less clear. Should NICE promote a choice in treatment when there is no net benefit, or when a relatively small net benefit is offset by guidance development and implementation costs?

Sequential therapy is often driven by the requirement of choice – that an effective alternative will improve the likelihood of controlling a disease, and therefore improve expected health outcomes. However, it is also driven by the desire for choice. Irrespective of potential to provide benefit, the cost, or the number of treatments already attempted, there is a strong belief that a patient requires active treatment. This could be extended to the extreme, where further active treatment may have no additional benefit on top of palliation or supportive care. This concept could potentially be the ‘arrow in the quiver’ argument; that there must always be a choice between alternative options at any point in time. Benefit may be provided by the delivery of care even if the treatment is not effective – the placebo effect.

* <http://www.nhs.uk/choiceintheNHS/Rightsandpledges/NHSConstitution/Pages/Overview.aspx> - Accessed June 2015

3.6 CONCLUSIONS

This chapter has highlighted the clinical reasons for why treatment sequences may be observed - effectiveness of a treatment, uncertainty regarding its effectiveness, patient variability, chronicity of a condition, and the ability to define a switching rule. The chapter also explored how the pharmaceutical market and the desire for healthcare choice may further promote the use of treatment sequences.

This chapter has highlighted several challenges for consideration by a research team when undertaking an economic evaluation of sequential therapies for chronic conditions.

Firstly, when the decision space for an economic evaluation includes treatment sequences, then the factorial rate of growth in the number of comparators becomes unfeasible for decision analytic models to evaluate every possible sequence. Also, the evidence requirements for populating a model with numerous treatment sequences as comparators grow rapidly. Not all treatment sequences are likely to be clinically plausible, and so scoping has to determine all clinically sensible comparator treatment sequences. Some explicit pre judgement has to be applied.

Secondly, when a model becomes more complex, it can lose face validity.⁹³ It is important that decision-makers and clinicians can believe the results of the model, and this can become a challenge when models require numerous assumptions or have a large amount of parameter uncertainty.

The validity of a model also comes at the expense of complexity that is required to provide estimates of the long term costs and effects due to a treatment sequence. The model requires complexity to account for the multiple sources of evidence which need to be incorporated, as well as the chosen methodology to account for uncertainty, and the consideration of alternative decisions such as switching rules.

Finally, the incorporation of sequential therapies in an economic evaluation may become unfeasible for decision-makers and for researchers. The evidence requirements and model development time may not fit in with the timeliness of developing clinical guidance. The evaluation may also require clinical evidence which may not be available when early decisions are made requiring the acceptance or rejection of potential treatments.

3.7 AIMS AND OBJECTIVES

The primary research question for this PhD thesis is ‘How can economic evaluations of sequential therapies for chronic conditions improve health resource allocation decision making?’ The PhD uses RA as a case study.

The aim of this thesis is to test the feasibility of simulation optimisation methods to find an optimal or near-optimal sequence of disease modifying treatments for RA in an economic evaluation framework.

The objectives to meet this aim are:

- To explore the key challenges when undertaking an economic evaluation of sequential therapies in chronic conditions
- To identify any published attempt to estimate the most cost-effective (optimal) sequence of treatments for patients with RA
- To develop a cost-effectiveness model which allows the evaluation of sequential disease modifying therapies for RA
- To review and assess the relative merits of methods of optimising a discrete event simulation model for a combinatorial decision problem when an incremental analysis of all possible alternative treatment sequences is not feasible.
- To implement and evaluate an identified method using RA as a case study condition
- To provide recommendations about the application of the implemented method
- To provide recommendations for further research.

3.8 STRUCTURE OF THE THESIS

The first chapter provided an overview of the motivation for this thesis. Chapter 2 explored the economic framework for Health Technology Assessment (HTA). This set the scene in terms of background theory and the current methodology that are applied for the allocation of finite healthcare resources.

This current chapter provides a rationale for this thesis. It explored why treatment sequences are used, and why they represent a challenge for undertaking an economic evaluation. In particular, it highlights how an incremental analysis is unlikely to be feasible when a large number of sequences are available for comparison.

The next chapter (Chapter 4) reports a systematic review of economic evaluations for disease modifying anti-rheumatic drugs (DMARDs) for RA. The objective of this review is to explore the

existing economic evidence for RA, assess the quality of this evidence, report to what extent they identify an optimal sequence of treatments, and identify where improved methods for economic evaluation may improve decision making in this context.

Chapter 5 reports the development of a discrete event simulation model for the evaluation of treatment sequences in RA. In Chapter 6, a systematic review of relevant simulation optimisation methods is reported, which are subsequently applied to the discrete event simulation model in Chapter 7. The thesis ends with Chapter 8, and a discussion regarding the strengths and limitations of the work undertaken. Recommendations for policy makers and further research are reported, before overall conclusions are drawn.

CHAPTER 4: A SYSTEMATIC REVIEW OF ECONOMIC EVALUATIONS OF DISEASE MODIFYING ANTI-RHEUMATIC DRUGS (DMARDS) FOR RHEUMATOID ARTHRITIS

4.1 CHAPTER OVERVIEW

The previous chapters have introduced the rationale for this thesis, by exploring why sequences of treatments represent a unique challenge when conducting an economic evaluation. This chapter contains a systematic review of economic evaluations of DMARDs for RA. The review determines the extent to which treatment sequences have been evaluated, and the methodological challenges that have arisen when developing a DAM. The review provides important information regarding the structure and parameterisation of a DAM developed in Chapter 5, and used in Chapter 7 as part of the simulation-optimisation analysis.

The scope of the systematic review is purposefully broad, enabling all economic evaluations of DMARDs for RA to be included, irrespective of if they did or did not model a treatment sequence. Only by including the full body of evidence can a conclusion be drawn regarding the extent to which treatment sequences have been explicitly modelled and considered, and only then can an inference be drawn to whether decision-makers are being provided with evidence required to potentially develop an optimal treatment sequence for clinical guidance.

There are many studies identified that did not formally consider the downstream sequence of treatments. While in many ways these studies are less useful for the objectives of the review, it is important to determine, where possible, the justification for why the scope of the economic evaluation was limited.

This systematic review was published in a peer-reviewed journal:

Tosh J, Stevenson M, Akehurst RL. Health Economic Modelling of Treatment Sequences for Rheumatoid Arthritis: A Systematic Review. *Current Rheumatology Reports* (2014). 16:447⁹⁴

4.2 INTRODUCTION

This chapter contains a systematic review of economic evaluations undertaken of DMARDs for RA. In Section 4.3 the methods of this systematic review are described. In particular, the information sources, eligibility criteria, and methods for data extraction and critical appraisal are reported. Section 4.4 presents the results of the search, including the characteristics of

each identified economic evaluation and a critical appraisal using the Drummond checklist.⁷⁸ Also included is a narrative synthesis of the identified studies. This includes the scope of the economic evaluation, the extent to which downstream costs, benefits and sequences are captured, the modelling methods used and the final health economic results of the evaluations. Section 4.5 provides a discussion of the systematic review, and finally Section 4.6 draws conclusions and implications for the rest of the thesis.

OBJECTIVES

The aim of this systematic review is to summarise the existing economic evidence for the use of DMARDs in the treatment of RA. The systematic review will assess the strengths and limitations of specific economic evaluations which compare DMARDs, and will draw generalised conclusions regarding the methodologies currently used to evaluate treatments for RA.

Specifically, the objectives of the review are:

1. To identify the existing economic evidence for disease modifying therapies for RA, and to assess where there are gaps in the existing evidence base.
2. To assess the health economic evaluation studies, with respect to their objectives and the methods used to meet these objectives.
3. To identify how improvement in the methods of economic evaluation may improve decision making regarding the treatment of people with RA.

The systematic review is reported to the PRISMA standards.⁹⁵ A completed PRISMA checklist can be found in Appendix A.1. The PRISMA criteria are applicable to clinical systematic reviews, and so some elements of the checklist are not directly applicable to this review of economic evaluations. However, the PRISMA standard provides a useful checklist of items to ensure all relevant details pertaining to the review are reported. The review has been registered with the International Prospective Register of Systematic Reviews (PROSPERO).

EXISTING SYSTEMATIC REVIEWS

Twenty-four existing systematic reviews of economic evaluations undertaken of disease modifying therapies for RA have been identified by the searches detailed in Section 4.3.⁹⁶⁻¹¹⁹ These reviews synthesised the general economic evidence, however the focus varied: some focussed on particular treatments, some focussed on particular patient populations, and some on particular aspects of the methodology and evidence used. For example, Bansback *et al.* (2008) focussed on HRQL in RA economic evaluations,⁹⁸ and Emery (2004) focussed on DAM methods.¹⁰³ However, none of the reviews focussed explicitly on the methods used to identify

and evaluate sequential therapies for patients with RA, although Sullivan *et al.* (2013) consider the economic impact of sequential DMARD treatment.¹²⁰ Therefore this review will add to the existing evidence base.

4.3 METHODS

ELIGIBILITY CRITERIA

Systematic searches of electronic databases were conducted to identify all published economic evaluations of DMARDs for RA. Because DMARDs for RA are numerous, using conventional methods for a systematic review would require the formulation of numerous search terms for each particular DMARD. Because of the significant number of interventions, and therefore search terms, a conventional search method could miss relevant interventions and economic evaluations. To ensure that the systematic search had high sensitivity (the identification of appropriate studies), a search strategy was developed by applying economics related terms to a set of clinical terms covering RA and DMARDs. The disease component of the electronic search was based on a previously used electronic search strategy for the NICE RA guideline.¹²¹ Database filters to identify economic evaluations were used from the InterTASC Information Specialists' Sub-Group (ISSG) website.*

A scoping search was undertaken using Google Scholar[®] to identify keywords for the search strategies. These keywords are listed in Table 4.1 and reported in the PICO (Population, Intervention, Comparator, Outcome) format.

Table 4.1: Keywords for systematic review (PICO format)

Population	Rheumatoid Arthritis, RA
Intervention/Comparator	Disease modifying, disease-modifying, DMARD, biologic, therapy, treatment, anti-rheumatic, anti rheumatic, TNF, tumor necrosis factor alpha, tumour necrosis factor alpha, TNF-alpha, TNF inhibitor, TNF blocker, interleukin 1, IL-1, monoclonal antibody, costimulation blocker, interleukin 6, IL-6
Outcomes	Economic, economics, cost, cost-effectiveness, cost-utility, cost-benefit, utility, health related quality of life, quality of life, quality adjusted life year, QALY

The search strategies used MeSH terms, including 'rheumatoid arthritis' and 'economics', and text string terms which were combined in the search strategy using Boolean logic. The search

* <http://www.york.ac.uk/inst/crd/intertasc/index.htm> - Accessed June 2015

strategies for all databases are provided in Appendix A.2. The search strategies were designed to maximise sensitivity, however this was at the cost of poor specificity (the rejection of inappropriate studies). This meant the search returned a large number of inappropriate studies, and therefore the review required extensive sifting of the results to filter out the appropriate studies.

INFORMATION SOURCES

Systematic searches were conducted in ten databases (Table 4.2). Conference abstracts were not included, however publications by the authors of any included studies were searched to identify any later publications. Reference and citation searching was undertaken on all included studies, including any identified reviews of published economic evaluations of DMARDs for RA. Published NICE Technology Appraisals and NICE Clinical Guidelines were searched to identify any studies not detected in the electronic searches.

Table 4.2: Systematic review databases

Database	Date *
BIOSIS (all databases)	1899 – Feb 2013
Cochrane Database of Systematic Reviews (CDSR)	All years – Feb 2013
Cochrane Database of Methodological Reviews	All years – Feb 2013
Cochrane Central Register of Controlled Trials (CCRCT)	All years – Feb 2013
Database of Abstracts of Reviews and Effects (DARE)	All years – Feb 2013
Cumulative Index to Nursing and Allied Health Literature (CINAHL)	1994 – Feb 2013
Embase	1974 – Feb 2013
MEDLINE	1945 – Feb 2013
NHS Economic Evaluations Database (NHSEED)	All years – Feb 2013
Science Citation Index: Web of Science	1899 – Feb 2013

All database searches were undertaken on 1st February 2013, and no date restriction was applied. No study type or language restrictions were applied to the electronic search. The search strategies were reviewed by an information specialist.

INCLUSION AND EXCLUSION CRITERIA

The primary objective of the systematic search is to identify any economic evaluations of DMARDs for RA. The search was irrespective of any decision-making context or geographical location. The eligibility criteria for the systematic review are presented in Table 4.3.

Studies were included in the review if they reported a comparative economic evaluation (cost-effectiveness (CEA), cost-utility (CUA) or cost-benefit analysis (CBA)). Cost-consequence

* "All years" when a formal start date for the Database is not provided.

analyses (CCA) were also included, because the evidence provided in the study can be used comparatively even if the comparison has not been made by the authors. Cost-minimisation analyses (CMA) were also included. However, economic evaluations where a full comparison between two or more alternatives was not conducted were excluded. This is because they do not consider an estimate of incremental costs or benefits between treatments. However, excluding these (likely rare) studies may ignore potentially relevant studies where the costs or benefits of sequential therapies have been estimated. This is a potential limitation of the study and will be discussed subsequently.

Table 4.3: Eligibility criteria

Inclusion Criteria
<ul style="list-style-type: none"> • Economic evaluation including a comparison of costs and health benefits based on outcomes data or undertaken using a decision analytic model • Economic evaluations of interventions targeting a change to the natural disease profile of people with rheumatoid arthritis (i.e. disease-modifying therapies) • Studies reporting costs and health outcomes
Exclusion Criteria
<ul style="list-style-type: none"> • Non-comparative/partial economic evaluations • Cost analyses/Cost-of-illness/Burden-of-illness studies • Methodological papers which do not report economic and health benefit outcomes • Commentaries, letters, editorials • Conference abstracts • Studies which claim cost-effectiveness but with no estimation of costs and effectiveness outcomes • Economic evaluations of therapies and treatments which do not modify the progression of RA • Non-English language studies

CRITICAL APPRAISAL AND DATA EXTRACTION

The identified studies were appraised using the validated and commonly used Drummond (2005) 'Critical appraisal of a published article' checklist.⁷⁸ This is a brief checklist and was used to appraise the key aspects of each economic evaluation.

While the checklist aided the extraction of data and appraisal of the quality of studies, it was not appropriate to use just a checklist for this particular review. Many systematic reviews of economic evaluations attempt to assess the quality and appropriateness of studies which are addressing a similar decision problem, and checklists are designed to aid this. For example, a systematic review of early treatments in patients with newly-diagnosed RA may be undertaken to appraise the relevant published cost-effectiveness evidence in that particular population,

and for that particular decision problem. Instead, the objective for this systematic review is to examine where studies have addressed or not addressed the sequential aspect of treatment for people with RA. Therefore data pertaining to treatment sequences and modelling methodology were extracted. Where studies were similar, themes were extracted to generalise the approach taken. Where studies were different but the decision context was apparently similar (e.g. the patient population or comparator treatments) then the difference in methodology was noted. The full data extraction template is provided in Table 4.4.

The data extracted from identified studies included general details regarding the economic evaluation (method, patient population, comparator interventions or sequences, time horizon, disease and treatment history, and health economic results). Data were also extracted regarding the sequential nature of the treatments for RA and how this was captured in the analysis. In particular, whether all relevant comparators were included in the sequence of treatments. Where relevant, information regarding the modelling methods used for the economic evaluation were extracted, to identify how data sources were used in reflecting the costs and benefits of particular treatments, at all points in the evaluated sequence. The principal summary measures for the review were the study name and year, country, evaluated interventions, time horizon of the study, the type of economic evaluation, the type of decision-analytic model used (if appropriate), and the basecase health economic results.

Table 4.4: Data extraction template

Evaluation information				
Study (name and year)				
Country				
Evaluation type (CEA, CUA, CCA, CMA, CBA)				
Comparators				
Previous treatment history				
Time horizon of analysis				
Analysis method (Trial evaluation, model)				
Basecase results				
Uncertainty analysis				
Conclusions				
Drummond (2005) Checklist⁷⁸				
Question	Yes	No	Not sure	Additional information
Was a well-defined question posed in an answerable form?				
Was a comprehensive description of the competing alternatives given?				
Was there evidence that the programme's effectiveness had been established?				
Were all important and relevant costs and consequences for each alternative identified?				
Were costs and consequences measured accurately in appropriate physical units?				
Were costs and consequences valued credibly?				
Were costs and consequences adjusted for differential timing?				
Was an incremental analysis of costs and consequences of alternatives performed?				
Was allowance made for the uncertainty in the estimates of costs and consequences?				
Did the presentation and discussion of study results include all issues of concern to users?				
Sequencing information				
If a sequence was evaluated, how many lines?				
Were all relevant treatments included in each sequence?				
Was the methodology used amenable to varying the sequence and comparing a full range of alternatives?				
Was an attempt to find an 'optimal' sequence undertaken?				
Modelling information				
If the analysis was undertaken using a decision analytic method, what method was used?				
How was initial treatment response modelled?				
What determined a switch to an alternative therapy?				
How were the costs and effectiveness of subsequent treatments in a sequence modelled?				
Where data were not available for a treatment in a sequence, how was this accounted for?				

4.4 RESULTS

STUDIES IDENTIFIED

From the systematic searching of electronic databases, 8,281 citations were identified (Quorum flow-diagram provided in Figure 4.1). After excluding 3,250 duplicate citations electronically, the remaining 5,031 citations were screened by their abstract. Of these, 4,913 abstracts did not meet the inclusion criteria and 118 full papers were retrieved for a full inspection. A total of 70 papers were excluded for not meeting the inclusion criteria, and 9 other papers were identified by reference and citation searches and searching any identified systematic reviews. 57 published studies were included in the systematic review. The full papers that were excluded, and the reason for exclusion, are provided in Appendix A.3. The data extraction tables and Drummond checklist results are provided in Appendix A.4 and Appendix A.5, respectively.

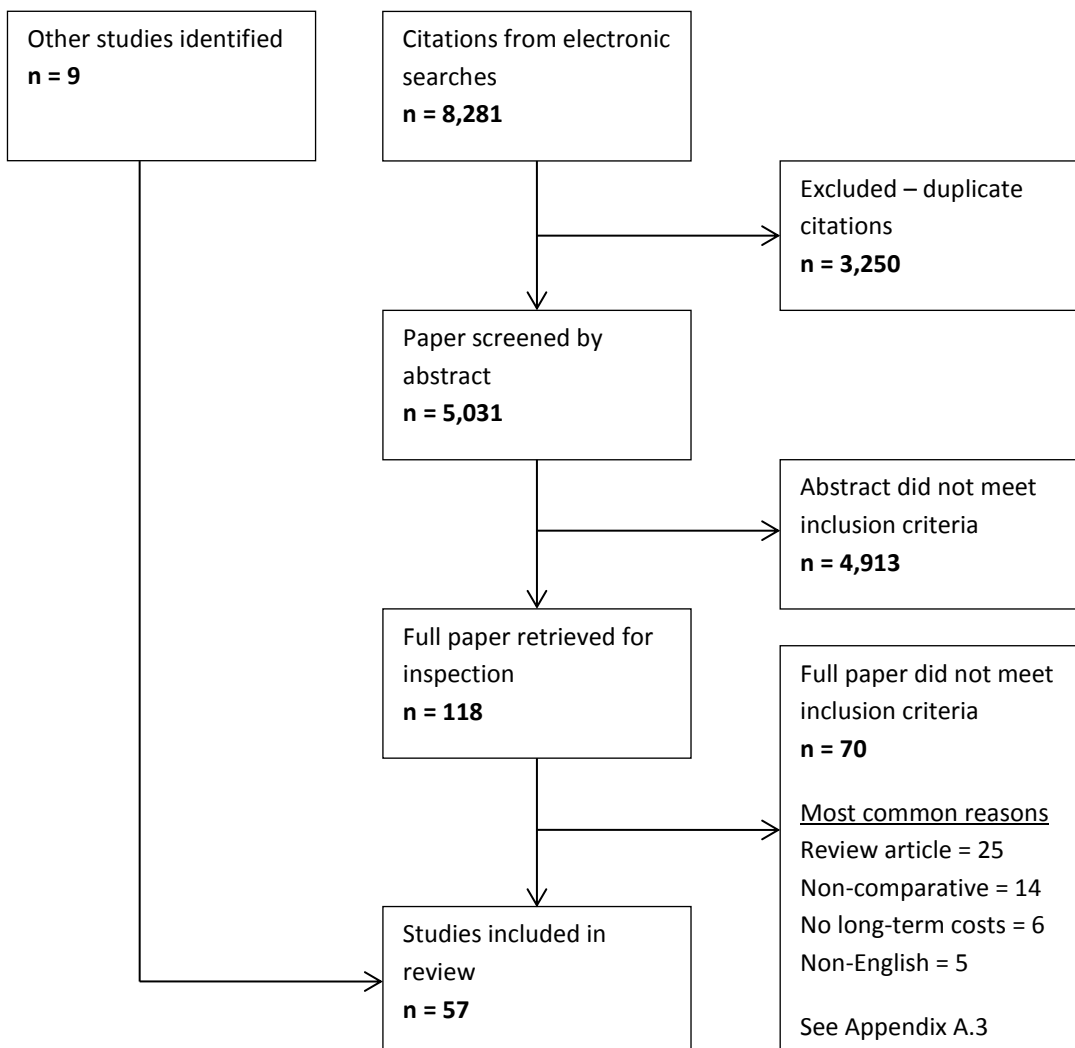


Figure 4.1: Quorum flow diagram

CHARACTERISTICS OF THE ECONOMIC EVALUATIONS

Of the 57 included studies, 43 (75%) were CUAs with QALYs as the unit of health benefit. Nine (16%) CEAs were conducted, with three (5%) CCAs, and two (4%) CMAs. Eleven (19%) studies were conducted with a UK perspective, and 11 (19%) with a US perspective. The remainder of the studies are mainly European (26 (46%)), along with six (11%) studies from Canada, and one (2%) study each from India, Japan and Thailand. The results reinforce the belief that QALYs are a common generic metric for health benefits when undertaking a CEA,^{122,123} and while CEAs are undertaken world-wide, they are more common in developed countries. The studies are summarised in Table 4.5.

Table 4.5: Summary of reviewed studies (Table also published in Tosh *et al.* (2014)⁹⁴)

Recent onset RA						
Study, year	Country	Interventions	Time horizon	Type	Model type	Incremental Cost-effectiveness Ratio (ICER)
Chen <i>et al.</i> 2006 ¹²⁴	UK	TNFa with or without MTX at first line or third line	Lifetime	CUA	Individual Level Model	ICERs for ETN, ADA and IFX after multiple cDMARD failure were £24k, £30k and £38k per QALY, respectively
Davies <i>et al.</i> 2009 ¹²⁵	US	MTX vs. ADA+MTX vs. ETN vs. IFX+MTX vs. ADA+MTX	Lifetime	CUA	Individual Level Model	IFX and ETN extendedly dominated by ADA. ADA+MTX ICER \$47k per QALY vs. cDMARDs. ADA+MTX then ETN ICER \$42k per QALY vs. cDMARDs
Finckh <i>et al.</i> 2009 ¹²⁶	US	Symptomatic therapy vs. MTX vs. bDMARDs	Lifetime	CUA	Individual Level Model	bDMARDs dominated by cDMARDs. cDMARDs ICER \$4k per QALY vs. symptomatic therapy
Hartman <i>et al.</i> 2004 ¹²⁷	NL	Placebo vs. folic acid vs. folinic acid. Adjunct to MTX	48 weeks	CUA	Trial analysis	Placebo dominates folic acid. Folinic acid dominates placebo
Kobelt <i>et al.</i> 2002 ¹²⁸	UK	MTX vs. SSZ vs. LEF	10 year	CUA	Markov model	Using Strand <i>et al.</i> , LEF dominates MTX. Using Emery <i>et al.</i> , MTX dominates LEF. Using Smolen <i>et al.</i> , LEF dominates SSZ.
Kobelt <i>et al.</i> 2011 ¹²⁹	Sweden	ETN+MTX vs. MTX	10 year	CUA	Markov model	ETN+MTX ICER is €13k per QALY vs. MTX
Korthals-de Bos <i>et al.</i> 2004 ¹³⁰	NL	MTX+SSZ+Prednisolone vs. SSZ	56 weeks	CUA	n/a	Combo cDMARDs dominates SSZ
Maetzel <i>et al.</i> 2002 ¹³¹	Canada	Adding LEF to a cDMARD sequence	5 year	CUA	Decision tree	Adding LEF ICER is Can\$71k per QALY vs. cDMARD sequence
Schadlich <i>et al.</i> 2005 ¹³²	Germany	Adding LEF to cDMARD sequences	3 years	CUA	Decision tree	ICER of adding LEF vs. cDMARD sequence is €8k per QALY
Schipper <i>et al.</i> 2011 ¹³³	NL	Sequential TNFa use	5 years	CUA	Markov model	ICER TNFa €138k per QALY vs. MTX. ICER MTX+LEF €439k per QALY vs. MTX
Spalding <i>et al.</i> 2006 ¹³⁴	US	MTX vs. bDMARD mono and combos	Lifetime	CUA	Markov model	ICERs ranged \$63k per QALY for ADA vs. MTX to \$409k per QALY for IFX vs. MTX.
Tosh <i>et al.</i> 2011 ¹³⁵	UK	Alternative cDMARD mono and combo therapies	Lifetime	CUA	Individual Level Model	Mono, Step-up, Parallel, Steroid are all dominated by step-down. Intensive ICER £27k per QALY vs. step-down
van den Hout <i>et al.</i> 2009 ¹³⁶	NL	Comparing cDMARD combos vs. IFX combo therapy	2 year	CUA	Trial analysis	Initial combo therapy with prednisone is likely to be the most cost-effective strategy at a WTP per QALY of <€100k
Verhoeven <i>et al.</i> 1998 ¹³⁷	NL	Step-down cDMARDs vs. SSZ	1 year	CUA	n/a	Combo. cDMARDs dominates SSZ

Established RA						
Study, year	Country	Interventions	Time horizon	Type	Model type	Incremental Cost-effectiveness Ratio (ICER)
Anis <i>et al.</i> 1996 ¹³⁸	Canada	CYA vs. AZA/PEN vs. placebo	1 year	CEA	Decision tree	CYA ICER \$11k per patient improved vs. placebo.
Bansback <i>et al.</i> 2005 ¹³⁹	Sweden	TNFa with or without MTX vs. cDMARDs	Lifetime	CUA	Individual level Markov model	For all TNFa strategies, ICERs using ACR50 response criteria are between €34k per QALY and €42k per QALY vs. cDMARDs. ADA+MTX likely to be the optimal strategy
Barbieri <i>et al.</i> 2005 ¹⁴⁰	UK	IFX+MTX vs. MTX	1 year and lifetime	CUA	Markov model	IFX-MTX ICER is £33k per QALY vs. MTX
Barton <i>et al.</i> 2004 ¹⁴¹	UK	ETN vs. IFX vs. cDMARD sequence	Lifetime	CUA	Individual Level Model	ETN ICER £50k per QALY vs. basecase. IFX ICER £68k per QALY vs. basecase. ETN ICER £28k per QALY vs. IFX
Benucci <i>et al.</i> 2009 ¹⁴²	Italy	ABT with LEF or MTX vs. ETN with LEF or MTX	2 years	CUA	Observational analysis	ETN+MTX had the lowest CER compared to baseline (non bDMARD tx) - €39k per QALY.
Benucci <i>et al.</i> 2011 ¹⁴³	Italy	RTX vs. constant disease	6 months, 1 year	CUA	Observational analysis	RTX ICER €15k per QALY vs. consistent disease comparator (6 months). ICER €23k in 1 year
Beresniak <i>et al.</i> 2011 ¹⁴⁴	Spain	ADA vs. IFX vs. ABT vs. RTX	2 years	CEA	Unclear	Highest effectiveness and lowest CER for ABT. LDAS and RS outcomes
Brennan <i>et al.</i> 2004 ¹⁴⁵	UK	ETN vs. cDMARD sequence	Lifetime	CUA	Individual Level Model	ETN ICER £16k per QALY vs. cDMARDs
Brennan <i>et al.</i> 2007 ¹⁴⁶	UK	TNFa vs. cDMARDs	Lifetime	CUA	Individual Level Model	TNFa ICER is £23k per QALY vs. cDMARDs
Chiou <i>et al.</i> 2004 ¹⁴⁷	US	ANA vs. ETN vs. ADA vs. IFX	1 year	CUA	Decision tree	ETN ICER \$7k per QALY vs. ANA. ADA and IFX dominated by ETN
Choi <i>et al.</i> 2000 ¹⁴⁸	US	cDMARD mono and combo vs. bDMARD mono and combo	6 months	CEA	Decision tree	ETN ICER \$42k per ACR20 responder vs. triple cDMARD therapy.
Choi <i>et al.</i> 2002 ¹⁴⁹	US	cDMARD mono and combo vs. bDMARD mono and combo	6 months	CEA	Decision tree	ETN ICER \$41k per ACR20 responder vs. MTX
Cimmino <i>et al.</i> 2011 ¹⁵⁰	Italy	ABT vs. ADA vs. RTX vs. IFX	2 years	CEA	Unclear	Highest effectiveness and lowest CER for ABT. LDAS and RS outcomes
Clark <i>et al.</i> 2004 ¹⁵¹	UK	Adding ANA in a treatment sequence	Lifetime	CUA	Individual Level Model	ANA ICER over £100k per QALY vs. standard care
Coyle <i>et al.</i> 2006 ¹⁵²	Canada	GLD vs. bDMARD mono and combo	5 years	CUA	Markov model	IFX and ETN had ICERS over \$100k per QALY vs. cDMARDs

Diamantopoulos <i>et al.</i> 2012 ¹⁵³	Italy	Sequential bDMARD use	lifetime	CUA	Individual Level Model	TCZ dominates replacing ETN or ADA. TCZ ICER €2k per QALY vs. IFX. TCZ ICER €17k when added first line.
Hallinen <i>et al.</i> 2010 ¹⁵⁴	Finland	Sequential bDMARD use	Lifetime	CUA	Individual level Markov model	RTX dominates ADA, ABT, ETN after TNFa failure. RTX ICER €30k per QALY vs. BSC.
Jobanputra <i>et al.</i> 2002 ¹⁵⁵	UK	Adding ETN and IFX into a cDMARD sequence	Lifetime	CUA	Individual Level Model	ETN ICER £83k per QALY vs. basecase. IFX ICER £115k per QALY vs. basecase. ETN ICER £44k per QALY vs. IFX.
Kavanaugh <i>et al.</i> 1996 ¹⁵⁶	US	GLD vs. MTX vs. bDMARDs	6 months	CCA	Decision tree	Efficacy reflected as costs. GLD = \$6k, MTX = \$5k, bDMARDs = \$9k
Kielhorn <i>et al.</i> 2008 ¹⁵⁷	UK	RTX+MTX vs. cDMARD sequence	lifetime	CUA	Individual level Markov model	RTX ICER £11k per QALY vs. cDMARDs. With no sequential bDMARD us, RTX ICER £14 per QALY vs. cDMARDs.
Kievit <i>et al.</i> 2009 ¹⁵⁸	NL	Comparing treatment guidelines	6 months	CCA	Trial analysis	All strategies had an equal cost. All variations to guideline generated more responders.
Kobelt <i>et al.</i> 2003 ¹⁵⁹	Sweden, UK	IFX+MTX vs. MTX	10 year	CUA	Markov model	IFX ICER is €3k per QALY vs. MTX in Sweden. £21k per QALY vs. MTX in UK
Kobelt <i>et al.</i> 2004 ¹⁶⁰	Sweden	TNFa vs. cDMARDs	1 year	CUA	Trial analysis	TNFa ICER is €43k per QALY vs. previous years' therapy
Kobelt <i>et al.</i> 2005 ¹⁶¹	Sweden	ETN vs. MTX vs. ETN+MTX	2 year/ 10 year	CUA	Markov model	ETN+MTX ICER is €37k per QALY vs. MTX (2 year horizon). ETN+MTX ICER is €46k per QALY vs. MTX (109 year horizon)
Lekander <i>et al.</i> 2010 ¹⁶²	Sweden	IFX vs. cDMARDs	20 year	CUA	Markov model	IFX ICER €22k per QALY vs. cDMARDs
Lindgren <i>et al.</i> 2009 ¹⁶³	Sweden	RTX vs. TNFa	lifetime	CUA	Discrete Event Simulation	RTX dominates TNFa
Maetzel <i>et al.</i> 2002a ¹⁶⁴	Canada	LEF vs. MTX vs. placebo	1 year	CCA	n/a	MTX dominates LEF and placebo
Malottki <i>et al.</i> 2011 ¹⁶⁵	UK	ADa vs. ETN vs. IFX vs. RTX vs. ABT vs. cDMARD sequence	Lifetime	CUA	Individual Level Model	RTX dominates ADA, ETN and IFX ABT ICER £130k per QALY vs. RTX
Marra <i>et al.</i> 2007 ¹⁶⁶	Canada	IFX+MTX vs. MTX	10 years	CUA	Markov model	IFX ICER between \$Can32k-70k per QALY vs. MTX.
Merkesdal <i>et al.</i> 2010 ¹⁶⁷	Germany	Adding RTX+MTX to a sequence	Lifetime	CUA	Individual level Markov model	RTX ICER €24k per QALY vs. TNFa
Nuijten <i>et al.</i> 2001 ¹⁶⁸	NL	ETN vs. IFX	1 year	CMA	Unclear	ETN dominates IFX
Osiri <i>et al.</i> 2007 ¹⁶⁹	Thailand	Comparing cDMARD strategies	1 year	CEA	n/a	MTX = \$2k (per 1 point HAQ change vs. AM). MTX + AM = dominates. MTX + SSZ = \$625. AM + SSZ = \$14k. AM + MTX + SSZ = \$1k. LEF = \$1k. Other DMARDS = \$16k

Rubio-Terrés <i>et al.</i> 2001 ¹⁷⁰	Spain	IFX+MTX vs. LEF	1 year	CMA	Unclear	LEF dominates IFX+MTX in the CMA
Russell <i>et al.</i> 2009 ¹⁷¹	Canada	Sequential TNFa use	2 years	CEA	Decision tree	1st bDMARD position: ABT dominates when using both remission and LDAS as outcomes. 2nd bDMARD position: ICER \$20k per LDAS and \$26k per remission
Saraux <i>et al.</i> 2010 ¹⁷²	France	Sequential TNFa use	2 year	CEA	Unclear	Lower costs per 'theoretical expected number of days in remission' with ABT after first TNFa compared with RTX. Consistent with remission criteria as well.
Shini <i>et al.</i> 2010 ¹⁷³	India	cDMARD mono and combo therapies	3 months	CEA	n/a	For mono, lowest CER was HCQ. For combo, lowest CER was MTX+HCQ
Soini <i>et al.</i> 2012 ¹⁷⁴	Finland	ADA vs. ETN vs. TCZ	Lifetime	CUA	Individual Level Model	Wholesale prices: TCZ dominates ADA and ETN and ICER €18k per QALY vs. MTX. Retail prices TCZ extendedly dominates ADA, ETN and ICER €17k per QALY vs. MTX.
Tanno <i>et al.</i> 2006 ¹⁷⁵	Japan	Adding ETN to a cDMARD sequence	Lifetime	CUA	Markov model	ETN ICER ¥3.5 per QALY vs. standard therapy
Vera-Llonch <i>et al.</i> 2008 ¹⁷⁶	US	ABT vs. cDMARDs	lifetime	CUA	Individual Level Model	ABT ICER \$45k per QALY vs. cDMARDs
Vera-Llonch <i>et al.</i> 2008a ¹⁷⁷	US	ABT+MTX vs. MTX	lifetime	CUA	Microsimulation	ABT+MTX ICER \$43k per QALY vs. MTX
Wailoo <i>et al.</i> 2008 ¹⁷⁸	US	ETA vs. ADA vs. ANA vs. IFX	Lifetime	CUA	Individual Level Model	ANA was the least effective and least costly strategy. ETN, IFX and ADA were similar in terms of effectiveness but IFX was more costly.
Welsing <i>et al.</i> 2004 ¹⁷⁹	NL	Usual care vs. LEF vs. TNFa vs. LEF, TNFa sequences	5 years	CUA	Markov model	Post-DMARD failure most cost effective position for TNFa, with ICER of €163k per QALY vs. usual care.
Wong <i>et al.</i> 2002 ¹⁸⁰	US	IFX+MTX vs. MTX	Lifetime	CUA	Markov model	IFX ICER is £30k per QALY vs. MTX

See glossary for definitions of abbreviations. NL = Netherlands, US = United States. Mono = Monotherapy, Combo = Combination therapy. LDAS = Low disease activity state. RS = Remission state

PATIENT POPULATION AND TREATMENTS

To categorise the patient population and decision point for which a new treatment is being considered, a conceptualisation of RA is required.

The current NICE Clinical Guideline for RA (CG90) defines one subset of the patient population as ‘recent-onset’ – where patients have active RA with no prior use of DMARDs.¹²¹ It states that the objective for health care professionals is to diagnose active RA as soon as possible so that DMARD therapy can be initiated within 3 months of the onset of persistent symptoms. Once patients have received DMARD therapy, then the guideline categorises patients as having ‘established’ disease. The studies will be categorised into these two broad groups, with subgroups defined where appropriate.

The NICE Clinical Guideline (CG90) provides a treatment algorithm to determine the appropriate strategy of care for a patient in either population (Figure 4.2). The development of the NICE Clinical Guideline was constrained by the fact that NICE Technology Appraisals had already developed guidance for the use of bDMARDs in patients with RA,¹⁸¹ and therefore this guidance was mandatory and the advisory remit of the Clinical Guideline could not alter this.¹⁸²

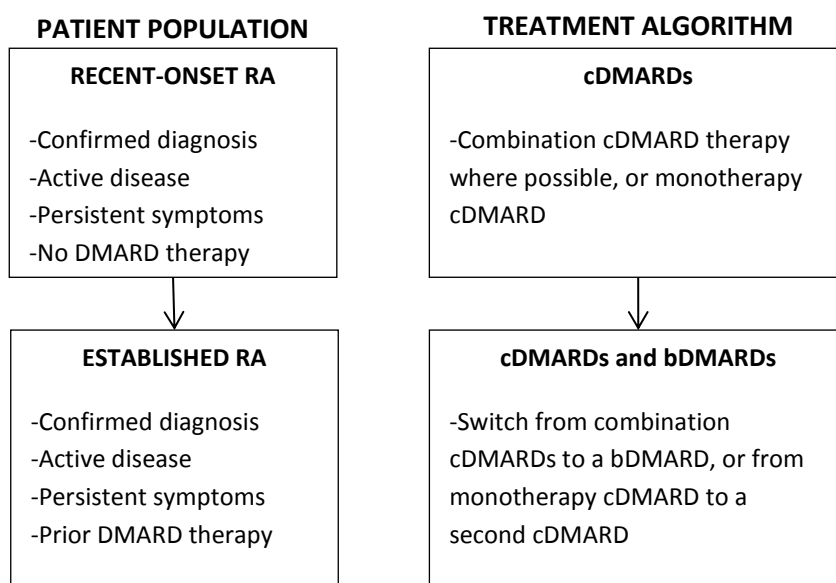


Figure 4.2: NICE Clinical Guideline (CG90) - Patient population and treatment algorithm (simplified)

The treatment algorithm highlights that in general the view of decision-makers has been that bDMARDs, due to their cost, should only be attempted when two less costly cDMARDs have been attempted first. The NICE Clinical Guideline commissioned a *de novo* CUA to compare different combination therapy and monotherapy cDMARD strategies in patients with recent-onset RA.¹³⁵ This CUA informed the treatment algorithm which recommends patients with

recent-onset RA receive combination cDMARD therapy as their first strategy. The downstream impact is that non-responders to combination cDMARDs move more quickly to bDMARDs when compared to a trial of cDMARDs in sequence.

For this systematic review, the patient population for each economic evaluation will be determined as either recent-onset or established RA. The position of the decision point in the economic evaluation will make reference to the standard treatment algorithm, where a recent onset population has no prior DMARD treatment, and established RA patients will have previously had DMARD treatment. If bDMARDs are evaluated for the use in treating recent onset RA patients, then this will be clearly reported.

The patient population for each economic evaluation is determined by the point in which the comparison between alternatives is made. Some researchers call this the divergence point,¹²⁴ because a DAM may explicitly model the previous treatment sequence from recent-onset RA, however the sequence only differs at the divergence point. For example a model may have an identical recent-onset RA sequence of cDMARDs, and then the divergence point occurs where there is a comparison of alternative bDMARDs in patients with established RA. Modelling upstream treatments (before the divergence point) may be undertaken to allow movement of the divergence point, allow alternative patient populations to be generated, or to allow screening or diagnostics to be evaluated.

There are fourteen (25%) studies of DMARD therapy in patients with recent-onset RA.¹²⁴⁻¹³⁷ 42 (74%) of studies are in patients with established RA.^{138-155,157-180} In one study (1%), it was unclear whether the interventions compared were for recent-onset or established RA.¹⁵⁶ The study references are provided in Table 4.6.

Table 4.6: Review patient population

Patient population	Studies	References
Recent - onset RA	14 (25%)	Chen <i>et al.</i> 2006 ¹²⁴ , Davies <i>et al.</i> 2009 ¹²⁵ , Finckh <i>et al.</i> 2009 ¹²⁶ , Hartman <i>et al.</i> 2004 ¹²⁷ , Kobelt <i>et al.</i> 2002 ¹²⁸ , Kobelt <i>et al.</i> 2011 ¹²⁹ , Korthals-de Bos <i>et al.</i> 2004 ¹³⁰ , Maetzel <i>et al.</i> 2002 ¹³¹ , Schadlich <i>et al.</i> 2005 ¹³² , Schipper <i>et al.</i> 2011 ¹³³ , Spalding <i>et al.</i> 2006 ¹³⁴ , Tosh <i>et al.</i> 2011 ¹³⁵ , van den Hout <i>et al.</i> 2009 ¹³⁶ , Verhoeven <i>et al.</i> 1998 ¹³⁷
Established RA	42(74%)	Anis <i>et al.</i> 1996 ¹³⁸ , Bansback <i>et al.</i> 2005 ¹³⁹ , Barbieri <i>et al.</i> 2005 ¹⁴⁰ , Barton <i>et al.</i> 2004 ¹⁴¹ , Benucci <i>et al.</i> 2009 ¹⁴² , Benucci <i>et al.</i> 2011 ¹⁴³ , Beresniak <i>et al.</i> 2011 ¹⁴⁴ , Brennan <i>et al.</i> 2004 ¹⁴⁵ , Brennan <i>et al.</i> 2007 ¹⁴⁶ , Chiou <i>et al.</i> 2004 ¹⁴⁷ , Choi <i>et al.</i> 2000 ¹⁴⁸ , Choi <i>et al.</i> 2002 ¹⁴⁹ , Cimmino <i>et al.</i> 2011 ¹⁵⁰ , Clark <i>et al.</i> 2004 ¹⁵¹ , Coyle <i>et al.</i> 2006 ¹⁵² , Diamantopoulos <i>et al.</i> 2012 ¹⁵³ , Hallinen <i>et al.</i> 2010 ¹⁵⁴ , Jobanputra <i>et al.</i> 2002 ¹⁵⁵ , Kielhorn <i>et al.</i> 2008 ¹⁵⁷ , Kievit <i>et al.</i> 2009 ¹⁵⁸ , Kobelt <i>et al.</i> 2003 ¹⁵⁹ , Kobelt <i>et al.</i> 2004 ¹⁶⁰ , Kobelt <i>et al.</i> 2005 ¹⁶¹ , Lekander <i>et al.</i> 2010 ¹⁶² , Lindgren <i>et al.</i> 2009 ¹⁶³ , Maetzel <i>et al.</i> 2002a ¹⁶⁴ , Malottki <i>et al.</i> 2011 ¹⁶⁵ , Marra <i>et al.</i> 2007 ¹⁶⁶ , Merkesdal <i>et al.</i> 2010 ¹⁶⁷ , Nuijten <i>et al.</i> 2001 ¹⁶⁸ , Osiri <i>et al.</i> 2007 ¹⁶⁹ , Rubio-Terrés <i>et al.</i> 2001 ¹⁷⁰ , Russell <i>et al.</i> 2009 ¹⁷¹ , Saraux <i>et al.</i> 2010 ¹⁷² , Shini <i>et al.</i> 2010 ¹⁷³ , Soini <i>et al.</i> 2012 ¹⁷⁴ , Tanno <i>et al.</i> 2006 ¹⁷⁵ , Vera-Llonch <i>et al.</i> 2008 ¹⁷⁶ , Vera-Llonch <i>et al.</i> 2008a ¹⁷⁷ , Wailoo <i>et al.</i> 2008 ¹⁷⁸ , Welsing <i>et al.</i> 2004 ¹⁷⁹ , Wong <i>et al.</i> 2002 ¹⁸⁰
Unclear	1 (1%)	Kavanaugh <i>et al.</i> 1996 ¹⁵⁶
Total	57 (100%)	

The studies cover a range of monotherapy and combination therapies utilising both cDMARDs and bDMARDs for patients with recent-onset RA. These are summarised in Table 4.7 and Table 4.8.

CRITICAL APPRAISAL OF THE ECONOMIC EVIDENCE FOR DMARDS IN RECENT-ONSET RA

Presented here is a critical appraisal of the economic evaluations conducted which compared treatment strategies for patients with recent-onset RA. Fourteen studies were identified in this population.^{124–137}

i) Scope of the economic evaluations in recent-onset RA

The summary details for all fourteen recent-onset RA economic evaluations are presented in Table 4.5. All fourteen studies were CUAs, with benefits quantified using QALYs. Ten (71%) of the studies considered the introduction of a particular DMARD to a treatment pathway,^{124,125,127,133,134,137} and four (29%) studies evaluated the adjustment or tapering of a treatment strategy.^{126,131,135,136} The specific treatments evaluated for people with recent-onset RA in the identified studies are reported in Table 4.7.

Table 4.7: Treatments evaluated for recent-onset RA

Type	Specific treatment	Studies	References
Monotherapy cDMARDs	LEF	2	Kobelt <i>et al.</i> 2002 ¹²⁸ , Schadlich <i>et al.</i> 2005 ¹³²
	MTX	7	Davies <i>et al.</i> 2009 ¹²⁵ , Finckh <i>et al.</i> 2009 ¹²⁶ , Kobelt <i>et al.</i> 2002 ¹²⁸ , Kobelt <i>et al.</i> 2011 ¹²⁹ , Schipper <i>et al.</i> 2011 ¹³³ , Spalding <i>et al.</i> 2006 ¹³⁴ , Tosh <i>et al.</i> 2011 ¹³⁵
	SSZ	2	Kobelt <i>et al.</i> 2002 ¹²⁸ , Korthals-de Bos <i>et al.</i> 2004 ¹³⁰
Combination cDMARDs	MTX+LEF	1	Schipper <i>et al.</i> 2011 ¹³³
	MTX+SSZ	1	Tosh <i>et al.</i> 2011 ¹³⁵
	MTX+Prednisone	1	Tosh <i>et al.</i> 2011 ¹³⁵
	MTX+SSZ+Prednisone	1	Korthals-de Bos <i>et al.</i> 2004 ¹³⁰
	Step-up combination	1	Tosh <i>et al.</i> 2011 ¹³⁵
	Step-down combination	1	Tosh <i>et al.</i> 2011 ¹³⁵
	Intensive combination	1	Tosh <i>et al.</i> 2011 ¹³⁵
Monotherapy bDMARDs	ADA	2	Chen <i>et al.</i> 2006 ¹²⁴ , Spalding <i>et al.</i> 2006 ¹³⁴
	ETN	3	Chen <i>et al.</i> 2006 ¹²⁴ , Davies <i>et al.</i> 2009 ¹²⁵ , Spalding <i>et al.</i> 2006 ¹³⁴
Combination bDMARDs	ADA+MTX	3	Chen <i>et al.</i> 2006 ¹²⁴ , Davies <i>et al.</i> 2009 ¹²⁵ , Spalding <i>et al.</i> 2006 ¹³⁴
	ETN+MTX	2	Chen <i>et al.</i> 2006 ¹²⁴ , Kobelt <i>et al.</i> 2011 ¹²⁹
	IFX+MTX	3	Chen <i>et al.</i> 2006 ¹²⁴ , Davies <i>et al.</i> 2009 ¹²⁵ , Spalding <i>et al.</i> 2006 ¹³⁴
Other	cDMARDs	2	Chen <i>et al.</i> 2006 ¹²⁴ , Schadlich <i>et al.</i> 2005 ¹³²
	bDMARDs	2	Finckh <i>et al.</i> 2009 ¹²⁶ , Schipper <i>et al.</i> 2011 ¹³³
	Non-DMARD	1	Finckh <i>et al.</i> 2009 ¹²⁶
	Placebo+MTX	1	Hartman <i>et al.</i> 2004 ¹²⁷
	Folic Acid+MTX	1	Hartman <i>et al.</i> 2004 ¹²⁷
	Folinic Acid+MTX	1	Hartman <i>et al.</i> 2004 ¹²⁷

See glossary for definitions of abbreviations.

The studies were diverse in their treatment considerations, and since 2006 seven of the fourteen studies (50%) have evaluated the use of bDMARDs in recent-onset RA.^{124-126,129,133,134,136} Prior to 2006, six studies (43%) were published which evaluated the economic impact of cDMARDs.^{127,128,130-132,137} This leaves one (7%) relatively recent study (Tosh *et al.*

(2011))¹³⁵ evaluating cDMARDs. This suggests that with bDMARDs coming on the market in early 2000s, there has understandably been a shift to evaluating their cost-effectiveness, at the expense of continuing evidence generation for the use of established and lower cost cDMARD treatments.

The disease severity in the patient population being evaluated was not clearly reported across the studies. Kobelt *et al.* (2011) evaluated ETN+MTX vs. MTX in a severe RA population.¹²⁹ Kobelt *et al.* (2002) was an evaluation of MTX+SSZ vs. LEF in any patient with RA.¹²⁸ Six (43%) of studies were explicitly reported as being in an active RA population.^{124,127,132,135-137} In the remaining six (43%) studies, the patient population and disease severity was not reported.^{125,126,130,131,133,134}

Only five (36%) of the studies had a lifelong time horizon for the economic evaluation.^{124-126,134,135} Of these five studies, four of them were evaluations of bDMARDs in recent-onset RA, and all four used decision-analytic modelling methods to estimate costs and effects.^{124-126,134} This included Chen *et al.* (2006), a publication of the independent submission made by the NICE Technology Appraisal Group based at Birmingham.¹²⁴ Only Tosh *et al.* (2011) considered the lifetime costs and effects of alternative cDMARD monotherapy and combination therapy strategies in recent-onset RA.¹³⁵ TNFa's were not considered at this divergence point, due to the evaluation being used to inform the NICE Clinical Guideline,¹²¹ and the NICE guidance at that time published from Technology Appraisals recommending that bDMARDs (specifically, TNF-a inhibitors) only be used after treatment failure with at least two cDMARDs.

Four studies (29%) had a time horizon of no more than 2 years.^{127,130,136,137} A truncated time horizon of this magnitude is likely to omit future costs and benefits that occur between alternative treatments, and in particular if a DMARD therapy is assumed to have a disease-modifying effect on the future course of a chronic condition like RA. Therefore, the short time horizon in these studies is likely to lead to biased estimates of cost-effectiveness.

Ten of the fourteen (71%) studies used DAM methods to estimate expected costs and QALYs.^{124-126,128,129,131,132,134,135} The remaining four studies (29%) were economic evaluations alongside clinical trials.^{127,130,136,137} Prior to 2006, six studies (43%) evaluated the economic impact of cDMARDs, with no evaluation having a time horizon of longer than 10 years.^{127,128,130-132,137} Three of the six studies (50%) undertook an economic evaluation alongside a clinical trial.^{127,130,137} This partially explains the short time horizon and why downstream implications are not fully considered. The extrapolation or modelling of costs and effects may not be the primary objective when reporting a clinical trial; however the results of these studies will be of limited use for resource allocation decision-making.

ii) Downstream costs and effects in recent-onset RA

In the five studies with a lifelong time horizon for the economic evaluation, only Chen *et al.* (2006) explicitly modelled a downstream sequence of treatments.¹²⁴ The analysis allowed a consideration of multiple positions of bDMARDs within the treatment sequence. However, the authors did not attempt to identify an optimal treatment sequence from the available treatment set.

Of the remaining four studies, Tosh *et al.* (2011) considered alternative cDMARD monotherapy and combination therapy strategies in recent-onset RA.¹³⁵ TNFa's were not considered at this divergence point, due to the evaluation being used to inform the NICE Clinical Guideline, and the NICE guidance at that time stating that bDMARDs can only be used after treatment with at least two cDMARDs.¹²¹ The lifelong time horizon would have allowed the implications of faster access to bDMARDs (but using combination rather than sequential monotherapy cDMARDs) to be quantified, however the downstream bDMARDs were not explicitly modelled, and instead estimates of expected costs and QALYs were added on. Spalding *et al.* (2006) used a pooled estimate of costs and effects to provide evidence of the downstream sequence after comparing the first line use of bDMARDs.¹³⁴ Finckh *et al.* (2009) compared symptomatic care with MTX and bDMARDs, and did not clearly report how future costs and QALYs after treatment failure were estimated.¹²⁶ Davies *et al.* (2009) evaluated bDMARDs at first line position in an explicit sequence (bDMARD, MTX+HCQ, LEF, GLD, PC), however they did not clearly report how evidence was used to determine the cost and QALY impact of these future treatments.¹²⁵

From the nine studies with a truncated time horizon, five explicitly included a downstream sequence of treatments.^{129,131-133,136} Kobelt *et al.* (2011) evaluated ETN+MTX vs. MTX over a 10 year time horizon, with a downstream sequence of two bDMARDs and then progression to a standard therapy extrapolation of costs and disease activity.¹²⁹ Both Maetzel *et al.* (2002)¹³¹ and Schadlich *et al.* (2005)¹³² evaluated the impact of adding LEF to a cDMARD sequence at second line, over a five year and three year time horizon, respectively. Neither study evaluated the cost-effectiveness of adding LEF at alternative positions in the sequence. Schipper *et al.* (2011) evaluated the cost-effectiveness of allowing sequential bDMARD use in recent-onset RA, over a five year time horizon.¹³³ After bDMARD use the model contained a transition to combination cDMARDs, however the impact of this on costs and effects was not reported. Van den Hout *et al.* (2009) compared cDMARD monotherapy and combination therapies with initial IFX+MTX therapy, over a two year time horizon.¹³⁶ The analysis was an economic evaluation alongside a clinical trial, and after switching treatment in the trial the patient progressed to another active therapy. The trial was reported as Intention to Treat (ITT), and so the costs and

effects of transition to downstream sequential therapies were included in the economic evaluation but only within the trial follow up period.

Four studies remain with a truncated time horizon and no explicit inclusion of downstream costs and effects.^{127,128,130,137} All four studies are relatively old (1998-2004) and are evaluations of cDMARDs. For these treatments, there was less of a focus on future benefits such as disease control and joint damage, and more of a focus on a short term reduction in disease activity. Three of the four studies were clinical trials,^{127,130,137} and only Kobelt *et al.*(2002) used a DAM to estimates costs and effects over a 10 year time horizon.¹²⁸ In their analysis, long term costs and effects are derived from an observational study (the Early Rheumatoid Arthritis Study).

iii) Decision-analytic modelling methods in recent-onset RA

Ten of the fourteen (71%) studies used DAM methods to determine expected costs and QALYs.^{124–126,128,129,131,132,134,135} Two of the ten models (20%) were a decision tree,^{131,132} four studies (40%) are cohort Markov/State-transition models,^{128,129,133,134} and four studies (40%) are ILMs.^{124–126,135}

The decision tree model by Maetzel *et al.* (2002) had a 5 year time horizon and was capable of modelling a sequence of six explicit treatments.¹³¹ However, this modelling method required simplifications which lead to limitations of the final analysis. In particular, only one level of treatment response was incorporated (ACR20), with the authors recognising that incorporating ACR50 would have allowed the potential superiority of newer DMARDs to be quantified in the model. Also, the model only incorporates approximate direct costs over the long term. The decision tree model by Schadlich *et al.* (2005)¹³² had a 3 year time horizon and was very similar to that of Maetzel *et al.* (2002).¹³¹ It also suffered from the same limitations,¹³¹ and additionally the fact that it did not account for disease duration or diminished clinical response for cDMARDs used at later points in the sequence.

The four Markov models defined health states and transition probabilities between different states. Two defined these health states by HAQ score, one by DAS score, and one simply by either being on an active treatment or dead, and with time dependent costs and utilities.^{128,129,133,134}

The four ILMs explicitly modelled sequential treatments.^{124–126,135} Tosh *et al* (2011)¹³⁵ and Davies *et al.* (2009)¹²⁵ used a regular 6-month time point to update costs and QALYs. This represents a simplification of evidence, in particular when events can occur at any time, or when regular events (such as treatment re-administration) occur outside of the 6-month cycle. Chen *et al.* (2006)¹²⁴ and Finckh *et al.* (2009)¹²⁶ overcome this limitation by being a time-to-

event model. The model time is updated when an event occurs which has an impact on costs or effects.

The six older studies evaluating cDMARDs in recent-onset RA were, unsurprisingly, less likely to meet the Drummond checklist for assessing the quality of the study.^{127,128,130–132,137} Only Maetzel *et al.* (2002)¹³¹ fully met the Drummond criteria. The other studies in general did not have a long enough time horizon to fully capture future costs and benefits,^{127,128,130,132,137} and did not report a fully incremental analysis between alternatives.^{127,132,137} Probabilistic Sensitivity Analysis was not commonly performed, however if detailed and comprehensive scenario and one-way sensitivity analyses were performed then it was considered that this was an appropriate level of testing for uncertainty.

Of the eight newer studies, five fully met the Drummond criteria.^{124–126,135,136} Kobelt *et al.* (2011)¹²⁹, Schipper *et al.* (2011)¹³³ and Spalding *et al.* (2006)¹³⁴ did not clearly detail the evidence to establish the programme's effectiveness, and the latter two studies did not report fully incremental results.

iv) Health economic results in recent-onset RA

Seven studies (50%) evaluated the economic impact of cDMARDs in patients with recent-onset RA.^{127,128,130–132,135,137} Three of these studies evaluated combination cDMARD strategies, and all three found that a combination of cDMARDs dominated monotherapy cDMARDs.^{130,135,137} Of the remaining four studies, three evaluated LEF monotherapy. Maetzel *et al.* (2002) estimated an ICER for LEF of \$Can71,000 per QALY compared to a cDMARD sequence.¹³¹ Kobelt *et al.* (2002) concluded that LEF either dominates or is dominated compared to SSZ and MTX depending on the clinical evidence used to derive effectiveness.¹²⁸ Schadlich *et al.* (2005) estimated that adding LEF to a cDMARD sequence generated additional QALYs, with an ICER of €8,000 per QALY.¹³² Hartman *et al.* (2004) estimated that, adjunct to MTX, folic acid was dominated by placebo, and folinic acid dominated placebo.¹²⁷

In the seven studies (50%) evaluating the economic impact of bDMARDs in patients with recent-onset RA, the general conclusion was that bDMARDs added both costs and benefits to cDMARD comparators.^{124–126,129,133,134,136} However, Finckh *et al.* (2009) estimated that bDMARDs would be dominated by cDMARDs in recent-onset RA.¹²⁶ Chen *et al.* (2006),¹²⁴ Schipper *et al.* (2011),¹³³ Spalding *et al.* (2006),¹³⁴ and van den Hout *et al.* (2009)¹³⁶ concluded that the ICERs comparing bDMARDs to cDMARDs are likely to be too high for decision-makers to approve. Only Davies *et al.* (2009),¹²⁵ with an ICER of \$47k per QALY for ADA+MTX vs cDMARDs, and Kobelt *et al.* (2011),¹²⁹ with an ICER of €13k per QALY for ETN+MTX vs MTX, are

potentially within the threshold for being cost-effective.* Both analyses are for countries (US and Sweden respectively) where cost-effectiveness thresholds are not established for health resource allocation decision-making.

Of the fourteen studies, six (43%) reported sensitivity analysis did not significantly alter the baseline estimates of cost-effectiveness.^{128,130,132,133,137,182} It was not possible to clearly identify what criteria were used to suggest the results were robust, and whether the sensitivity analysis was comprehensive enough. Eight studies reported significant decision uncertainty, with four studies (29%) reporting specific model parameters which lead to decision uncertainty. These were the progression rate of HAQ whilst on treatment,^{124,125} the mapping algorithm from HAQ to utility,¹²⁵ the initial effectiveness,¹²⁴ and withdrawal rate for cDMARDs,¹²⁵ and the initial change in HAQ score after a treatment response.¹³⁴

CRITICAL APPRAISAL OF THE ECONOMIC EVIDENCE FOR DISEASE-MODIFYING THERAPIES IN ESTABLISHED RA

Presented here is a critical appraisal of the economic evidence identified within established RA. 42 studies provided economic evidence for treatments of established RA.^{138–155,157–180}

i) Scope of the economic evaluations of disease-modifying therapies in established RA

All 42 studies were economic evaluations of DMARD therapies for people with established RA. The summary details for all of the established RA economic evaluations are presented in Table 4.5. 29 of the 42 studies (69%) were CUAs, with effects quantified as QALYs.^{139–143,145–147,151–155,157,159–163,165–167,174–180} Nine studies (21%) were CEAs,^{138,144,148–150,169,171–173} with four using low disease activity score (LDAS) or remission as the unit of effect,^{144,150,171,172} two with ACR70 weighted response,^{148,149} and one study apiece using per patient improved,¹³⁸ HAQ improvement,¹⁶⁹ and DAS improvement.¹⁷³ Two studies (1%) were CCAs,^{158,164} and two studies (1%) were CMAs.^{168,170} The specific treatments evaluated for people with established RA in the identified studies are reported in Table 4.8.

* Assuming a threshold of £30,000 (or €40,000 or \$50,000) per QALY.

Table 4.8: Treatments evaluated for established RA

Type	Treatment	Studies	References
Monotherapy cDMARDs	AZA	1	Shini <i>et al.</i> 2010 ¹⁷³
	CYA	1	Anis <i>et al.</i> 1996 ¹³⁸
	GLD	1	Barton <i>et al.</i> 2004 ¹⁴¹
	HCQ	2	Osiri <i>et al.</i> 2007 ¹⁶⁹ , Shini <i>et al.</i> 2010 ¹⁷³
	LEF	3	Maetzel <i>et al.</i> 2002a ¹⁶⁴ , Rubio-Terrés <i>et al.</i> 2001 ¹⁷⁰ , Welsing <i>et al.</i> 2004 ¹⁷⁹
	MTX	10	Barbieri <i>et al.</i> 2005 ¹⁴⁰ , Choi <i>et al.</i> 2000 ¹⁴⁸ , Choi <i>et al.</i> 2002 ¹⁴⁹ , Kobelt <i>et al.</i> 2003 ¹⁵⁹ , Kobelt <i>et al.</i> 2005 ¹⁶¹ , Maetzel <i>et al.</i> 2002a ¹⁶⁴ , Marra <i>et al.</i> 2007 ¹⁶⁶ , Osiri <i>et al.</i> 2007 ¹⁶⁹ , Shini <i>et al.</i> 2010 ¹⁷³ , Wong <i>et al.</i> 2002 ¹⁸⁰
	SSZ	1	Shini <i>et al.</i> 2010 ¹⁷³
Combination cDMARDs	AZA+Pen	1	Anis <i>et al.</i> 1996 ¹³⁸
	HCQ + SSZ	2	Osiri <i>et al.</i> 2007 ¹⁶⁹ , Shini <i>et al.</i> 2010 ¹⁷³
	HCQ + SSZ + MTX	3	Choi <i>et al.</i> 2000 ¹⁴⁸ , Choi <i>et al.</i> 2002 ¹⁴⁹ , Osiri <i>et al.</i> 2007 ¹⁶⁹
	MTX + HCQ	2	Osiri <i>et al.</i> 2007 ¹⁶⁹ , Shini <i>et al.</i> 2010 ¹⁷³
	MTX + SSZ	2	Osiri <i>et al.</i> 2007 ¹⁶⁹ , Shini <i>et al.</i> 2010 ¹⁷³
	MTX + LEF	1	Shini <i>et al.</i> 2010 ¹⁷³
	MTX + CYA	2	Choi <i>et al.</i> 2000 ¹⁴⁸ , Choi <i>et al.</i> 2002 ¹⁴⁹
Monotherapy bDMARDs	ABT	6	Beresniak <i>et al.</i> 2011 ¹⁴⁴ , Cimmino <i>et al.</i> 2011 ¹⁵⁰ , Hallinen <i>et al.</i> 2010 ¹⁵⁴ , Malottki <i>et al.</i> 2011 ¹⁶⁵ , Russell <i>et al.</i> 2009 ¹⁷¹ , Saraux <i>et al.</i> 2010 ¹⁷²
	ADA	11	Bansback <i>et al.</i> 2005 ¹³⁹ , Beresniak <i>et al.</i> 2011 ¹⁴⁴ , Chiou <i>et al.</i> 2004 ¹⁴⁷ , Cimmino <i>et al.</i> 2011 ¹⁵⁰ , Diamantopoulos <i>et al.</i> 2012 ¹⁵³ , Hallinen <i>et al.</i> 2010 ¹⁵⁴ , Malottki <i>et al.</i> 2011 ¹⁶⁵ , Russell <i>et al.</i> 2009 ¹⁷¹ , Saraux <i>et al.</i> 2010 ¹⁷² , Soini <i>et al.</i> 2012 ¹⁷⁴ , Wailoo <i>et al.</i> 2008 ¹⁷⁸
	ANA	2	Chiou <i>et al.</i> 2004 ¹⁴⁷ , Wailoo <i>et al.</i> 2008 ¹⁷⁸
	ETN	18	Bansback <i>et al.</i> 2005 ¹³⁹ , Barton <i>et al.</i> 2004 ¹⁴¹ , Brennan <i>et al.</i> 2004 ¹⁴⁵ , Chiou <i>et al.</i> 2004 ¹⁴⁷ , Choi <i>et al.</i> 2000 ¹⁴⁸ , Choi <i>et al.</i> 2002 ¹⁴⁹ , Clark <i>et al.</i> 2004 ¹⁵¹ , Coyle <i>et al.</i> 2006 ¹⁵² , Diamantopoulos <i>et al.</i> 2012 ¹⁵³ , Hallinen <i>et al.</i> 2010 ¹⁵⁴ , Jobanputra <i>et al.</i> 2002 ¹⁵⁵ , Kobelt <i>et al.</i> 2005 ¹⁶¹ , Malottki <i>et al.</i> 2011 ¹⁶⁵ , Nuijten <i>et al.</i> 2001 ¹⁶⁸ , Russell <i>et al.</i> 2009 ¹⁷¹ , Saraux <i>et al.</i> 2010 ¹⁷² , Soini <i>et al.</i> 2012 ¹⁷⁴ , Tanno <i>et al.</i> 2006 ¹⁷⁵ , Wailoo <i>et al.</i> 2008 ¹⁷⁸
	IFX	13	Barton <i>et al.</i> 2004 ¹⁴¹ , Beresniak <i>et al.</i> 2011 ¹⁴⁴ , Chiou <i>et al.</i> 2004 ¹⁴⁷ , Cimmino <i>et al.</i> 2011 ¹⁵⁰ , Diamantopoulos <i>et al.</i> 2012 ¹⁵³ , Hallinen <i>et al.</i> 2010 ¹⁵⁴ , Jobanputra <i>et al.</i> 2002 ¹⁵⁵ , Lekander <i>et al.</i> 2010 ¹⁶² , Malottki <i>et al.</i> 2011 ¹⁶⁵ , Nuijten <i>et al.</i> 2001 ¹⁶⁸ , Russell <i>et al.</i> 2009 ¹⁷¹ , Saraux <i>et al.</i> 2010 ¹⁷² , Wailoo <i>et al.</i> 2008 ¹⁷⁸
	RTX	7	Benucci <i>et al.</i> 2011 ¹⁴³ , Beresniak <i>et al.</i> 2011 ¹⁴⁴ , Cimmino <i>et al.</i> 2011 ¹⁵⁰ , Hallinen <i>et al.</i> 2010 ¹⁵⁴ , Lindgren <i>et al.</i> 2009 ¹⁶³ , Malottki

			<i>et al. 2011</i> ¹⁶⁵ , <i>Saraux et al. 2010</i> ¹⁷²
	TCZ	2	<i>Diamantopoulos et al. 2012</i> ¹⁵³ , <i>Soini et al. 2012</i> ¹⁷⁴
Combination bDMARDs	ABT+LEF	1	<i>Benucci et al. 2009</i> ¹⁴²
	ABT+MTX	3	<i>Benucci et al. 2009</i> ¹⁴² , <i>Vera-Llonch et al. 2008</i> ¹⁷⁶ , <i>Vera-Llonch et al. 2008a</i> ¹⁷⁷
	ADA+MTX	1	<i>Bansback et al. 2005</i> ¹³⁹
	ETN+LEF	1	<i>Benucci et al. 2009</i> ¹⁴²
	ETN+MTX	5	<i>Bansback et al. 2005</i> ¹³⁹ , <i>Benucci et al. 2009</i> ¹⁴² , <i>Choi et al. 2000</i> ¹⁴⁸ , <i>Choi et al. 2002</i> ¹⁴⁹ , <i>Kobelt et al. 2005</i> ¹⁶¹
	IFX+MTX	7	<i>Bansback et al. 2005</i> ¹³⁹ , <i>Barbieri et al. 2005</i> ¹⁴⁰ , <i>Coyle et al. 2006</i> ¹⁵² , <i>Kobelt et al. 2003</i> ¹⁵⁹ , <i>Marra et al. 2007</i> ¹⁶⁶ , <i>Rubio-Terrés et al. 2001</i> ¹⁷⁰ , <i>Wong et al. 2002</i> ¹⁸⁰
	RTX+MTX	2	<i>Kielhorn et al. 2008</i> ¹⁵⁷ , <i>Merkesdal et al. 2010</i> ¹⁶⁷
Other	cDMARDs	8	<i>Bansback et al. 2005</i> ¹³⁹ , <i>Brennan et al. 2004</i> ¹⁴⁵ , <i>Brennan et al. 2007</i> ¹⁴⁶ , <i>Clark et al. 2004</i> ¹⁵¹ , <i>Osiri et al. 2007</i> ¹⁶⁹ , <i>Tanno et al. 2006</i> ¹⁷⁵ , <i>Vera-Llonch et al. 2008</i> ¹⁷⁶ , <i>Vera-Llonch et al. 2008a</i> ¹⁷⁷
	Constant disease	4	<i>Benucci et al. 2011</i> ¹⁴³ , <i>Kobelt et al. 2004</i> ¹⁶⁰ , <i>Lekander et al. 2010</i> ¹⁶² , <i>Lindgren et al. 2009</i> ¹⁶³
	bDMARDs	4	<i>Brennan et al. 2007</i> ¹⁴⁶ , <i>Kobelt et al. 2004</i> ¹⁶⁰ , <i>Merkesdal et al. 2010</i> ¹⁶⁷ , <i>Welsing et al. 2004</i> ¹⁷⁹
	Treatment Guidance	1	<i>Kievit et al. 2009</i> ¹⁵⁸

The studies were diverse in the treatments considered, however only four (9%) studies were exclusively for cDMARDs.^{138,164,169,173} This potentially reflects the development of bDMARD therapies in the last 15 years, and their relatively high cost requiring a formal economic evaluation to determine if they offer value for money for use in patients with established RA.

In fourteen (33%) of the 42 studies, the disease severity in the patient population being evaluated was not clearly reported.^{138,141,145,148,149,151,159,160,165,168,169,173,175,178} Eleven (26%) studies were reported as being in an active RA patient population.^{146,155–157,162–164,167,170,179,180} Four (9%) studies were in a severe RA patient population,^{140,152,154,161} leaving thirteen (24%) studies in a moderate-severe RA patient population.^{139,142–144,147,150,153,158,171,172,174,176,177}

Only 19 (45%) of the studies had a lifelong time horizon for the economic evaluation.^{139–141,145,146,151,153–155,157,163,165,167,174–178,180} All of these studies used decision-analytic modelling methods. None of the cDMARD exclusive studies in established RA had a lifelong time horizon. 17 (40%) studies had a time horizon of no more than two years.^{138,142–144,147–150,158,160,164,168–173}

36 (86%) of the 42 studies used DAM methods to determine the expected costs and QALYs.^{138–141,144–155,157–159,161–163,165–168,170–172,174–180} This includes prospective studies with a model to extrapolate estimates into the longer-term. Of the six remaining studies, five were

observational studies,^{101,142,160,169,173} and one was an economic evaluation alongside a clinical trial.¹⁶⁴ None of these six studies had a time horizon longer than two years.

ii) Downstream costs and effects in established RA

In the 19 studies with a lifelong time horizon for the economic evaluation, 13 (68%) explicitly modelled a downstream sequence of treatments.^{114,139,141,145,146,151,153–155,157,163,165,175} None of these studies attempted to estimate the optimal sequence of treatments from the available treatment set.

Bansback *et al.* (2005) evaluated bDMARDs with or without adjunct MTX vs cDMARDs in patients who had already failed on two previous cDMARDs.¹³⁹ The downstream cDMARD sequence was explicitly modelled; however the sequence was fixed for all comparisons. Hallinen *et al.* (2010) compared alternative sequences of bDMARDs after failure on one bDMARD.¹⁵⁴

Jobanputra *et al.* (2002),¹⁵⁵ and Barton *et al.* (2004)¹⁴¹ evaluated ETN and IFX in a cDMARD sequence. ETN and IFX were evaluated in three different positions in a sequence of 10 active therapies. The same decision analytic model was used by Clark *et al.* (2004) to evaluate anakinra in alternative positions in a cDMARD sequence,¹⁵¹ and by Malotki *et al.* (2011) to evaluate bDMARDs after failure on a previous bDMARD.¹⁶⁵

Brennan *et al.* (2004) evaluated ETN in a cDMARD sequence.¹⁴⁵ ETN was only evaluated in one position, after 2 cDMARDs had failed. However, alternative downstream sequences were modelled in scenario analyses. This was the same for a latter evaluation by Brennan *et al.* (2007) comparing TNFa's as a class to a cDMARD sequence.¹⁴⁶ Tanno *et al.* (2006) evaluated ETN in a sequence of three cDMARDs over a patient's lifetime,¹⁷⁵ after failure on bucillamine. The downstream sequence is likely to be too short and omits other cDMARD options and sequential bDMARD use for this patient population.

Diamantopoulos *et al.* (2012) compared alternative positions of TCZ in a bDMARD naïve and experienced population.¹⁵³ Kielhorn *et al.* (2008) evaluated the introduction of RTX+MTX after people had failed on two previous bDMARDs.¹⁵⁷ The downstream sequence, or position of RTX+MTX, was not altered. Lindgren *et al.* (2009) evaluated the introduction of RTX after failure on one previous bDMARD.¹⁶³ The subsequent sequence of treatments was not specified, and was not altered. Merkesdal *et al.* (2010) evaluated the introduction of RTX after failure on one previous bDMARD.¹⁶⁷ The subsequent sequence of cDMARDs was not altered, and no comparison to other bDMARDs was made.

Six studies (72%) had a lifelong time horizon but did not explicitly model the downstream treatments.^{140,174,176–178,180} Barbieri *et al.* (2005) simulated HAQ states with associated costs and utilities.¹⁴⁰ Soini *et al.* (2012) modelled progression to best supportive care, and did not clearly report how costs and HRQL were estimated.¹⁷⁴ Vera-Llonch *et al.* (2008) used the same model for two analyses, and after treatment withdrawal moved onto a linear extrapolation of HAQ with mapped estimates of costs and utilities.^{176,177} Wailoo *et al.* (2008) also extrapolated HAQ after treatment withdrawal.¹⁷⁸ Wong *et al.* (2002) estimated future costs and health effects by simulating a worsening of HAQ score via movement of the modelled cohort through Markov health states.¹⁸⁰

23 of the 42 studies (55%) in established RA did not have a lifelong time horizon. Of these, only six (26%) explicitly modelled a downstream sequence of treatments.^{144,150,152,171,172,179} The time horizon for these studies was no longer than 5 years, and only Coyle *et al.* (2006) considered more than one downstream treatment in the sequence (the other five modelling only a switch onto one other active therapy).¹⁵²

17 studies remain with a truncated time horizon and no explicit inclusion of downstream costs and effects.^{131,138,142,143,147–149,158–162,166,168–170,173} The justification for omitting long-term future costs and effects is not clear in any of the studies. Six studies are either observational analyses,^{101,142,160,169,173} or evaluations alongside a trial,¹⁶⁴ and long-term modelling may not have been the primary research objective.

iii) Decision-analytic modelling methods in established RA

As already mentioned, 36 (86%) of the 42 studies used DAM methods to determine the expected costs and QALYs.^{138–141,144–155,157–159,161–163,165–168,170–172,174–180} Five (14%) of the 36 models were a decision tree,^{138,147–149,171} nine (25%) were cohort Markov models,^{140,152,159,161,162,166,175,179,180} and 16 (44%) were individual level models.^{139,141,145,146,151,153–155,157,163,165,167,174,176–178} For the remaining six (17%) studies, the method of decision-analytic modelling was unclear.^{144,150,158,168,170,172}

Of the five decision tree models,^{138,147–149,171} none had a time horizon of over 2 years. Only Russell *et al.* (2009) considered sequential use of therapies.¹⁷¹ Moving onto a second therapy occurred after an inadequate response, and the evidence for this was not clearly reported.

The nine Markov models were also limited in considering the costs and effects of future treatments.^{140,152,159,161,162,166,175,179,180} Only three had a lifelong time horizon,^{140,175,180} and only three considered sequential use of treatments.^{152,175,179}

The 16 individual level simulations all had a lifelong time horizon.^{139,141,145,146,151,153-155,157,163,165,167,174,176-178} 12 of these studies also considered sequential use of therapies in patients with established RA.^{139,141,145,146,151,153,154,157,163,165,167} All 12 determined a treatment switch by either a short-term lack of response, or a long-term withdrawal due to a loss of efficacy or an adverse event. Initial response was modelled using an ACR response mapped to a HAQ improvement in six models.^{139,145,153,154,157,167} Brennan *et al.* (2007) modelled initial treatment response using the EULAR response categories and mapping the response to EQ-5D (or SF-6D) via a multivariate regression.¹⁴⁶

Only 17 of the 42 (40%) met the Drummond checklist for assessing the quality of the study.^{141,145,146,151,154,157,159,162,165-167,175-178,180} The most common reason for not meeting the Drummond criteria were: not providing a comprehensive description of the competing alternatives;^{174,179} not providing evidence that the programme's effectiveness had been established;^{148,155,168,170} not including all important and relevant costs and consequences;^{144,148,150,158,170,179} not measuring costs and consequences appropriately;^{144,150} not undertaking a fully incremental analysis;^{138,139,144,150,152,153,158,161,168,170-172} not allowing for uncertainty;^{140,147-149,170,172} and not including all issues of interest.^{138-140,144,148,150,152,153,155,158,161,168,170-172}

iv) Health economic results in established RA

The health economic results are provided for each study in Table 4.5. None of the studies looked to identify the optimal sequence of treatments from the treatment set included in the analysis.

Four of the 42 studies (10%) were exclusively for cDMARDs in patients with established RA.^{138,164,169,173} In Maetzel *et al* (2002)¹⁶⁴ observed in a one year economic evaluation alongside a clinical trial that MTX dominates LEF and placebo, with Osiri *et al* (2007) also concluding that MTX+AM dominates AM, and non-MTX strategies are unlikely to be cost effective.¹⁶⁹ Shini *et al.* (2010) performed a CEA with change in HAQ as the unit of health benefit.¹⁷³ Their study suggests that HCQ is the most cost effective monotherapy cDMARD strategy, with MTX+HCQ the most cost effective combination strategy. Anis *et al.* (1996) estimated an ICER of CYA therapy of \$1,000 per patient improved compared to placebo.¹³⁸

19 studies (45%) were non-sequential evaluations of bDMARDs in patients with established RA.^{140,142,143,147-149,158,161,162,166,168,170,174,176-178,180} In general, the studies found that bDMARDs were more effective but also more costly compared to cDMARDs in patients with established disease. This conclusion was consistent across all studies, irrespective of country, patient population or method of evaluation. Six of the 19 studies were decision-analytic models with a

lifelong time horizon.^{140,174,176–178,180} Barbieri *et al.* (2005)¹⁴⁰ and Wong *et al.* (2002)¹⁸⁰ estimated an ICER for IFX+MTX vs MTX of £33,000 and £30,000 per QALY, respectively. Likewise, two analyses performed by Vera-Llonch *et al.* (2008) estimated an ICER for ABT+MTX vs MTX of \$43,000 per QALY and \$45,000 per QALY, in a TNFa naïve¹⁷⁷ and TNFa experienced¹⁷⁶ patient population, respectively.

19 studies (45%) were evaluations of alternative sequences of bDMARDs in patients with established RA.^{139,141,144–146,150–155,157,163,165,167,171,172,175,179} 13 of these studies had a lifelong time horizon, and as before these studies found sequential bDMARD use to be more effective but also more costly.^{139,141,145,146,151,153–155,157,163,165,167,175}

Four studies evaluated the introduction RTX into a sequence of DMARDs. Hallinen *et al.* (2010),¹⁵⁴ Lindgren *et al.* (2009),¹⁶³ and Merkesdal *et al.* (2010)¹⁶⁷ concluded that RTX was cost-effective after TNFa failure compared to TNFa's. Kielhorn *et al.* (2008) concluded that RTX after two TNFa failures was cost-effective.¹⁵⁷ None of the studies considered the optimal position of RTX, at the very least by comparing RTX after one or two TNFa failures.*

Of the nine remaining studies, nearly all were consistent in concluding that bDMARDs were likely to be cost effective. The studies by Barton *et al.* (2004)¹⁴¹ and Jobanputra *et al.* (2002)¹⁵⁵ were the only studies to conclude that, after two cDMARDs, bDMARDs were unlikely to be cost effective compared to further cDMARD treatment.

There were six studies with an explicitly modelled sequence of downstream treatments, but with a truncated time horizon.^{144,150,152,171,172,179} These studies reported that bDMARDs were less likely to be cost effective. The truncated time horizon may therefore omit important downstream health benefits from bDMARDs, such as delayed joint erosion or disease progression.

22 of 42 (52%) of studies reported the results were robust when undertaking sensitivity analyses.^{140,141,144,148,150,152–154,157–160,162,163,165,168,170,172,176,177,179,180} As with the similar conclusion from the recent-onset RA population, it was not clear what criteria had been used to suggest that the results were robust, and whether rigorous enough testing had been performed. Eight studies (19%) reported significant uncertainty,^{139,145–147,149,155,166,174} with six studies (14%) reporting specific model parameters which lead to significant sensitivity in the economic model. These were the baseline age in the model,¹³⁹ the standardised mortality ratios,¹³⁹ the algorithm to estimate HRQL,^{139,146,166} the rate of disease progression,^{145,146} discounting rates,¹⁴⁶ ACR response rates,¹⁴⁷ and cost parameters.¹⁴⁹

* RTX is not licensed for used prior to a TNFa therapy

4.5 DISCUSSION

The review identified 57 unique economic evaluations of DMARD therapy for people with RA. However, none of the identified studies have considered identifying the most cost effective sequence from the full treatment set available. This has therefore led to clinical guidance being developed without published economic evidence being available to ensure that health resource allocation decisions are fully informed. Where models have been developed that consider a lifelong time horizon and downstream treatment sequences, evidence gaps have been identified, and evaluations have not fully considered optimising the sequence. These evidence gaps include the efficacy of treatments in downstream positions, and the long term impact of treatments on costs and HRQL in the future. The review has identified that methods have not been consistently applied, which has led to varied estimates of cost-effectiveness and uncertainty with respect to the most appropriate analyses to address particular decision questions.

A number of key themes have been identified from this systematic review of economic evaluations of disease modifying therapies for rheumatoid arthritis.

Firstly, the review highlights the significant decision space within rheumatoid arthritis. Fourteen economic evaluations were undertaken of therapies within a recent-onset RA population, and 42 undertaken within an established RA population. Evaluations were undertaken when people have had no prior treatment, up to patients having had cDMARDs and two bDMARDs. There are several potential positions for each DMARD therapy, and the review identified approximately 30 discrete treatments. Therefore the decision space on a very crude level is every potential sequence constructed from that set of 30 treatments.* Understandably, the vast decision space and therefore huge number of potential comparators led to no study attempting to determine the optimal sequence of therapies. The decision space could be broadly divided into recent-onset RA and established RA populations. The evaluation by Chen *et al.* (2006) represents the only attempt from 57 evaluations to determine whether bDMARDs should be used in recent-onset or established RA.¹²⁴ However, the evaluation only considers a small subset of all feasible treatment sequences. Therefore the review has identified a significant number of constrained or pair-wise evaluations, the majority of which did not conduct a fully incremental analysis or discuss the possibility of alternatives positions other than the primary analysis. This is not particularly surprising, because each

* $30! = 265,252,859,812,191,058,636,308,480,000,000$. If each sequence took one second to enumerate, it would take over 8 years to solve

study was undertaken for its own particular decision-making context. The heterogeneity in terms of comparators, sequences and methodology reflect both local/national variation and also the context in which health economic evaluation is conducted.

Secondly, the modelling methodology was a significant predictor with respect of the quality of the study and the ability to evaluate alternative sequences. Models with a lifelong time horizon were more likely to be an individual level simulation, and Markov and decision tree models were less likely to evaluate the impact of switching onto another therapy. In all studies, the quality of reporting about the impact of future treatments on costs and health benefits was varied. Within a short peer-reviewed journal article it is understandable that not every detail regarding a model can be fully explained. However, in sensitivity analyses undertaken in several studies, the long term progression of disease was shown to be a key parameter that determines cost-effectiveness.

Finally, when downstream treatments were explicitly modelled, the evidence used to parameterise this part of the model was not consistent, and also poorly reported. Evidence used was often referred to rather than explicitly stated. In several evaluations assumptions of equal efficacy between treatments, or potential treatment decrements for later positioning within a sequence, was referred to when direct evidence was not identified. However, the quantitative or qualitative evidence to support these assumptions was not provided. The assumptions used to determine differences in impact of alternative treatments lead to significant uncertainties in the evaluations, and also highlighted that when cDMARDs or bDMARDs can largely be considered a class, with similar costs and health effects, small assumptions can have a significant impact on a treatment's cost-effectiveness. Therefore it is important to identify and synthesise all relevant evidence to inform models, not just at the divergence point, but also throughout the complete model pathway.

As with any systematic review, there are limitations that should be considered. The review does not include non-economic evaluations, or purely disease modelling studies. Some studies which modelled a sequence of treatments may have been omitted. These studies may have provided data regarding modelling methods, but their usefulness would have been limited by not being a comparative evaluation, which will determine model structure, evidence and assumptions.

Secondly, there were some aspects of the data extraction which relied on a certain level of subjectivity. Where possible, checkbox choices and the Drummond checklist were used to ensure bias was minimised. However, when considering particular modelling methodologies and the ability of the model to estimate sequences or identify an optimal sequence, the

knowledge of the reviewer was required. To minimise bias, the reviewer relied on what was reported by the author as fact. Where details were missing, this was noted, rather than assuming what had been undertaken. Also the identified systematic reviews of economic evaluations in RA were cross-checked when data extraction overlapped. Some of this subjectivity could have been accounted for by having multiple reviewers to ensure consistency. However, for this PhD it was not feasible.

Finally, manufacturer's submissions to organisations such as NICE were not included, because full text versions of their reports are not publicly available. These submissions would have potentially been a very informative source of evidence regarding some of the published studies identified, as well as unpublished models which have been used substantially to inform health resource allocation decisions within RA.

The electronic searches were conducted in February 2013. There is the potential for relevant studies to have been published after this date. The review has not been formally updated due to time constraints; however non-systematic searches were conducted to ensure that any recently published studies would not materially affect the conclusions of this systematic review. These searches were performed when the systematic review was published, when it contributed to the NICE biologics TA, and when the thesis was finalised in 2015. No major papers were identified.

4.6 CONCLUSIONS

This chapter highlights that treatment sequences represent a challenge when undertaking an economic evaluation of DMARDs for people with RA. The methods used to model treatment sequences have a significant impact on the final estimates of cost-effectiveness, and these methods have not been consistently applied. This has led to varied estimates of cost-effectiveness, which may potentially alter decisions regarding reimbursement if used in practice. The level of reporting of the methods and data used to assess the impact of downstream treatments in a sequence was poor, and when downstream treatments have been modelled, evidence gaps have been identified.

This systematic review has demonstrated the significant challenges faced when attempting to estimate cost-effectiveness of competing treatment sequences. There is therefore a requirement for methods that allows all relevant sequences to be evaluated. The remainder of this thesis attempts to address this, by developing a flexible decision-analytic model in Chapter 5 which addresses the model limitations identified in this review.

(Blank page)

CHAPTER 5: DEVELOPING A HEALTH ECONOMIC MODEL FOR THE NICE TECHNOLOGY APPRAISAL OF BIOLOGICS FOR RHEUMATOID ARTHRITIS

5.1 CHAPTER OVERVIEW

The first three chapters provided the background, context and rationale for this thesis. In Chapter 4, a systematic review was undertaken to highlight the extent to which treatment sequences have been evaluated in RA, and some of the methodological challenges which have arisen in previous attempts to develop decision analytic models for RA economic evaluations.

This chapter introduces the current NICE Technology Appraisal of biologics for RA in Section 5.2, and the decision problem is defined in Section 5.3. NICE decided not to appraise downstream (post 1 bDMARD) sequences within this appraisal. This has led to the potential for a sub-optimal sequence to be recommended. The reasons for this are detailed, and subsequently discussed in Section 5.7.

For the appraisal, a health economic model was developed. The chapter reports the conceptualisation of the model in Section 5.4, and highlights how key decisions regarding the structure of the model were reached. A discrete event simulation model was developed for the appraisal, and this is reported in Section 5.5. The model evaluates three main populations, patients with severe RA and no previous DMARD treatment (Population 1), patients with severe RA and treatment with two previous DMARDs (Population 2), and patients with moderate to severe RA and treatment with two previous DMARDs (Population 3). These populations further divided into two, one sub-population with people eligible for methotrexate (MTX) therapy, and one sub-population who are not.

The results of the analysis are reported, and the model finds that the ICER for bDMARD use in populations 2 & 3 are over £60,000 per QALY gained compared to a cDMARD treatment sequence.* In patients who are ineligible for MTX, the ICERs are higher, at approximately £90,000 per QALY gained. bDMARD therapy in Population 1 is unlikely to be cost-effective, with ICERs of over £300,000 per QALY gained.

The key component of the model that significantly affects the estimated ICER is the growth model used to estimate HAQ progression whilst on cDMARD therapy. Using a previously used

* The original results for the analysis reported in the Assessment Group report – viewable here: <http://www.nice.org.uk/guidance/gid-tag313/documents/rheumatoid-arthritis-adalimumab-etanercept-infliximab-certolizumab-pegol-golimumab-abatacept-and-tocilizumab-review-assessment-report2> - Accessed June 2015

linear model brings the ICER down to £37,000 per QALY for Population 2 patients who can receive MTX. However, this linear model is less valid compared to the latent class growth model used within this analysis.

The results and the ongoing nature of the appraisal are discussed in Section 5.7, before conclusions are drawn in Section 5.8.

5.2 INTRODUCTION

This chapter has two objectives. The first is to report the development of a health economic model for the NICE Multiple Technology Appraisal (MTA) of bDMARDs for RA. The second is to explain how NICE determined the scope for this appraisal and their rationale for not fully evaluating all possible treatment sequences.

At the start of the PhD, an opportunity arose to develop a Technology Assessment Group (TAG) cost effectiveness model for the NICE MTA update (now referred to as the 'NICE RA biologics appraisal') of bDMARDs in people with moderate and severe RA.*

With the PhD requiring the development of a flexible health economic model for RA, it was decided that developing the model for the NICE RA biologics appraisal would be beneficial to the PhD. In particular, the opportunity to engage with the NICE appraisal process for a complex chronic condition with treatment sequences would add value to this PhD.

This chapter is divided into four sections. Section 5.3 explains the decision problem for the NICE RA biologics appraisal. Section 5.4 explains the conceptualisation process of the health economic model. Sections 5.5 and 5.6 explain how the model was developed, reports the input parameters and assumptions used, and provides the results of the analysis. Section 5.7 presents a discussion before conclusions are drawn in Section 5.8.

INDEPENDENT ASSESSMENT GROUP

The University of Sheffield School of Health and Related Research (SchARR) Technology Assessment Group (SchARR-TAG) are assigned the role of independent assessment group for some NICE single technology appraisals (STAs) and MTAs. For both forms of appraisal, SchARR-TAG provides a critique of the clinical and cost-effectiveness evidence submitted by sponsor organisations, as well as a systematic review of the wider published clinical and cost-

* Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying anti-rheumatic drugs (DMARDs) and after the failure of conventional disease-modifying anti-rheumatic drugs (cDMARDs) only

effectiveness evidence. For MTAs, a *de novo* cost effectiveness analysis is submitted by the TAG for consideration alongside the sponsor submissions by the NICE Technology Appraisal Committee.

The SchARR-TAG assessment for this NICE RA biologics appraisal was understandably a large and collaborative research project, led by Professor Matt Stevenson, and included a team comprised of systematic reviewers, statisticians, information specialists, clinicians, health economic modellers and health economists. As a member of the team, my specific role was to design, develop and validate the *de novo* cost effectiveness analysis, and to include the systematic review of economic evaluations (from this PhD research). Many parameter values used as inputs into the model were identified by other colleagues, which included the development of econometric models to estimate certain patient-level parameters. The final validation and debugging of the model, running the simulations, and reporting the results, was undertaken by Professor Matt Stevenson.

DMARDS: CONVENTIONAL AND BIOLOGIC

The disease modifying anti-rheumatic drugs (DMARDs) licensed for RA are numerous. Table 5.1 provides a list of 16 commonly used DMARDs for RA. They are classified into two groups, conventional DMARDS (cDMARDs), and biologic DMARDs (bDMARDs).

Table 5.1: Disease modifying anti-rheumatic drugs (DMARDs)

Conventional DMARDS (cDMARDs)		Biologic DMARDS (bDMARDs)	
Methotrexate	MTX	Infliximab	IFX
Sulfasalazine	SSZ	Etanercept	ETN
Hydroxychloroquine	HCQ	Adalimumab	ADA
Leflunomide	LEF	Certolizumab pegol	CTZ
Azathioprine	AZA	Golimumab	GOL
Penicillamine	PEN	Rituximab	RTX
Gold injections	GLD	Tocilizumab	TCZ
Cyclosporin	CYA	Abatacept	ABT

Conventional DMARDs are in general older and less expensive. They can be used as monotherapy treatments, in combinations with each other (double or triple cDMARD therapy), or with steroids. Their effectiveness varies, however many patients with RA benefit from their use.

Biologic DMARDs are newer and more expensive. Evidence suggests that bDMARD monotherapies are superior to cDMARDs monotherapies.^{183–185} However, recent evidence suggests non-inferiority between triple cDMARD combination therapy and bDMARDs.¹⁸⁶ The

significant cost of bDMARDs (over £10,000 per annum) is why NICE have been required to appraise their clinical and cost-effectiveness. Some bDMARDs are used as monotherapy, and some are only licensed for use with concomitant methotrexate (Table 5.3). At present there is little evidence to support combination therapies of multiple bDMARDs and their high cost would suggest that any clinical use is unlikely.

5.3 DECISION PROBLEM

As discussed in Chapter 3, a decision problem is where a decision maker is faced with competing options. NICE faces a complex decision problem in the treatment of people with RA. The fundamental problem is - which treatment(s) to approve for people with a diagnosis of RA? The utility function for the outcome of each competing option is the net benefit for each competing alternative. However, the chronic nature of RA, the numerous treatments available (both cDMARDs and bDMARDs), and the uncertain and limited efficacy of these treatments means that another problem emerges – what treatment(s) to approve for patients with RA who have failed a DMARD? This decision problem is contingent on the recommendations that NICE have made ‘upstream’, because treatments will not be available if they have been used previously.

The decision problem is therefore compounded by the multiple treatment histories that will be faced by a rheumatologist. Some patients will have a new diagnosis of RA and no previous DMARD treatment, and some will have established RA and have failed on several DMARDs.

When considering RA, and reflecting on the findings from the previous chapters, there is a treatment sequencing issue which leads to another decision problem – what is the optimal sequence of treatments for a patient with RA? However, this decision problem is not addressed in either this appraisal, or the previous guidance developed by NICE. The systematic review in Chapter 4 showed that it has not been addressed by any published cost-effectiveness analysis, and exploring this decision problem is an objective of this PhD.

SCOPE

The NICE RA biologics appraisal does not attempt to address the fundamental decision problems stated above. Instead, the scope is a complex subset of questions which have been formed due to the existing NICE guidance which are being reviewed, and the scoping process undertaken by NICE to determine the most important questions to be addressed within the timescales of the appraisal.

The NICE RA biologics appraisal is a review of NICE TA guidance 130, 186, 224, 234, and a part review of TA guidance 225 and 247. Table 5.2 outlines these individual pieces of NICE guidance. All guidance is for patients with a Disease Activity Score (DAS28) greater than 5.1; classified as severe RA by the European League against Rheumatism (EULAR).

Table 5.2: Technology Appraisals being updated by NICE RA biologics appraisal

Appraisal	Type	bDMARDs	Guidance
TA130	MTA	Adalimumab, etanercept, infliximab	Adalimumab, etanercept and infliximab, in combination with methotrexate, recommended in patients with active RA who have tried methotrexate and one other DMARD. Adalimumab or etanercept can be given as monotherapy if intolerant to methotrexate.
TA186	STA	Certolizumab pegol	As per TA130
TA224	STA	Golimumab	Suspended – no evidence submitted by manufacturer
TA234	STA	Abatacept	As per TA130
TA225	STA	Golimumab	As per TA130, and recommended in combination with methotrexate in patients who have failed a previous TNF inhibitor
TA247	STA	Tocilizumab	As per TA130, and recommended in combination with methotrexate in patients who have failed a previous TNF inhibitor and are contraindicated to, or suffer an adverse event from, rituximab.

MTA = Multiple Technology Appraisal. STA = Single Technology Appraisal

TA195 appraised adalimumab, etanercept, infliximab, rituximab and abatacept after TNF inhibitor failure. The recommendation was that rituximab was the preferred treatment in this position, and the other treatments were only recommended if the patient was contraindicated to, or suffer an adverse event from, rituximab. TA195 was not to be updated within this appraisal (for reasons detailed later in this section). Therefore this guidance represented a constraint to the sequences which could be evaluated within this appraisal.

A NICE Clinical Guideline was published in 2009, recommending the use of combination cDMARDs* in patients with early active RA.¹²¹ However, this is only a recommendation, and is not mandatory guidance for the NHS.

* Specifically an intensive cDMARD combination which combines three cDMARDs and steroids, as used in the TICORA trial³³⁶

From these seven NICE Technology Appraisals, as well as a NICE Clinical Guideline, a complex sequence of treatments has emerged for patients with severe active RA (DAS>5.1). This sequence is shown in Figure 5.1.

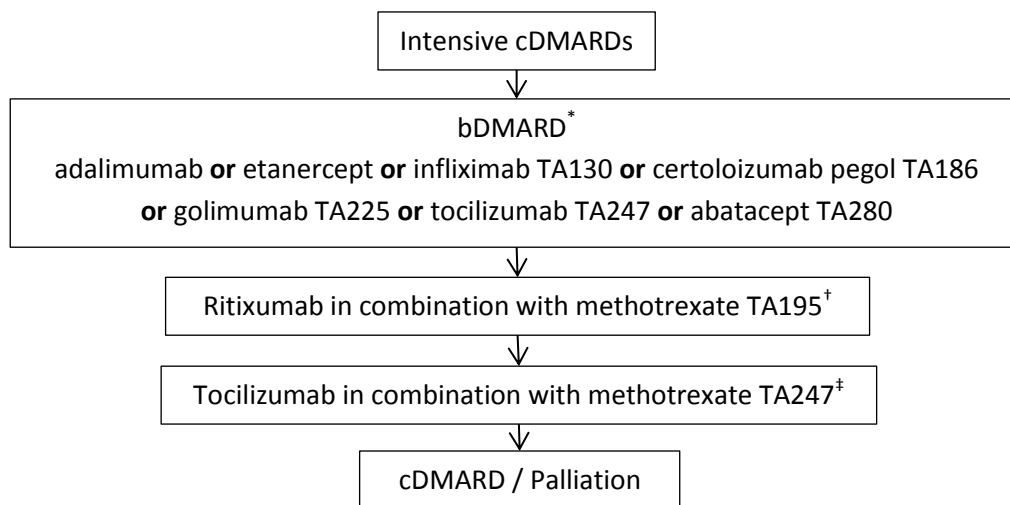


Figure 5.1: Summary of biologics within NICE TA guidance

For the NICE RA Biologics appraisal, the remit of the appraisal was to “appraise the clinical and cost effectiveness of adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept within their licensed indications for the treatment of rheumatoid arthritis.”[§]

The scope was to re-appraise the recommendations made for severe active RA patients in the light of new evidence, and also to assess whether the interventions were cost-effective in moderate-to-severe active RA patients (DAS28 3.2 - 5.1). All previously appraised bDMARDs have a licensed indication which covers both moderate-severe and severe RA patients. In the previous NICE TAs, bDMARDs were not approved for use in patients with moderate-to-severe RA, only severe RA. Not all bDMARDs are licensed to be used without concomitant methotrexate, or prior to the use of methotrexate. The licenses for bDMARDs are summarised in Table 5.3.

* In combination with methotrexate

† If rituximab and MTX is contraindicated or withdrawn due to adverse events then the following can be used: adalimumab or etanercept or infliximab or abatacept in combination with MTX; adalimumab or etanercept monotherapy TA195 : tocilizumab in combination with MTX TA 247, assuming these have not been used previously in the sequence

‡ Would not be used if tocilizumab has been used previously in the sequence

§ <http://guidance.nice.org.uk/TAG/313#keydocs> Final Scope – Accessed June 2015

The scope was constrained by the fact that post bDMARDs positions were not to be included in this appraisal (TA195 is to be reviewed separately at a later date). This means that after a bDMARD is used in a sequence, the sequence then follows the existing NICE TA195 guidance. This limits the ability for the appraisal to identify and recommend an optimal sequence of DMARDs for patients with severe RA.

There were three reasons why NICE made the decision not to include sequential bDMARD use within the scope of the appraisal. Firstly, there was a desire by the manufacturers to see a re-appraisal of biologics in the moderate-severe active RA population in the light of the original rejection by NICE and a maturing of the evidence base. This guided the scoping discussions, and placed the focus on this patient population. Secondly, manufacturers were not interested in addressing sequential positions of their treatments. If manufacturers identify an optimal sequence, it exposes them to the risk of their treatment not being recommended, or being ‘relegated’ to a less attractive downstream position. Thirdly, NICE believed that the appraisal was already complex enough without adding in the sequential biologics question. The scope of the appraisal (seven bDMARDs) across multiple patient populations represents the largest NICE appraisal to date.

Table 5.3: Licenses for RA bDMARDs

Intervention	Is the intervention licensed...			
	Prior to the use of MTX?	As a monotherapy?	For patients with severe RA?	For patients with moderate to severe RA?
Abatacept			✓	✓
Adalimumab	✓	✓	✓	✓
Certolizumab pegol		✓	✓	✓
Etanercept	✓	✓	✓	✓
Golimumab	✓		✓	✓
Infliximab	✓		✓	✓
Tocilizumab		✓	✓	✓

5.4 MODEL CONCEPTUALISATION

For the *de novo* cost-effectiveness model developed by the TAG, a conceptual model was firstly developed. This process involves translating the decision problem with the aim of guiding the development of a mathematical model.^{54,187}

The conceptualisation of the model for the TAG submission involved understanding the scope of the appraisal and making decisions regarding the sequences of treatments to be compared

within a cost-utility analysis. It also involved the conceptualisation of a health economic model which could estimate the costs and QALYs of each defined DMARD treatment sequence.

FINAL DECISION PROBLEM

Having consulted with NICE, the TAG project team made the decision to fix the downstream sequence of drugs beyond the first use of bDMARDs. This was for two reasons: firstly to reflect the NICE scope for the appraisal, where existing guidance was available for sequential biologics (TA195) and was not changeable within the remit of this appraisal, and secondly to reduce the workload of the TAG by reducing the number of evaluations required. It avoided the undertaking of analyses which would not be considered by the appraisal committee and therefore would not have a bearing on the final guidance.

The TAG defined three patient populations for their report and analyses.

- **Population 1:** Adults with severe active RA not previously treated with cDMARDs
- **Population 2:** Adults with severe active RA that have been previously treated with cDMARDs but not bDMARDs
- **Population 3:** Adults with moderate to severe active RA that have been previously treated with cDMARDs only (including MTX unless contraindicated or inappropriate)

Population 1 is patients with severe active RA who are newly diagnosed and have not been treated with cDMARDs. Population 2 and 3 follow this point, and are patients who have been treated with cDMARDs, but not bDMARDs. Population 2 is severe patients with RA, and Population 3 moderate to severe patients with RA.

These populations were run for two different analyses. One was the comparison of bDMARDs in combination with MTX, and one was the comparison of bDMARD monotherapy. Therefore across three populations, six baseline analyses were undertaken (Table 5.4).

Table 5.4: Populations and analyses

Population	Analysis code	
	In combination with MTX	Monotherapy
1: Adults with severe active RA not previously treated with cDMARDs	1	2
2: Adults with severe active RA that have been previously treated with cDMARDs but not bDMARDs	3	4
3: Adults with moderate to severe active RA that have been previously treated with cDMARDs only (including MTX unless contraindicated or inappropriate)	5	6

For each analysis, a baseline cDMARD treatment sequence was required which represented the baseline comparator. Each bDMARD would then be compared to this baseline cDMARD sequence, as well as each alternative bDMARD sequence.

For Population 1, the NICE RA Clinical Guideline recommends combination cDMARDs.¹²¹ This guidance was based on a cost-effectiveness analysis comparing combination cDMARDs to monotherapy cDMARDs.¹³⁵ Therefore, it was assumed that patients will have had combination cDMARDs as first line treatment, and subsequently the comparator cDMARD sequence would be sequential monotherapy cDMARD use (specifically called non-biologic therapy (NBT)).

After first line bDMARD use (the comparison of the interventions in the appraisal) patients progress to RTX+MTX and then TCZ+MTX (if not received at first line bDMARD position), as per existing NICE guidance. Patients who can receive MTX are eligible to receive TCZ+MTX, which is only licensed for use after MTX (Table 5.5).

For patients who are unable to receive MTX, tocilizumab is not eligible for any sequence, as well as abatacept and certolizumab (only licensed for concomitant use with MTX). Intensive cDMARD treatment is possible with alternative cDMARD treatments (Table 5.6).

Table 5.5: Sequences (patients who could receive MTX)

Population	Treatment sequence (line)						
	1 st	2nd	3rd	4th	5th	6th	7th
Population 1	MTX	Intensive cDMARDs	Non-biologic therapy				
	MTX	Intensive cDMARDs	bDMARD + MTX	RTX + MTX	TCZ + MTX	MTX	Non-biologic therapy
	MTX	Intensive cDMARDs	TCZ + MTX	RTX + MTX	MTX	Non-biologic therapy	
	bDMARD + MTX ¹	RTX + MTX	TCZ + MTX	MTX	Intensive cDMARDs	Non-biologic therapy	
Population 2 & 3	MTX	Non-biologic therapy					
	bDMARD ² + MTX	RTX + MTX	TCZ + MTX	MTX	Non-biologic therapy		
	TCZ + MTX	RTX + MTX	MTX	Non-biologic therapy			

¹Excluding abatacept, certolizumab and tocilizumab. ²Excluding tocilizumab

Table 5.6: Sequences (patients who could not receive MTX)

Population	Treatment sequence (line)				
	1 st	2nd	3rd	4th	5th
Population 1	Intensive cDMARDs	cDMARD	Non-biologic therapy		
	Intensive cDMARDs	bDMARD	bDMARD	cDMARD	Non-biologic therapy
	bDMARD ¹	bDMARD ²	Intensive cDMARDs	cDMARD	Non-biologic therapy
Population 2 & 3	cDMARD	cDMARD	Non-biologic therapy		
	bDMARD ¹	bDMARD ²	cDMARD	Non-biologic therapy	

¹Excluding abatacept, certolizumab and tocilizumab. ²Excluding tocilizumab

Table 5.7 to Table 5.10 provide the full sets of sequences evaluated in each analysis.

Table 5.7: Sequences evaluated for Populations 2 and 3 for those who can receive MTX

	First line	Second line	Third line	Fourth line	Fifth line
1	MTX	NBT			
2	ABT iv+	RTX+	TCZ+	MTX	NBT
3	ABT sc+	RTX+	TCZ+	MTX	NBT
4	ADA+	RTX+	TCZ+	MTX	NBT
5	CTZ+	RTX+	TCZ+	MTX	NBT
6	ETN+	RTX+	TCZ+	MTX	NBT
7	GOL+	RTX+	TCZ+	MTX	NBT
8	IFX+	RTX+	TCZ+	MTX	NBT
9	TCZ+	RTX+	MTX	NBT	

'+' with MTX

Table 5.8: Sequences evaluated for Populations 2 and 3 for those who cannot receive MTX

	First line	Second line	Third line	Fourth line
1	SSZ	NBT		
2	ADA	ETN	SSZ	NBT
3	CTZ	ETN	SSZ	NBT
4	ETB	ADA	SSZ	NBT
5	TCZ	ETN	SSZ	NBT

Table 5.9: Sequences evaluated for Population 1 for those who can receive MTX

	First line	Second line	Third line	Fourth line	Fifth line	Sixth line
1	MTX	Int CD+	NBT			
2	ETN+	RTX+	TCZ+	MTX	Int CD+	NBT

'+' with MTX; Int CD+ = Intensive cDMARDs

Table 5.10: Sequences evaluated for Population 1 for those who cannot receive MTX

	First line	Second line	Third line
1	SSZ	NBT	
2	ETB	ADA	NBT

CONCEPTUAL MODEL

A conceptual model was developed by the project team, which provided an agreed representation of RA, and the treatment of RA, which would be represented mathematically in the decision analytic model.

The conceptual model was developed during project meetings, which included drawing upon the team's experience at developing other health economic models for RA. Time constraints meant that the conceptual model was not formally stated or recorded, but there was agreement as to the important aspects of the condition and its treatments which would need to be captured in the health economic model.

Rheumatoid arthritis is a chronic autoimmune disease which results in inflammation and damage to synovial joints. The conceptual model for the disease focussed on the relapsing-remitting nature of RA, along with an insidious worsening of disease activity leading to irreversible joint damage and permanent disability. Similar to other RA models summarised by Madan *et al.* and Tosh *et al.*, the conceptual model identified three phases of the disease for each treatment, which were repeated as the sequence of treatments progressed.^{188,189}

Phase 1 – Initial response to treatment and improvement in Health Related Quality of Life (HRQL)

Phase 2 – Long-term progression of the disease while on treatment causing a gradual worsening of HRQL

Phase 3 – Loss of efficacy or adverse event causing a worsening of HRQL

A treatment therefore has the opportunity to improve the long term HRQL of patients with RA via the initial response, the progression of disease on treatment, and the time spent on a treatment. After a loss of efficacy or adverse event, the treatment is switched to a subsequent therapy, and the phases are repeated.

DISCRETE EVENT SIMULATION

The decision was made at the conceptualisation stage to develop an individual level simulation model for the NICE RA biologics appraisal. Individual level simulation models use probability distributions for a set of patient characteristics (e.g. age, gender, disease severity), which can be sampled using Monte Carlo methods to simulate an individual patient. Further sampling of future events is conducted for each patient, to simulate their disease and the engagement with health services, treatments and related events. Each patient simulation allows the estimation of lifetime costs and QALYs, and by repeated sampling and simulation of patients,

the expected costs and QALYs for the simulated population can be estimated. As discussed in Chapter 2, there are benefits to using an individual level model methodology instead of cohort model methods, and these benefits were seen as important when looking to develop a lifetime model for RA. In particular, it was *a priori* seen as important to incorporate patient covariates throughout the model process, because certain model parameters were dependent on patient covariates which when repeatedly sampled would provide an accurate estimate of the expected costs and QALYs for the patient population.

Secondly, a patient level approach was seen as important to allow flexibility in assumptions and methods underpinning the analysis, with the expectation that the appraisal committee might request alternative analyses, and also the ability to adapt the model to provide comparison to the models provided by the manufacturers.

Finally, the team was confident that the model could be programmed efficiently using Microsoft Excel and Visual Basic for Applications (VBA) so that computational time for running the patient simulations, probabilistic sensitivity analysis (PSA) and alternative scenarios was not excessive. TAGs are limited to Microsoft Excel, TreeAge, WinBugs and R when developing cost-effectiveness models for NICE appraisals.^{190,191} If they wish to use a bespoke simulation package (such as Simul8 or Arena) then they are expected to provide licences for each stakeholder so that the model can be accessed, which is not feasible.

It was decided at the conceptualisation stage to develop the patient level health economic model using discrete event simulation (DES) methods. A discrete event simulation is often called a 'time to event' model within health economics. An event is scheduled to occur at a particular instance in time, and mark a change of state in the system. For health economic modelling, this means that events can occur to a patient (e.g. begin treatment, experience an adverse event, disease progresses) which will impact both on their instant costs and HRQL, but also on the future competing events which may occur, and when they will occur. A list of all possible events is developed, with a time of when each event will occur. The event with the shortest time is then selected, and this triggers logic in the model code which updates the costs, HRQL and also the list of event times. The process is then repeated until the death event occurs and the simulation ends.

An alternative to DES for a patient level simulation is to use a fixed Markovian time cycle (classified by Brennan *et al.* as simulated patient-level Markov model (SPLMM))⁴⁴ occurring at some arbitrarily time (e.g. 1 month, 6 months, 1 year). The decision between using a DES approach compared to a fixed Markovian approach was made by weighing up the advantages and disadvantages of the two. These are summarised in Table 5.11.

Table 5.11: Discrete vs Markovian patient level simulation

	Discrete event simulation	Fixed Markovian time-cycle
Advantages	Model calculations only occur at actual events – potential improved efficiency Time to event data more accurately incorporated	Provides a regular time cycle for updating costs and QALYs.
Disadvantages	Events may frequently occur which requires a frequent update of the simulation time	Running the simulation model at time cycles when no events are occurring is computationally inefficient May require manipulation of time-specific data to fit into time cycle

Due to the size of the appraisal and number of anticipated model simulations and evaluations, it was decided to use a DES method to try to minimise the run time.

5.5 MODEL DEVELOPMENT

MODEL

The model informed a cost-utility analysis with costs from a NHS and Personal Social Services (PSS) perspective, and outcomes measured as QALYs. The model employs a lifetime patient time horizon (limited to 101 years), with costs and QALYs discounted at 3.5% per annum as recommended by NICE.⁴⁷ A comprehensive set of sensitivity analyses were undertaken, as well as PSA.

The model estimated a Health Assessment Questionnaire (HAQ) score for the patient at each event point because HAQ was used to subsequently estimate costs and a patient's HRQL. HAQ is not a continuous score, but has 25 possible scores from 0-3 at 0.125 intervals. Sampled HAQ scores were continuous, and were rounded to a legitimate discrete HAQ score by using the inverse relation to their distance from legitimate score and estimating probabilities from which to sample from.*

At the start of the model, a patient was simulated with a baseline age, disease duration, HAQ score, Disease Activity Score (DAS), number of DMARDs previously used, and life expectancy.

* A non-legitimate HAQ sample of 1.600 has a 20% chance of being rounded to 1.500 and 80% chance of being rounded to 1.625, because 1.600 is 80% of the distance between 1.500 and 1.625

Their life expectancy was adjusted with a hazard ratio defined by their baseline HAQ to account for the reduction life expectancy experienced by people with RA.

A patient would then begin their first DMARD treatment, with a EULAR response (good, moderate or none) estimated at 6 months.* If a good or moderate EULAR response was simulated, treatment was continued until a loss of efficacy or adverse event occurred and treatment was withdrawn. If no EULAR response occurred, then the treatment would be withdrawn. Each EULAR response is associated with a change in a patient's HAQ score. After withdrawal, the patient would lose any gain in HAQ obtained in the first 6 months and would switch to the next treatment in the sequence.

POPULATION

The model sampled patients who had experienced MTX treatment (Populations 2 & 3) from the British Society for Rheumatology Biologics Registry (BSRBR). A multivariate regression analysis was undertaken using the patient level data which provided an econometric model from which to sample patients with accurately correlated characteristics. Individuals were resampled until they met DAS score for the population being modelled (DAS 3.2-5.1 for moderate-severe, DAS > 5.1 for severe). This required significant resampling for moderate-severe DAS patients because they were a minority in the BSRBR dataset.† Multivariate sampling was undertaken using the University of Sheffield Centre for Bayesian Statistics in Health economics Excel Functions.‡ This allowed the correlation between characteristics to be maintained via the variance-covariance matrix from the regression analysis.

MTX naïve patients (Population 1) were very rarely seen in the BSRBR, due to almost all patients on a bDMARD having been previously treated with MTX. Therefore the COMET trial was used to sample patient characteristics for population 1.¹⁹² Because a covariance matrix was not available in the published journal article, the correlation structure between patient characteristics could not be maintained. The mean parameters for all Population 1 (COMET trial) and Populations 2&3 (BSRBR) are provided in Table 5.12.

* see latter section – Short Term Response, for further information regarding EULAR response

† The AG was not provided with spate databases for the two patient populations

‡ <http://www.shef.ac.uk/chebs> - Accessed June 2015

Table 5.12: Population parameters

Characteristic	COMET trial (population 1)		BSRBR (populations 2 & 3)	
	Mean (unless specified)	s.d.	Mean (unless specified)	s.d.
Age	51.4	0.6	56.2	12.2
Proportion female	73%	-	76.3%	-
Disease duration	0.8	0.0	13.3	9.6
DAS	6.5	1.0	6.6	1.0
Previous DMARDs	0.0	0.0	3.9	1.6
HAQ	1.7	0.7	2.0	0.6
Weight	73.1 ¹	17.61 ¹	73.1	17.6

¹ not reported by COMET so set equal to BSRBR

SIMULATION

A simulation ‘engine’ was developed for the DES model in Microsoft Excel and Visual Basic for Applications (MS VBA). The engine maintained the simulation event list, which contained all possible events, and identified the event on the list with the shortest time to occurrence (Time to Next Event (TTNE)). Once a next event is identified by the engine, the simulation clock is updated and the event occurs, triggering logic code in the model to update costs and QALYs. The event also impacts on the time to other events occurring, and therefore each event can update the simulation event list. The process of sampling the TTNE is then repeated until the next event to occur is death, and the patient simulation is completed.

The simulation included five competing events: HAQ progression; death; administration of a treatment; response to a treatment; and withdrawal from a treatment. More details on each event are provided in Table 5.13.

A logic diagram of the simulation model is provided in Figure 5.2. The diagram shows the different events which can occur in the model, and how simulation time and TTNE are used to determine the flow of a patient through the sequence of treatments.

Table 5.13: Simulation event list

Event List	Method of estimating time to event
Time to HAQ change	Both costs and utility are related to HAQ (and explicitly related in the model). Therefore each change in HAQ score is modelled as an explicit event.
Time to death	Hazard ratios associated with baseline HAQ scores are used to appropriately reduce the sampled life expectancy for each simulated patient at model entry.
Time to administration	The majority of treatments modelled are given in a very frequent (continuous) dosing regimen. However, IFX, RTX and TCZ are infrequent infusions and therefore their dosing is modelled as an event, to ensure accurate costs are calculated in the model.
Time to response	Set for six months until response decision point (good, moderate, none), and then large number until next treatment initiated
Time to withdrawal from treatment	For bDMARDs, uses a statistical model that estimates expected time on treatment using patient covariates. For cDMARDs, published Weibull distributions are sampled from to estimate the time on treatment for different cDMARDs (MTX, combination cDMARDs, SSZ etc)

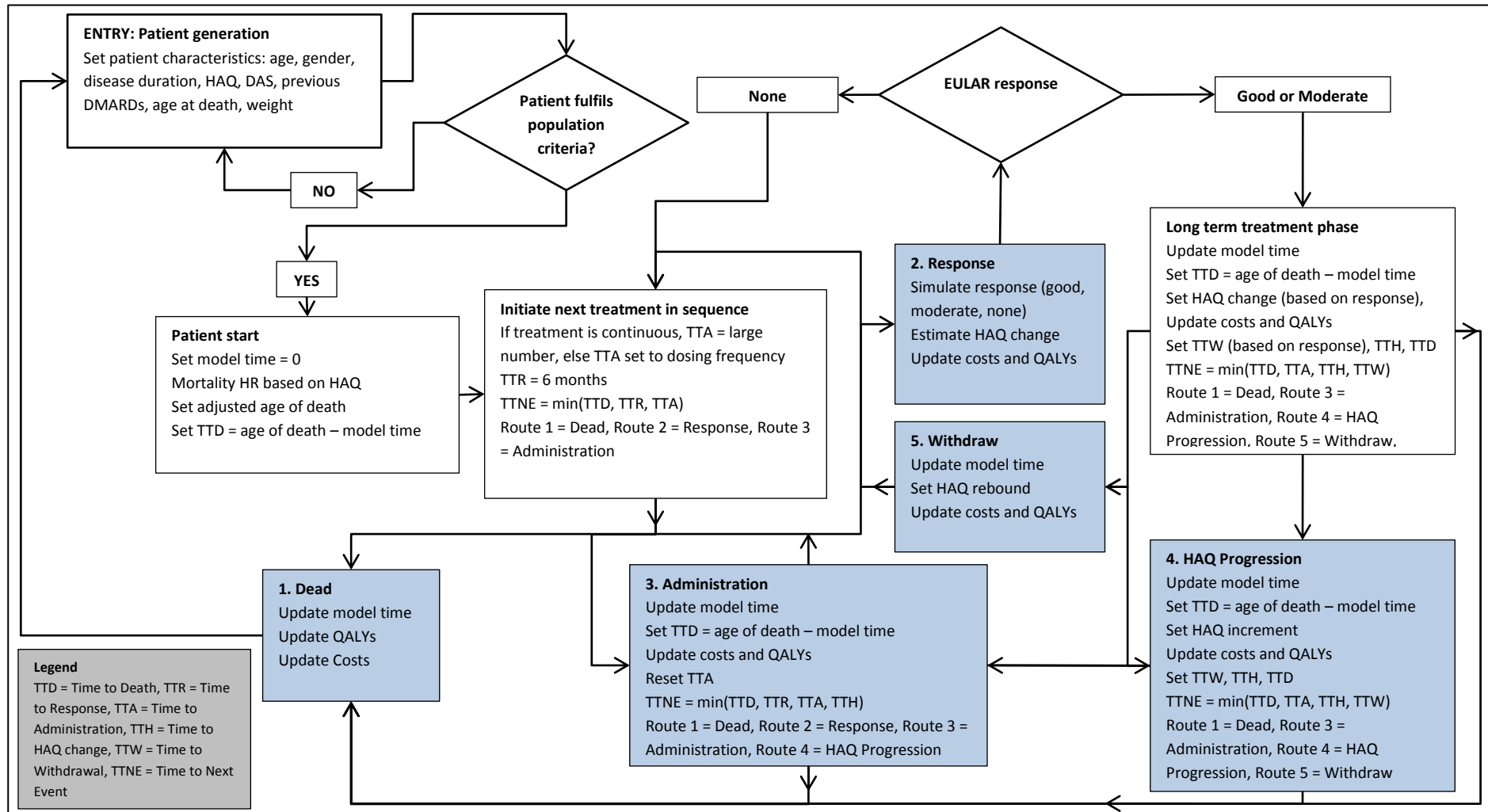


Figure 5.2: Simulation logic for RA model

SHORT TERM RESPONSE

Short term response was estimated using the EULAR criteria for RA. The EULAR criteria defines response based on both the magnitude of a DAS change observed, and also the final DAS score (see Table 5.14). It is a response criterion which is used in current NICE guidance for bDMARDs in RA and is aligned with UK clinical practice. The DAS is a routinely collected measure in the NHS and therefore EULAR response is very easily collected and reported. Alternative response criteria include the American College of Rheumatology (ACR) criteria (ACR20/50/70), which is not routinely used in the NHS, but is very commonly reported in RA clinical trials. In the decision-analytic model, response was assumed to occur during the first six months on treatment, which was aligned with the standard time point for observing response in both clinical practice and the reviewed clinical trials.

Table 5.14: EULAR response criteria

Final DAS28	DAS28 improvement		
	> 1.2	> 0.6 and ≤ 1.2	≤ 0.6
≤ 3.2	Good	Moderate	No response
> 3.2 and ≤ 5.1	Moderate	Moderate	No response
> 5.1	Moderate	No response	No response

The EULAR response probabilities were taken from a network meta-analysis of all included clinical trials. Not all reported EULAR response, however all reported ACR response. Where EULAR was missing, a model using the US Veterans Affairs Rheumatoid Arthritis (VARA) dataset was used to provide an empirical relationship between EULAR response and ACR response. This model allowed all ACR trials to be interpreted using the EULAR response criteria, however these trials were not combined into one analysis, but evaluated separately.

The main analysis used the mean EULAR response from all trials (Figure 5.3). Scenario analyses included the inclusion of trials which did not meet the strict inclusion criteria, and an analysis using the ACR trials.

A sampled EULAR response in the patient level simulation was converted to an appropriate improvement in the patients HAQ score. These changes in HAQ scores were identified from the BSRBR. The BSRBR data were restricted to patients who had a full set of baseline characteristics, and had at least two measurements of HAQ whilst on bDMARD therapy. This resulted in 10,186 included patients, with 2,417 (24%) good EULAR responders, 5,492 (54%) moderate EULAR responders, and 2,277 (22%) EULAR non-responders.

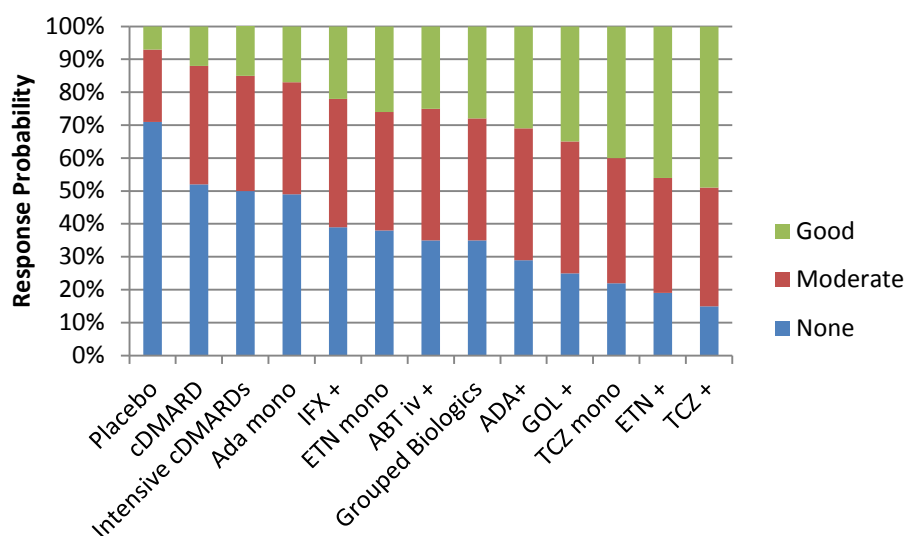


Figure 5.3: Mean EULAR response probabilities for comparator treatments (Population 2&3, main analysis - all ACR trials) - reproduced from SchARR Technology Appraisal Report*

It was assumed that the change in HAQ due to a response would be the same, irrespective of treatment, and therefore was applied to cDMARDs as well as bDMARDs. This assumption was considered acceptable by the clinical advisors to the project. The change in HAQ score are provided in Table 5.15.

Table 5.15: Mean HAQ improvement by EULAR response - BSRBR dataset

EULAR response baseline – 6 month	HAQ change	
	Mean	s.e.
None	0	-
Moderate	-0.317	0.048
Good	-0.672	0.112

HAQ PROGRESSION – CONVENTIONAL DMARDS

After the initial 6 month improvement due to a treatment response, those patients who had either a good or moderate EULAR response remained on a cDMARD until withdrawal due to an adverse event or loss of efficacy. While on treatment, a worsening (progression) of their HAQ over time was modelled.

In many other health economic evaluations of RA therapies, an annual rate of HAQ progression is used for patients receiving cDMARDs to account for disease progression. Estimates include

* '+' with methotrexate

0.08 per annum over 5 years,¹⁹³ 0.06 per annum over 3 years,¹⁹⁴ and 0.05 per annum over 5 years.¹⁹⁵ However, the challenges with estimating this are well recognised and the clinical plausibility is weak, especially for tolerated cDMARDs which can be administered for many years.¹⁹⁶ In models with a lifetime horizon, it's common for HAQ with a linear trajectory to 'bottom out' at 3, which is not clinically realistic, and is a health state usually valued as worse than death.

To overcome these limitations, a non-linear growth mixture model was developed by Norton *et al.* using the Early Rheumatoid Arthritis Study (ERAS) inception cohort study (n=1460, 10 years follow up).¹⁹⁷ This model was corroborated using two other datasets.* The growth mixture model produced four latent classes of HAQ progression, and the probability of membership of each of these classes given the patient descriptors from the DES model. Therefore, given the four latent classes, and a patients set of baseline characteristics, the expected HAQ at any time point can be estimated. This allowed the growth model to be implemented within the DES model.

The baseline characteristics sampled in the cost-utility model provided the probabilities of latent class membership, which when applied to each of the latent classes provided an expected HAQ profile for a given patient. This profile was used to estimate the time to a HAQ increase over the longer-term. The growth mixture model provided a HAQ profile up to 15 years. Patients who stayed on a cDMARD for over 15 years were assumed to remain on a constant HAQ score (no progressive worsening).

HAQ PROGRESSION – BIOLOGIC DMARDS

To estimate the HAQ progression of patients while on a bDMARD treatment, the dataset from the BSBRB to estimate the initial HAQ change due to a EULAR response was used. An Autoregressive Latent Trajectory (ALT) model was fitted to moderate and good EULAR responders.† The model uses baseline patient covariates, including baseline HAQ, to estimate both initial HAQ response (6 months) and the longer term progression of a patient's HAQ in a single statistical model. The predictions of HAQ over time, across different EULAR responders is provided in Figure 5.4. With no worsening of HAQ observed over the three year time period, this assumption was used within the health economic model, for the whole period while on a bDMARD therapy.

* The Norfolk Arthritis Register (NOAR) and the Early Rheumatoid Arthritis Network (ERAN).

† Full details are provided in the Assessment Group report – viewable here:

<http://www.nice.org.uk/guidance/gid-tag313/documents/rheumatoid-arthritis-adalimumab-etanercept-infliximab-certolizumab-pegol-golimumab-abatacept-and-tocilizumab-review-assessment-report2> - Accessed June 2015

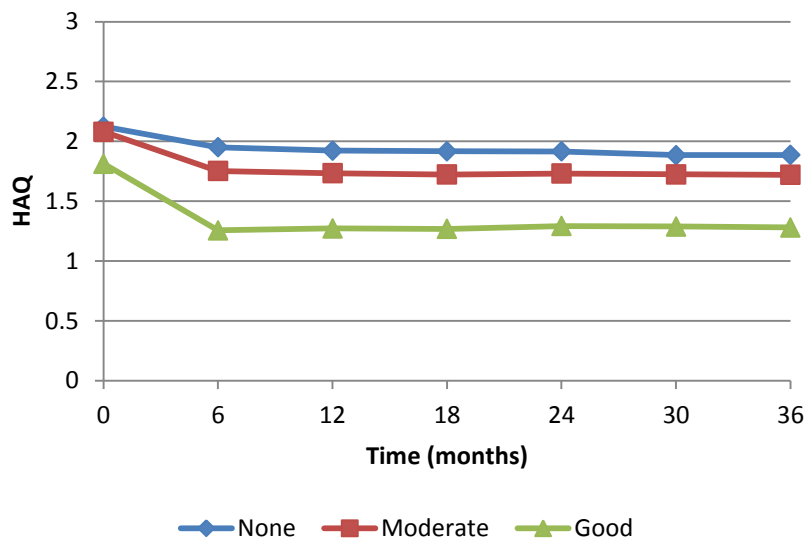


Figure 5.4: Mean HAQ by EULAR response for bDMARD patients

TIME TO TREATMENT WITHDRAWAL

After the six month response period, patients remain on treatment until they either die or withdraw due to a loss of efficacy or adverse event. The BSRBR database was used to estimate survival times on treatments, using the dates on which therapies are initiated and ended. Separate models were fitted for those patients obtaining good and moderate EULAR responses at 6 months.

A range of parametric survival models were considered, and based on the Akaike information criteria (AIC) and Bayesian information criteria (BIC) the generalised gamma distribution was selected for moderate EULAR responders, and the log normal distribution for good EULAR responders.

It was assumed that treatment duration would be unaffected by whether or not cDMARDs were used prior to bDMARDs. It was also assumed that the treatments included in the BSRBR (ETN, IFX, and ABT) would be very similar to the newer bDMARDs being modelled (GOL, TCZ, ADA, CTZ). Due to a lack of data regarding duration of treatment for patients receiving cDMARDs, it was assumed that the survival duration for each EULAR response category for bDMARDs would be applicable to cDMARDs. This may be an unfavourable assumption for cDMARDs, which in general are less toxic compared to bDMARDs.

COSTS

Direct drug costs were taken from the BNF January 2013 (BNF65). Illustrative costs are provided in Table 5.16. Often, a manufacturer will negotiate a confidential discount with the

Department of Health, which reduces the treatment cost and therefore increases the likelihood of a positive recommendation by NICE. These discounts are called a Patient Access Scheme. Details regarding the Patient Access Schemes are confidential, and are not applied to the treatment costs shown in Table 5.16 for this reason. However, the Patient Access Schemes are included in the model and subsequent cost-effectiveness results.

Table 5.16: Drug costs for all included treatments

Treatment	Dose regimen	Cost per dose ¹	Cost (first 6 months) ²	Subsequent annual cost
Abatacept (intravenous)	500 mg below 60 kg, 750 mg between 60-100 kg, 1000 mg above 100 kg; 0, 2 and 4 weeks then every 4 weeks thereafter	169.34 (250mg)	£6,350.40	£10,886.40
Abatacept (subcutaneous)	125mg weekly following loading dose 500 mg below 60 kg, 750 mg between 60-100 kg, 1000 mg above 100 kg.	169.34 (125mg)	£8,796.44	£15,778.48
Adalimumab	40 mg; every other week	£352.14 (40mg)	£4,593.45	£9,186.89
Certolizumab pegol	400 mg per week initially, repeated at weeks 2 and 4 weeks followed by a maintenance dose of 200 mg every 2 weeks	£357.50 (200 mg)	£5,440.59	£9,326.73
Etanercept	50 mg; every week	£178.75 (50mg)	£4,663.36	£9,326.73
Golimumab	50 mg below 100 kg, 100 mg above 100 kg, per month	£762.97 (50mg)	£4,557.82	£9,115.64
Infliximab ³	3 mg/kg: 0, 2, 6 then every 8 weeks	£419.62 (100mg)	£6,294.30	£8,222.40
Tocilizumab	8 mg/kg every four weeks	80.38 (80mg)	£5,222.40	£11,673.60
Rituximab ⁴	2000mg every 9 months	£3,492.60 (2000 mg)	£3,492.60	£3,492.60 per dose
Hydroxychloroquine ⁵	6.5mg/kg per day (max. 400mg per day)	£0.17 (400mg)	£31.35	£62.70
Methotrexate	7.5mg per week escalated by 2.5mg per week up to 20mg per week	£0.80 (20mg)	£19.32	£41.57
Prednisolone	7.5mg per day	£1.07 (7.5mg)	£196.25	£392.50
Sulphasalazine	500mg per day escalated by 500mg per week up to 3000 mg per day	£0.79 (3000mg)	£131.38	£290.17
Intensive combination DMARD therapy	Hydroxychloroquine + methotrexate + prednisolone + sulfasalazine (doses as per monotherapy treatments)	NA	£378.31	£786.94
Palliative Care / Rescue Therapy ⁶	N/A	Assumed £60 per month	£360	£720

¹Treatment can be daily or weekly. Assumes weight distribution from the BSRBR and chooses the least expensive way of meeting the dose requirement. No vial sharing assumed. No PAS schemes included. ²No administration or monitoring costs included. ³Assuming 8 doses in year 1 and 6.5 in subsequent years. ⁴Rituximab provided every 9 months. ⁵Using a BSRBR average weight of 73kg for illustration. ⁶An approximation of monthly 'post bDMARD' cDMARD therapy (leflunomide, gold, cyclosporine). N/A = not applicable

As well as direct drug costs, administration costs and monitoring costs were included. Infusions were assumed to cost £154, as reported in TA247.¹⁹⁸ It was assumed that 10% of subcutaneous injections would require administration by a district nurse, costing an average administration cost per subcutaneous injection of £2.61.

Monitoring of treatment toxicity was assumed equal between cDMARDs and bDMARDs. The monitoring costs assumed are provided in Table 5.17.

Table 5.17: Monitoring costs

Monitoring component	FBC ¹	ESR ²	Biochemical profile	Chest X-Ray	Outpatient attendance	Total cost
Cost	£2 ³	£3 ³	£3 ³	£33 ³	£128 ³	
Methotrexate: pre-treatment	1	1	1	1	1	£170
Methotrexate: first 6 months	10	0	10	0	10	£1,700
Monthly monitoring	1	0	1	0	1	£134

¹FBC = Full Blood Count, ²ESR = Erythrocyte Sedimentation Rate

³NHS Reference Costs 2012.

It is plausible that hospitalisation costs increase as HAQ increases, due to diminished functional ability, and damage to joints requiring surgery. Values of the cost per HAQ value were taken from the Norfolk Arthritis Register (NOAR), and their number of inpatient hospital stays and joint replacement surgeries. * These values are provided in Table 5.18.

Table 5.18: HAQ related costs

HAQ score	Cost
0.0 – 0.5	£167.41
0.5 – 1.0	£102.54
1.0 – 1.5	£364.68
1.5 – 2.0	£523.68
2.0 – 2.5	£1,246.26
2.5 – 3.0	£2,687.97

* <http://www.nice.org.uk/guidance/indevelopment/gid-tag313> - data used from Roche submission. Accessed June 2015

UTILITIES

Utility values were estimated each time a HAQ change was simulated. By taking the time between two HAQ changes, the associated utility values were used calculate QALYs via the trapezium rule.

To estimate utility values, a mixture model developed and published by Hernandez *et al.* was used in the DES model, which estimated utility as a function of HAQ, HAQ², pain, age, age² and gender.^{199,200} All of these variables were tracked in the discrete event simulation, apart from pain. Pain was estimated from an independent model based on HAQ. Many models estimate utility from only HAQ, but the evidence from Hernandez *et al.* shows that pain is a significant independent predictor of HRQL.^{199,200}

Hernandez *et al.* fitted the mixture model to data from a US observational database with over 100,000 observations. The model for pain based on HAQ score was developed using the ERAS dataset, and over 13,357 observations. The model is quadratic in HAQ, with HAQ and HAQ² the dependent variables to determine pain.

This two-step approach of calculating pain based on HAQ, and then utility based on HAQ, pain, age, involves significant calculations at each HAQ change event. These events can occur frequently when HAQ progression rates are high. This is potentially inefficient for the patient level simulation.

UNCERTAINTY

Parameter uncertainty was quantified using PSA. This involved assigning probability distributions to parameters and undertaking Monte Carlo sampling from these. Repeated sampling reduces the Monte Carlo error and allows the estimation of the expected costs and QALYs, and also the quantification of parameter uncertainty via Cost Effectiveness Acceptability Curves and the probability of being cost-effective given a particular ICER threshold.

As noted above, many parameters were the coefficients from econometric models. Where possible, multivariate normal distributions were assumed for correlated parameters, otherwise independent normal distributions were assumed

Costs were known with certainty and therefore not subject to probabilistic sensitivity analysis, and the short term effectiveness probabilities from the network meta-analysis were sampled using the CODA (Convergence Diagnostic and Output Analysis) output from the WinBUGS software package.

Numerous methodological and structural changes were tested using scenario analyses. These included using alternative methods to estimate utility values, using EULAR data only (rather than including mapped EULAR from ACR trials), various inclusions and exclusions of particular heterogeneous trials, and alternative discount rates.

5.6 MODEL RESULTS AND DISCUSSION

MODEL RUN TIME AND CONVERGENCE

The model was found to provide relatively stable ICERs at between 3,000-5,000 patient simulations (see Figure 5.5). Two tests of 10,000 simulation patients was undertaken, with a difference of approximately £1,000 per QALY in the ICERs from the tests. 10,000 patients were simulated (taking approximately 1 hour) for the severe population, and 1,000 patients were simulated for the moderate-severe population (also taking approximately 1 hour, due to the high amount of resampling to identify an eligible patient). For the PSA, 1,000 patients were simulated for the severe population, and 100 for the moderate-severe population, and 100 probabilistic samples were evaluated. This resulted in PSA taking approximately 10 hours to run. Undertaking more simulations was not feasible due to the large number of models to run and the time constraints of the NICE appraisal.

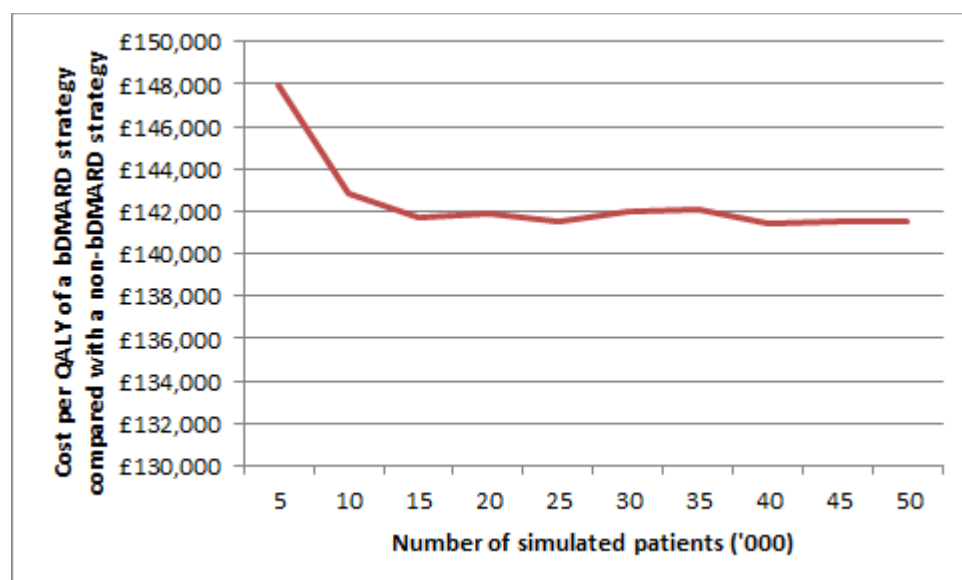


Figure 5.5: Discounted ICER of bDMARD sequence compared to non bDMARD sequence in Population 1 (from Assessment Group report)

RESULTS

The full results are published online and for that reason they are not replicated here.* Also it should be noted that the appraisal is still ongoing, and therefore the results may change if the committee requests changes to the analysis, and future publications from this appraisal may have different results. A summary of the deterministic and probabilistic basecase results and key scenario analyses is provided in Table 5.19 to Table 5.22.

The ICER for bDMARD treatment in Population 2 (severe RA) is approximately £60,000 per QALY gained. In Population 3 (moderate to severe RA) the ICER increases to over £70,000 per QALY gained. If a patient cannot have MTX, then the ICER increases further (Table 5.20), to approximately £90,000 per QALY gained.

The ICER for Population 1 is £300,000 per QALY gained in patients who can receive MTX (Table 5.21), and £400,000 per QALY gained in patients who cannot receive MTX (Table 5.22).

The key parameter within the model that significantly affects the estimated ICER is the method used to estimate HAQ progression whilst on cDMARD therapy. If a linear progression rate is used (as per previous NICE appraisals), then the ICER falls significantly. The ICER for Population 2 in patients who can receive MTX falls to approximately £37,000 per QALY gained (Table 5.19).

* <https://www.nice.org.uk/guidance/gid-tag313/documents/rheumatoid-arthritis-adalimumab-etanercept-infliximab-certolizumab-pegol-golimumab-abatacept-and-tocilizumab-review-assessment-report-addendum2> - Accessed June 2015

Table 5.19: Summarised results: Median ICERS for all bDMARD strategies compared with MTX alone strategy. Populations 2 & 3 who can receive MTX

	Response Measure	Assumed HAQ Progression	Basecase		Scenario analysis						
			RCTs with small %age of bDMARD prior use , adequate MTX-history	RCTs with small %age of bDMARD prior use (irrespective of MTX-history)	Trials with inadequate MTX history	Malottki mapping of HAQ to utility	Discount rates (6% costs, 1.5% QALYs)	Impact of AEs assumed to be 100-fold higher	Relationship between HAQ and pain taken from ERAS	PSA results	
Population 2 (severe MTX – experienced) (ANALYSIS 3)	EULAR	ERAS	£61,200	£61,400	No data	No data	£49,700	£39,500	£62,200	£73,700	£61,700
		Linear	£37,900	£36,300	No data	No data	£32,400	£22,300	£38,300	£46,300	£37,600
	ACR	ERAS	£62,200	£62,200	£62,600	£68,900	£49,700	£39,500	£62,200	£73,700	£62,700
		Linear	£35,500	£35,100	£35,700	£36,400	£30,900	£21,400	£35,600	£43,700	£35,900
Population 3 (moderate MTX-experienced) (ANALYSIS 5)	EULAR	ERAS	£75,000	£74,200	No data	No data	£53,400	£46,600	£78,100	£87,300	£76,800
		Linear	£37,500	£36,600	No data	No data	£31,300	£21,800	£39,300	£48,300	£35,800
	ACR	ERAS	£77,100	£77,500	£77,300	£79,200	£53,900	£48,300	£79,800	£89,300	£79,000
		Linear	£38,000	£36,700	£38,000	£39,200	£30,000	£21,800	£39,100	£46,700	£38,400

All numbers rounded to the nearest £100.

Table 5.20: Summarised results: Median ICERs for all bDMARD strategies compared with SSZ alone strategy. Populations 2 & 3 who cannot receive MTX

	Response Measure	Assumed HAQ Progression	Basecase		Scenario analysis						
			RCTs with small %age of bDMARD prior use , adequate MTX-history	RCTs with small %age of bDMARD prior use (irrespective of MTX-history)	Trials with inadequate MTX history	Malottki mapping of HAQ to utility	Discount rates (6% costs, 1.5% QALYs)	Impact of AEs assumed to be 100-fold higher	Relationship between HAQ and pain taken from ERAS	PSA results	
Population 2 (severe MTX – experienced) (ANALYSIS 4)	EULAR	ERAS	£87,600	£89,000	No data	No data	£71,600	£58,200	£89,100	£107,000	£88,400
		Linear	£39,600	£38,000	No data	No data	£34,800	£24,800	£40,200	£49,200	£39,100
	ACR	ERAS	£94,800	£93,900	£99,600	£94,700	£79,000	£64,700	£97,200	£117,400	£90,000
		Linear	£38,500	£37,300	£37,200	£37,200	£34,100	£23,600	£39,300	£47,800	£38,800
Population 3 (moderate MTX- experienced) (ANALYSIS 6)	EULAR	ERAS	£104,800	£108,100	No data	No data	£74,400	£65,100	£108,700	£121,900	£105,400
		Linear	£41,400	£39,300	No data	No data	£32,800	£23,900	£41,600	£49,700	£41,700
	ACR	ERAS	£106,400	£107,900	£110,500	£107,900	£77,200	£70,000	£105,900	£120,300	£108,200
		Linear	£38,800	£38,500	£38,000	£37,200	£31,100	£23,800	£40,500	£47,100	£39,600

All numbers rounded to the nearest £100.

Table 5.21: Summarised results: Median ICERs for all bDMARD strategies compared with MTX alone strategy. Population 1 who can receive MTX

	Response Measure	Assumed HAQ Progression	Basecase		Scenario analysis				
			RCTs with small %age of MTX prior use	Malottki mapping of HAQ to utility	Discount rates (6% costs, 1.5% QALYs)	Impact of AEs assumed to be 100-fold higher	Relationship between HAQ and pain taken from ERAS	PSA results	
Population 1 (severe MTX – naïve) (ANALYSIS 1)	ACR	ERAS	£308,700	£571,700	£214,800	£185,000	£326,100	£344,800	£295,700
	mapped to EULAR	Linear	£296,300	£432,800	£216,400	£192,900	£323,600	£344,700	£296,700

All numbers rounded to the nearest £100.

Table 5.22: Summarised results: Median ICERs for all bDMARD strategies compared with SSZ alone strategy. Population 1 who cannot receive MTX

	Response Measure	Assumed HAQ Progression	Basecase		Scenario analysis				
			RCTs with small %age of MTX prior use	Malottki mapping of HAQ to utility	Discount rates (6% costs, 1.5% QALYs)	Impact of AEs assumed to be 100-fold higher	Relationship between HAQ and pain taken from ERAS	PSA results	
Population 1 (severe MTX – naïve) (ANALYSIS 2)	ACR	ERAS	£414,700	£140,418	£340,500	£295,400	£382,000	£438,700	£404,500
	mapped to EULAR	Linear	£378,000	£139,800	£357,700	£291,200	£375,300	£460,000	£408,800

All numbers rounded to the nearest £100.

5.7 DISCUSSION

This chapter reports the development of a health economic model for the NICE MTA of bDMARDs for RA. The chapter details the NICE appraisal process and decision problem for this particular MTA. It reports how the health economic model was conceptualised and developed, before provide the basecase results from the evaluation.

The MTA continues to be in development. After the first appraisal committee meeting in October 2013, the standard NICE process would be to publish draft guidance in the form of an Appraisal Consultation Document (ACD). This guidance would be subject to stakeholder comments and review, and often the manufacturers and assessment group are requested to undertake further analysis. The committee would then meet one or more times to develop the final appraisal document (FAD), which provides the final published guidance.

Instead, an ACD has not yet been published. The ICERs initially reported in the analysis by the assessment group (at least £60,000 per QALY for bDMARDs) were above the normal NICE threshold. Following these figures through to a natural conclusion would lead to NICE reversing their original decision to recommend bDMARD therapies for funding by the NHS. Instead, NICE recognised that the key sensitivity in the model estimates was the growth model used to estimate HAQ progression in cDMARD treatment sequences.

The previously used linear method of HAQ progression results in an ICER that is more likely to see a positive recommendation, whereas a more methodologically robust method of latent class growth models resulted in a much higher ICER. The decision was made by NICE to temporarily halt the MTA while they requested further independent analysis regarding the cDMARD HAQ progression modelling.

At this time of writing (May 2015), the second appraisal committee meeting is imminent; however the delay has meant that the final guidance cannot be reported in this thesis.

As reported earlier in the chapter, NICE decided at the scoping stage to not include downstream sequential treatments within the decision problem for the MTA. The reasons for this decision were that the manufacturers did not request an appraisal which focussed on sequential or post-bDMARD use of alternative bDMARDs. Also, the existing size of the MTA appraisal meant it was not seen as feasible to include an attempt to optimise the treatment sequence.

This raises fundamental questions about the objective of NICE and their appraisal process. If 'partial' evaluations are being undertaken, then inconsistent and potentially sub-optimal

guidance will be published. However, on the other hand, if feasible methods exist to inform an optimal or near-optimal treatment sequence, then they have the potential to be utilised within an appraisal process and make a significant positive impact to the development of NICE guidance and the optimal allocation of finite health care resources.

This thesis is therefore well placed to identify and evaluate how untried methods of simulation optimisation may help inform future NICE appraisals and resource allocation decisions.

It should be noted that the model is relatively slow to run. It was developed in Microsoft Excel and requires a large amount of data manipulation to use the right data and parameters for every possible patient population and analysis. From experience, Microsoft Excel is many times slower than a bespoke simulation software package to evaluate a DES model. Rebuilding the model in a bespoke package is likely to be a worthwhile endeavour due to the significant speed-up that would be gained. The slowness of the model meant that precise ICERs were not estimated, due to persisting Monte Carlo error (noise) at 1,000 to 10,000 patient simulations. Therefore the bDMARDs were treated as a class and incremental analyses comparing specific bDMARDs were not seen as being robust. A much faster model would allow the evaluation of more patients and a more precise estimate of each sequence's costs and QALYs.

5.8 CONCLUSIONS

The chapter has highlighted that the process for guidance development used by NICE allows the possibility of sub-optimal treatment sequences to be recommended for NHS funding. Methods which allow the evaluation of all potential treatment sequences may have significant value for the formulation of guidance based on economic evaluation evidence. However, these methods need to be appropriate and feasible within the NICE appraisal process. The following two chapters will seek to identify and implement a simulation optimisation method to inform an economic evaluation of RA bDMARDs. In Chapter 8, the feasibility of these methods within the context of a NICE appraisal will be discussed.

(Blank page)

CHAPTER 6: A CITATION PEARL GROWING SYSTEMATIC REVIEW OF THE METHODOLOGICAL LITERATURE

6.1 CHAPTER OVERVIEW

This chapter reports a systematic review to identify methods which are relevant for a treatment sequencing problem. The chapter begins by framing the treatment sequencing problem, and defining it as a combinatorial discrete simulation optimisation problem. By formally characterising the properties of the problem, the relevance of the methods identified can be judged. The systematic review uses citation pearl growing methods.²⁰¹ This involves identifying key references, and undertaking citation and reference searches to ensure all relevant literature are obtained.

The appropriateness of each identified method is judged by a bespoke framework. The development and theoretical basis of each method is reported. The practical applicability of each method to the treatment sequencing problem is judged. The search methodology is pragmatic, and methods are considered even if not identified in the first instance but become known during the process of searching, synthesis and reflection. This is to account for the fact that the literature is reported across many academic disciplines.

The chapter provides a range of potential methods which could be taken forward for implementation and evaluation in Chapter 7 and Chapter 8. These methods are broad in scope: published in different disciplines and applied to a range of combinatorial problems. Many methods have only been recently developed, and therefore the evidence to support their use is relatively scant. However, there are methods of genuine promise, and numerous routes for further research have been identified. Contrastingly, some methods have been established for a long time, in particular methods applicable to general optimisation which have shown good performance for simulation optimisation methods. The robustness of these methods, including simulated annealing (SA) and genetic algorithms (GA), across a range of problem contexts has been proven.

6.2 INTRODUCTION

This chapter contains a systematic search and review to identify relevant methods for finding an optimal or near-optimal sequence of treatments in an economic evaluation using DES.

Section 6.3 describes the methods used to undertake the systematic review, and Section 6.4 describes the search strategy to identify relevant studies. Sections 6.5 and 6.6 detail how the

quality of the papers was assessed, how data were extracted and how it was synthesised. The results of the search are provided in Section 6.7, and Section 6.8 is a narrative synthesis of the identified methods. Section 6.9 provides discussion and conclusions, including the implications of this chapter for the remaining thesis.

PROBLEM DESCRIPTION

When undertaking a systematic review of methods (in contrast to a systematic review of evidence), there must be a clearly defined problem for which the applicability of identified methods can be evaluated.

The problem to be addressed within this thesis is how to find an optimal sequence of treatments for a chronic condition when a discrete event simulation is required to evaluate the objective function for an economic evaluation. The objective function in this case is the net monetary benefit (NMB). Maximising net monetary benefit represents the optimal configuration of health care resources (treatments) for a particular condition. Therefore this problem can be represented as an optimisation problem, where a configuration of treatments is sought that maximises NMB.

A general optimisation problem:

$$\max_{x \in X} g(x) \quad [6.1]$$

Where $x \in X$ represents a vector of input variables x from the potentially feasible solution space X . Therefore x is a particular permutation of a sequence of treatments from all potentially feasible treatment sequences X . $g(x)$ is the objective function, which cannot be determined directly (analytically or observed), but instead must be estimated via simulation. In the case of an infinite number of simulations, the simulation model provides an estimation of the objective function $g(x)$:

$$g(x) = E_{\omega}[G(x, \omega)] \quad [6.2]$$

The performance measure estimated via the simulation model $G(x, \omega)$ is stochastic, with ω the randomness exhibited in each run of the simulation.

For N simulation runs (i), the sample average is:

$$\bar{G}(x) = \frac{1}{N} \sum_{i=1}^N G(x, \omega_i) \quad [6.3]$$

This can be used as an approximation of the objective function $g(x)$, as $\bar{G}(x) \rightarrow g(x)$ for $N \rightarrow \infty$:

$$E_{\omega}[G(x, \omega)] \approx \bar{G}(x) \equiv \frac{1}{N} \sum_{i=1}^N G(x, \omega_i) \quad [6.4]$$

By the strong law of large numbers (Billingsley 1995) when N is sufficiently large, the sample average can approximate the objective value $g(x)$.

The problem is specifically a discrete combinatorial optimisation problem -the set of available sequences (the feasible solution space) may be very large, but it is discrete and finite. The size of the feasible solution space (\mathbf{X}) is very complex to formally estimate, given the various rules regarding the eligibility of the position of each treatment. However, an upper bound (\mathbf{X}_{max}) can be derived using the following formula for k -permutations of n objects, where k is the length of the sequence (up to a maximum sequence length L) and n is the total number of treatments in a set.*

$$\mathbf{X}_{max} = \sum_{k=1}^L \frac{n!}{(n-k)!} \quad [6.5]$$

For 12 treatments[†] this equates to an upper bound on the feasible solution space (\mathbf{X}_{max}) of over 10 billion unique solutions.

As discussed in previous chapters, full enumeration and comparison of every possible solution (treatment sequence) is not possible when the decision space is large. Therefore this particular decision problem requires a method which can find a good enough solution within a feasible amount of time, rather than a method that can search every feasible solution to find a true optimum. In this instance, whether a solution is good enough is a judgement to be made by the decision maker, however methods can look to ensure that the good enough solution cannot be improved upon and therefore increase confidence that it is a true optimum.

6.3 METHODS

The methods for systematically reviewing published health evidence (for example, health economic evidence, or clinical trial evidence) are well established.²⁰² The classical approach to information retrieval involves matching a search query with the relevant literature. Increasing the sensitivity of the search increases the likelihood of finding relevant literature, but at the cost of finding irrelevant literature. Refining the specificity of the search will reject irrelevant

* (a selection of k objects from a list of n , where $k \leq n$), and where the order of selection matters and selections cannot be repeated.

[†] $n = L = 12$

literature, and by increasing the sensitivity and specificity of the search, strive towards the ‘impossible ideal’ as shown in Figure 6.1.

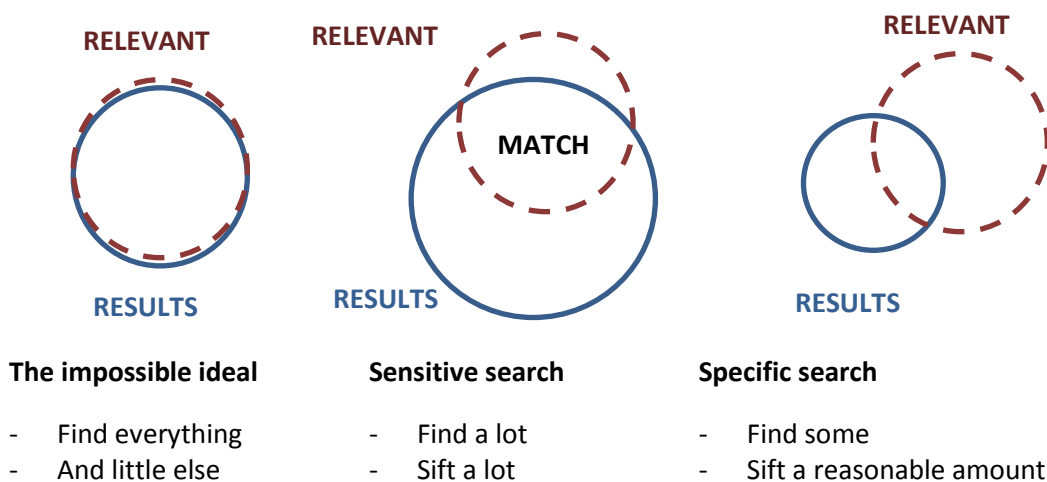


Figure 6.1: Classical search model (Levay 2012)*

Systematic reviews of clinical effectiveness literature compare the results of different clinical studies and take into account the quality of each study. The quality is judged using an explicit framework (randomisation, blinding, allocation etc).^{203,204} However, a systematic review of methodological literature is different from reviewing published effectiveness evidence.^{202,205} In particular, methodological literature may be published in disparate or unexpected fields, may have been applied in one particular problem but may be relevant for another, and may be difficult in general to identify. Most importantly, there is no ‘gold standard’ that different methods can be compared against, and therefore alternative methods must be judged upon other factors, such as their theoretical suitability and their practical suitability. In the book by Black, Brazier and Fitzpatrick (1998), two chapters are dedicated to the issue of searching and reviewing health services research methods: Edwards *et al.* (1998); and Hutton and Ashcroft (1998).²⁰²

The chapter by Edwards *et al.* (1998) proposes that the review of methods must be considered according to an explicit framework, as would be the case with any systematic review of clinical research evidence. Also, the authors propose that any search spans a range of academic disciplines. Because a systematic review of methodological literature is influenced by the topic of interest, there is no ‘best practice’ set of methods or processes for the review. Instead, Edwards *et al.* (1998) propose that any systematic review of methods should be ‘objective’ and

* <http://www.ihs.manchester.ac.uk/events/pastworkshops/2012/ESN131212/paullevay.pdf> - Accessed June 2015

that the process by which literature are obtained and synthesised should be methodological and explicit.²⁰²

Therefore, a framework for this systematic review of methods has been developed. This framework contains three key factors of interest. Firstly, it is important to understand whether the method was developed specifically for a simulation optimisation context, and also whether it was developed for a discrete and combinatorial problem. If the method is an adaptation of an alternative method, then it is important to understand how the method has been adapted. Secondly, the theoretical basis of the identified method contains the key assumptions, limitations and possible biases associated with the method. Understanding these ensures that the method is suitable for application within the problem context. Thirdly, the practical applicability of an identified method considers how it performs when used to solve a real-world problem. Although the review is focussed on methodological papers, any applications of a particular method that are identified will help inform the suitability of a method in practice. These factors are summarised within the review framework in Table 6.1.

Table 6.1: Framework for the methods review

Framework factor	Issues to consider
Development	What problem was the method originally developed for? Has the method been adapted from its original context? Was the method designed to address discrete event simulation (DES) optimisation? Or is it a general optimisation method that could be suitable for DES?
Theoretical basis	How does the method address optimising a DES with a combinatorial problem? What assumptions does the method require? What are the theoretical limitations of the method? What are the potential biases associated with the method?
Practical applicability	Has the method been used to optimise a DES with a combinatorial problem? If so, how did it perform? (Speed, optimality, ease of implementation) If not, are there any suggestions to its practical applicability?

The use of a bespoke framework for classifying the results and evidence introduces bias, due to there being many different ways to classify information. In particular, subjective issues regarding the limitations and biases of an identified method may be challenging to classify and summarise, and therefore a framework ‘cannot be completely impartial.’²⁰² In an ideal situation, multiple researchers would be employed to identify and classify information using the predefined framework. However, in this PhD thesis it was not possible to employ multiple researchers to ensure that inter-observer reliability was maintained.

6.4 SEARCH STRATEGY

A systematic review of methodological literature may cover a number of different disciplines.²⁰² For a systematic review of discrete event simulation optimisation methods for combinatorial problems, it is expected that methods may emerge from operational research, computer science, mathematics, and other academic areas. Therefore a systematic search must be performed across all potentially relevant disciplines.

Rather than a systematic review of clinical trials, where each trial must be identified, a systematic search of methods must identify all appropriate methods, rather than all published instances of each appropriate method. Therefore rather than a global search identifying an infeasible amount of potentially relevant studies, an iterative search focussing on citations and references was assumed to be valid in this instance. Edwards et al. (1998) refers to this approach as 'theoretical saturation', where an iterative approach is used due to the unknown quantity of relevant literature. In this model of searching, methodological topics tend to frequently return a large quantity of theoretical articles, and the marginal benefit of adding further articles decreases rapidly beyond a certain point. The model focuses on truncating the search when new information is not forthcoming, rather than pursuing every last possible reference. The model therefore needs a net that is cast wide across many types of literature, to ensure that relevant methods are not missed.

While the classical approach to searching is perfectly acceptable for a tightly defined set of information, they rely on pre-defined queries which suggest that there is a subset of knowledge which can be defined as "all the relevant evidence" (see Figure 6.1). More complex queries or evidence searches (including 'dynamic queries') allow new questions and answers to emerge from the evidence.

This breaks down the process of searching for evidence, and allows reflection and thinking to occur based on the documents retrieved before either varying the next search query, or deciding that information saturation has been reached. It is also a pragmatic process, which avoids an unmanageable volume of results.

Citation pearl growing is a method of searching citation indexes to iteratively explore the published evidence. It is particularly useful where terminology or indexing to categorise evidence varies (perhaps across fields or disciplines), which has significant benefits for searching methodological articles.²⁰⁵ Pearl growing is similar to qualitative research methods, where key documents (called 'pearls') are identified and then references citing these documents are also reviewed to assess their relevance. However, as with qualitative research,

the method relies on the prior selection of key records to begin the search process. The process is also dependent on relevant literature being relatively well cited.

SEARCH

A search was designed based on specific terms already identified in the topic area. These terms were generated from the title, abstracts and keywords of already identified relevant papers. In particular, Andradottir's (2006) review of simulation optimisation via random search review paper,²⁰⁶ and Fu's (1994) review of simulation optimisation.²⁰⁷ These search terms are shown in Table 6.2. The electronic search was conducted in March 2014 within the ISI Web of Science (ISI WoS) database. All databases within ISI WoS were searched including the Science Citation Index and the Social Science Citation Index. ISI WoS was selected due to its excellent coverage across the full spectrum of science, including the social sciences, computer science and technology. It is an index of over 5000 journals, 54 million records which span 100 years.

Table 6.2: Search terms

Search step	Search terms	Results
1	TITLE: (optimi* AND simulation)	8,996
2	TOPIC=(method* OR approach)	51,549,272
3	TOPIC =(simulation)	3,505,384
4	TOPIC=(optimi*)	2,138,034
5	TOPIC: (discrete OR combinatorial)	824,381
6	#1 AND #2 AND #3 AND #4 AND #5	462
7	#6 Refined by: DOCUMENT TYPES=(ARTICLE OR REVIEW OR BOOK)	170

Timespan = All years

ISI Web of Science database (all databases)

Search language = English

INCLUSION/EXCLUSION CRITERIA

From the search, abstracts of identified papers were screened before ordering the full-text papers of relevant citations. The inclusion and exclusion criteria used for the screening of relevant papers are provided in Table 6.3.

Table 6.3: Inclusion and Exclusion criteria

Inclusion criteria	Exclusion criteria
- Methodological papers relevant to the problem formulation	- Papers describing or implementing a method which is not relevant to the problem formulation
- Methodological papers which are potentially generalisable	- Conference abstracts or editorials
- Papers applying a method to a broadly similar problem	- Methods not applicable to a simulation approach to estimate the objective function
	- Methods for obtaining a local optima

- Statistical methods for small problems, or to inform the comparison of simulation output
- Methods solely for multi-objective problems
- Naïve methods (enumeration, trial and error)

CITATIONS AND REFERENCES

The identification of papers from the search was confirmed by searching the references and citations of each paper. By searching citations, any future development of methods could be identified. By searching the references, it allowed the identification of the methods' origins, or alternative methods. This process allowed the set of pearl papers to be finalised.

6.5 QUALITY ASSESSMENT

A published criterion for critically appraising the quality of identified methods for this review was not identified. Therefore the framework presented in Table 6.1 was used as a quality assessment tool. The synthesis of each identified method includes a critical appraisal based on the application of this framework.

6.6 DATA EXTRACTION AND SYNTHESIS

Data extraction for the review was based upon the framework presented in Table 6.1, and is provided in Table 6.4. Full evidence tables are contained in Appendix B.4. A narrative review methodology was used to synthesise the details of each identified methodology. A narrative review provides a discussion and summary of a particular topic, and the review framework was used to provide a structure to this narrative synthesis. Where identified methods were broadly similar (or a modification) then they were grouped together.

Table 6.4: Data extraction form

Reference	Details
Development	
What is the method?	
What problem was the method originally developed for?	
Has the method been adapted from its original context?	
Was the method designed to address discrete event simulation (DES) optimisation?	
Or is it a general optimisation method that could be suitable for DES?	

Theoretical basis	
How does the method address optimising a DES with a combinatorial problem? How does the method work? What assumptions does the method require? What are the theoretical limitations of the method? What are the potential biases associated with the method?	
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem? If so, how did it perform? (Speed, optimality, ease of implementation) If not, are there any suggestions to its practical applicability?	

6.7 SEARCH RESULTS

The initial search of the database identified 170 citations. 84 of these citations were excluded immediately for not being a full peer-reviewed journal article, or not being English language. Of the remaining 86 citations, 49 were excluded and a full list of citations and reasons for exclusion is provided in Appendix B.1.

37 full articles were retrieved for assessment. 21 were excluded and a full list of articles and the reason for exclusion is provided in Appendix B.2. The remaining 16 articles (pearls) were included in the review, and a reference and citation search was applied to these articles. 16 citations/references were of interest and full papers were ordered. Nine of these articles were excluded, and full details are provided in Appendix B.3. The seven remaining articles were added to the original 16 pearl articles. Four reviews were cross-checked to identify any relevant articles that may have been missed by the database search.^{206–209} No relevant articles were identified. The full process of sifting and exclusion is detailed in the QUORUM diagram in Figure 6.2.

In total, 23 papers were identified which either developed or applied a method for a combinatorial simulation optimisation problem. The references of the 23 studies included in the systematic review are provided Appendix B.3.

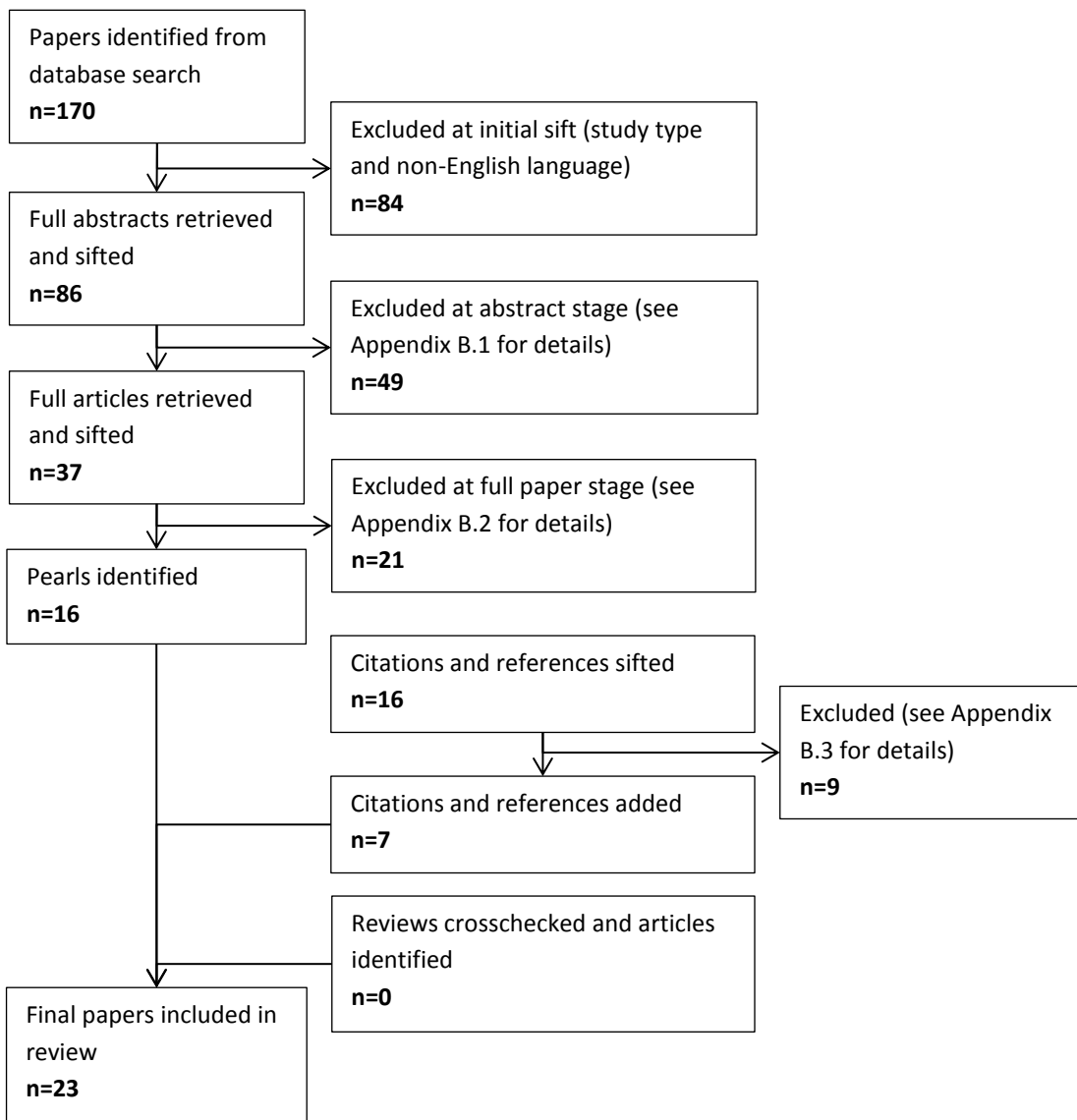


Figure 6.2: QUORUM diagram for methods review

6.8 NARRATIVE SYNTHESIS OF IDENTIFIED METHODS

This section of the chapter classifies and describes the methods identified. For each method, the development, theoretical basis and practical applicability to the treatment sequencing problem are described in detail. Methods of most relevance to the treatment sequencing problem (i.e. a combinatorial problem with discrete parameters and a large, finite decision space which is evaluated using a discrete event simulation model) are the focus of this section, because these will potentially be taken forward for implementation in Chapter 7. Identified methods which are not deemed to be as relevant are described in less detail. The reasons for why they are not deemed to be as relevant or appropriate are clearly specified.

Full tables reporting data extraction from all 23 studies are provided within Appendix B.4. Many of the studies identified used similar methods. Where the methods used overlap across studies, then some studies are discussed in less detail than others to avoid repetition. Also, many methods for simulation optimisation are variations or applications of established optimisation methods which do not require a simulation model for solution evaluation. The papers included within this review are methods specifically for simulation optimisation, however much of the development or theoretical basis for these methods may come from general optimisation methods. Where required, reference will be made to these.

OPTIMISATION

Optimisation is the task of making the best decision among various alternatives.⁵⁵ An introduction to optimisation is provided in chapter Section 2.6, and the treatment sequencing problem is defined as an optimisation problem in Section 6.2. The treatment sequencing problem can be defined as a discrete optimisation problem, as opposed to a continuous optimisation problem, because alternative solutions are determined by discrete variables. The problem is also finite in size, because there is a limited, albeit very large, set of available treatments, providing a finite limit on the number of feasible sequences. The number of feasible solutions is defined as the potentially feasible solution space \mathbf{X} , where each potential solution is $x \in \mathbf{X}$

The global optima within a feasible search space is the configuration of each design parameter which maximises (or minimises) an objective function. There may be one global optimum, or many configurations which are all global optima. Local optima represent the best solution within a particular local neighbourhood of potential solutions. Figure 6.3 provides a representation of a continuous search space which contains local and global maxima, where $g(x)$ is the objective function value of vector of input variables x from the potentially feasible space \mathbf{X} .

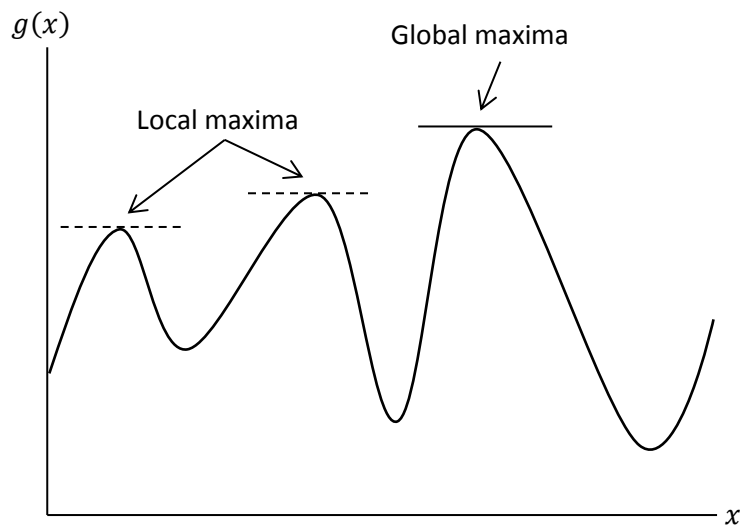


Figure 6.3: Local and global optima

LOCAL SEARCH FOR OPTIMISATION

If the feasible search space is small enough, then a brute force (or exhaustive) search algorithm can be used to systematically enumerate all possible solutions, and therefore identify the optimal.*

However, many optimisation problems are combinatorial (including the treatment sequencing problem), which means that the feasible space of potential solutions is finite (but often very large).²¹⁰ As the size of the problem increases (in terms of number of input variables), the computational difficulty of enumerating each solution increases. These problems are computationally hard to solve,[†] and therefore beyond small-sized problems, full enumeration is unlikely to be feasible.²¹¹ In this situation, a local search method can be applied to identify a good enough solution within a feasible amount of time.

Local search (LS) involves the movement from solution to solution by applying local changes. Specifically, LS looks for a nearby solution which is better or as good as the current solution. The algorithm continually makes moves to better solutions until no further improvement can be found. The key to the performance of LS is the neighbourhood function which determines how to identify a nearby solution. Formally, the neighbourhood function is $N(x)$, where $x \in \mathbf{X}$. For a minimisation problem, a solution x , is a local optimum with respect to the neighbourhood function N , if $f(x) < f(y)$ for every y in $N(x)$. Neighbourhood functions are often problem specific and the performance of a search can be closely related to the specification of the neighbourhood function.

* More formally, all optima

† NP-complete or NP-hard, as defined by computational complexity theory

LS will find a local optimum, but they are unlikely to find a global optima if local optima are present. For example a hill climbing algorithm is a very simple form of LS. The algorithm starts with a randomly selected solution. Incremental changes to an element of the current solution are made until the change produces a better solution, at which point the new solution is accepted. The process is then repeated until no further improvements can be found.

Because only local optima can be guaranteed, many LS methods incorporate modifications to overcome this, which allow them to become global search methods. These modifications include repeated local search, which just repeats the LS with numerous starting points, or an iterative local search, which allows the algorithm to jump to another point in the search space when a local optimum is found. Alternative modifications have focused on the ability to randomly allow worsening moves to be accepted, a form of stochastic optimisation (more detail in the following sub-section). This allows the algorithm to avoid becoming trapped in a local optimum and to continue to seek a global optimum.

SIMULATION OPTIMISATION

Simulation optimisation was the most commonly used term to present the use of a simulation model to evaluate the objective function for an optimisation problem. Andradottir (2006) defines simulation optimisation as *"...a special case of stochastic optimization where the required objective function values $g(x)$ are estimated via computer simulation, and hence involve some noise."*²⁰⁶

Stochastic optimisation is defined as optimisation of a problem where there is random noise in the measurement of the objective function $g(x)$. Somewhat confusingly, stochastic optimisation is also used to define LS algorithms where there is a random (often Monte Carlo) choice made in the search direction as the LS algorithm iterates.²¹² As was explored in the previous section.

Many of the identified simulation optimisation methods identified in this review met both of these definitions. They used a simulation model to determine an approximate estimate of $g(x)$, and they used a stochastic optimisation process which allowed worsening solutions to be accepted during the search.

Many heuristics and search methods are designed for deterministic optimisation. These are general optimisation methods, and may be applied to a problem where there is an analytic value of the objective function. For example the distances between cities for a travelling salesman problem are known. Therefore the objective function value (the total route distance) can be calculated with certainty.

When used in a deterministic setting, a stochastic optimisation algorithm will return the solution identified with the best performance as the 'best solution'. However, within simulation optimisation, it is common for the best performing solution to be attributed to the solution with the best sample mean objective value. Simulation optimisation may therefore attribute a wrong best solution, due to stochastic or Monte Carlo error.

With a large combinatorial search space, it is not possible to guarantee that an identified solution is the global optimum. In particular, for a stochastic simulation setting, even if the search algorithm does visit the true global optimum, there is no assurance that the algorithm will correctly identify this, due to the stochastic nature of the objective function. For this reason, simulation optimisation methods often have to balance a search/selection trade-off.²¹³ If there is a fixed computation budget, how to allocate computing resources between searching the feasible space for better solutions, and evaluating the performance of each solution (via a computationally expensive simulation model) to ensure the search is sensible and results are useful for decision-making purposes?

Taxonomies of simulation optimisation methods are provided in Appendix B.5.

METHODOLOGICAL STUDIES AND CATEGORIES

The papers identified represent a body of literature concerning simulation optimisation of a combinatorial problem. Finding key methodological papers within this literature was a difficult task, because many methods were first developed for optimisation problems without simulation required to evaluate performance, or for continuous optimisation problems which were then applied to combinatorial problems. Therefore the literature identified studies which either applied established methods, or adapted established methods, as well as novel methods. Within this review, the development of each discussed method is detailed, including the background to established methods where necessary. However, the focus remains on the suitability of all methods to our particular treatment sequencing problem, and the practical applicability of each method revealed by the reviewed studies.

The 23 identified studies can be classified into broad method areas and specific method categories. This classification is detailed in Table 6.5. These classifications are used to guide the narrative review. Where several identified studies are within one class, the review is much more detailed, compared with classes with just one study. This is to ensure that the differences in approaches and specific details are emphasised, while minimising repetition and unnecessary detail. The review provides detail about the background, development and practical applicability for each identified method. Details regarding how each method works are provided in Appendix B.6.

Table 6.5: Methodological studies and categories

Class	Category	Specific methodology	Brief details	Papers included*
Random Search	Random Search	Random search hill climbing	A stochastic optimisation method where neighbourhood solutions are selected using a probabilistic method.	Jacobson et al. (1998) ²¹⁴ Kamrani et al. (2012) ²¹⁵
	Adaptive Random Search	Balanced Explorative and Exploitative Search	An adaptive random search method where a switch between local and global search is incorporated into the algorithm	Andradottir et al. (2009) ²¹⁶
		Convergent Optimisation via Most-Promising-Area Stochastic Search	A random search algorithm with a unique neighbourhood function	Hong et al. (2006) ²¹⁷ Huang et al. (2012) ²¹⁸
Metaheuristics	Simulated annealing	A stochastic metaheuristic which mimics the annealing process of a crystalline solid. It accepts worsening random moves with a decreasing probability based on cooling the 'temperature' parameter	Ahmed et al. (1997) ²¹⁹ Ahmed et al. (2002) ²²⁰ Alrefaei et al. (1999) ²²¹ Ghiani et al. (2007) ²²² Haddock et al. (1992) ²²³ Lacksonen et al. (2001) ²²⁴ Rosen et al. (2005) ²²⁵	
	Genetic algorithms	Population based metaheuristics using crossover and mutation to replicate evolution and natural selection	Ding et al. (2005) ²²⁶ Jun et al. (2010) ²²⁷ Korytkowski et al. (2013) ²²⁸ Lacksonen et al. (2001) ²²⁴ Yang et al. (2007) ²²⁹	
	Tabu search	A metaheuristic that forbids movement to recently visited (tabu) solutions	Azadeh et al. (2010) ²³⁰ Yang et al. (2004) ²³¹	
	Ordinal Optimisation	A metaheuristic to identify solutions with a high probability of being 'good enough'	Ho et al. (2000) ²³²	
	Nested partitions	A metaheuristic where the feasible region is partitioned and searching focusses on regions of most promise	Shi et al. (2000) ²³³	
	Particle Swarm Optimisation	A population based metaheuristic where the direction of movement through the search space is influenced by the current best solution	Kuo et al. (2011) ²³⁴	
Hybrid and other methods	Averaging framework for simulated annealing	A variation of simulated annealing which records the performance of previous solutions to estimate how much simulation effort is required	Prudius et al. (2012) ²³⁵	
	Empirical stochastic branch-and-bound	A hybrid method of nested partitioning and branch and bound.	Xu et al. (2013) ²³⁶	

*some studies repeated if include multiple methods

Random search

Random search methods are a broad class of optimisation techniques. There is a general distinction between two main types of random search – traditional random search, and adaptive random search. These two types of methods are used to provide the structure in this section of the narrative review. Also, random search methods are the foundation for many of the metaheuristic methods which will be reviewed in the subsequent section of this chapter.

Search results

Two studies were identified which used a traditional random search method for a combinatorial simulation optimisation problem.^{214,215} Four studies were identified which used an adaptive random search method for a combinatorial simulation optimisation problem.^{216–218,236}

Traditional Random Search

Development

Traditional Random Search (RS) methods were in general developed to solve deterministic optimisation problems, where there is no uncertainty in the value of the objective function for a given solution.²⁰⁶ In particular, they were developed as a gradient free method, which enables use for non-differentiable discrete and continuous problems. Early development coincided with the development of computers, and Rastrigin is often attributed with the first use of the term ‘random search’.²³⁷

Practical applicability

Two studies have reported the use of RS for a combinatorial simulation optimisation problem.^{214,215} Jacobson et al. (1998) use RS for a discrete manufacturing process design optimisation problem.²¹⁴ They apply generalised hill climbing algorithms which incorporate different random variables to determine the selection of an inferior solution. Each random variable has a different mechanism for accepting an inferior solution at a particular iteration, and these are provided in Table 6.6). For instance, one algorithm is a local search, where the probability of accepting a worse solution is always returned as 0. This is in contrast to the Monte Carlo search (often called a random walk). Here, the algorithm will always accept any neighbouring solution, irrespective of its performance.

One of the random search algorithms applied is a simulated annealing metaheuristic. However, the details are limited and therefore this method is not formally reviewed here, and simulation annealing is reviewed within the metaheuristics section of this review.

Table 6.6: Jacobson et al. algorithm formulations²¹⁴

Generalised Hill Climbing algorithms	Parameter(s) determining the acceptance of an inferior solution
Simulated Annealing	$-t_k \ln(u)$
Threshold accepting	Q_k
Monte Carlo search	1
Local search	0
Weibull accepting	$-t_k (-\ln(u))^{1/\alpha}$

Where t_k is a temperature parameter, α is a shape parameter, Q_k is a threshold constant, $u = U(0,1)$ uniform variable. More details regarding Simulated Annealing provided in the Metaheuristics section

The study incorporated three different neighbourhood rules, which are used to estimate a neighbouring solution from a current solution.

The study found that simulated annealing, threshold accepting and Weibull accepting methods all found comparable results which were superior to Monte Carlo search and local search. The local search results tended to yield higher variance. The first neighbourhood rule was very conservative, and took a large number of iterations before the stopping rule was met. The third rule was aggressive and often terminated very quickly. The authors conclude that rule two provided an acceptable balance between time taken to run, and the quality of the solution found. The study is limited in providing information about convergence and stopping rules, along with the computational burden of the problem and efficiency of the algorithms. However, they conclude quite positively by stating that these results are a useful and practical tool for a complex manufacturing sequencing problem.

Kamrani et al. (2012) use RS for a business process optimisation problem.²¹⁵ The problem involves finding the most beneficial assignment of tasks to agents. Tasks can be defined as critical and non-critical, and assignments of workers to tasks must avoid invalidating a pre-defined work process. In one formulation of the problem, assignment of any task to any agents does not affect the flow of the business process, and the Hungarian algorithm is applied.* In a separate formulation of the problem, the assignment of tasks does affect workflow, and a RS heuristic is applied to solve this. The RS algorithm is specifically a hill climbing algorithm, where any improving move is accepted (irrespective of being the best move in the evaluated neighbourhood), and no worsening moves are accepted.

* The Hungarian method is an algorithm that solves the assignment problem. It was developed by Harold Kuhn in 1955, and is named after two influential Hungarian mathematicians - Kőnig and Egerváry

The random search method was applied with three initial solutions. The algorithm reached near optima after 80 iterations. The relative deviation for a number of problem sizes is less than 0.5% from the analytically proven optimal value. The authors noted that the algorithm showed good performance in their problem context, but is not generalisable to combinatorial problems outside of assignment problems. This is because it applies rules based on a determination between critical and non-critical tasks in the problem. The optimisation process is one relatively minor component of the overall study, and therefore specific details are relatively brief.

Adaptive Random Search – Introduction

Adaptive Random Search (ARS) is a search method which is designed to address the limitations of having a fixed neighbourhood structure. In particular, when an algorithm revisits a solution, the candidate solutions are drawn from the same neighbourhood. ARS looks to change the neighbourhood structure based on information generated during the algorithm process. With an ARS, the neighbourhood will in general shrink as information regarding the objective function is gathered. In particular, many problems will have a cluster of good solutions within an particular area. Algorithms with an adaptive neighbourhood structure will often perform better than those with a fixed neighbourhood structure.²¹⁷

The method was initially developed for continuous optimisation problems, and research involved experimenting with neighbourhood structures (then referred to as ‘variable step sizes’) throughout the 1960s and 1970s. Schumer and Steiglitz developed the ‘Adaptive Step-Size Random Search’,²³⁸ which was extended by Kregting and White and their ‘Adaptive Directional Random Search’.²³⁹

Search results

Three studies used an adaptive random search method for a combinatorial simulation optimisation problem.^{216–218} One study developed the Balanced Explorative and Exploitative Search (BEES) framework.²¹⁶ The two remaining studies report the development and application of the Convergent Optimisation via Most-Promising-Area Stochastic Search (COMPASS) adaptive search method.^{217,218}

Adaptive Random Search – Balanced Explorative and Exploitative Search (BEES)

Development

The BEES algorithm was developed by Andradottir & Prudius,²¹⁶ with the aim of providing almost surely* convergent random search algorithms which are simple and general enough to provide applicability to a range of combinatorial simulation optimisation problems.²¹⁶

Practical applicability

The algorithms have only be applied to test functions, and not to a real world combinatorial simulation optimisation problem. The performance of these algorithms appears to be promising, and convergence within a finite search space has been proven by the authors under certain conditions. A limitation is that the methods all required user-specified tuning parameters (for the deterministic, stochastic and adaptive variants of the Adaptive Random Search methods), and these parameters require extensive experimentation.

Adaptive Random Search – Convergent Optimisation via Most Promising Area Stochastic Search (COMPASS)

Development

Convergent Optimisation via Most-Promising-Area Stochastic Search (COMPASS) is a random search algorithm with a unique neighbourhood function. It was developed by Hong & Nelson (2006).²¹⁷ The method is therefore relatively new and unproven, but has already garnered interest in the simulation optimisation field.²⁴⁰

Practical applicability?

In the original paper by Hong & Nelson (2006), the algorithm was applied to an assemble-to-order manufacturing problem, and not to a combinatorial simulation optimisation problem.²¹⁷ This particular problem uses mixed value decision parameters and therefore is not strictly included within the parameters of this systematic review. However, the algorithm showed good performance within this particular problem type.

Huang et al. (2012) applied the COMPASS algorithm to a vehicle allocation problem. The application was for the purpose of testing the method, rather than solving a real work problem. The problem is to maximise the throughput of vehicles through an intrabay system for semiconductor manufacturing. 11 vehicles can be allocated to 10 intrabay systems. As the algorithm iterates, the same size for each solution in the visited-solution set is increased. This increases the precision of the objective function estimates. From an initial solution, an extra 10

* Almost sure convergence implies that the probability of convergence on a target value (the global optima) is 1.

designs are sampled. The best solution is selected and the algorithm iterates. The algorithm converges to the optimal solution after only 30 iterations of the algorithm.

A limitation of the study is that the feasible space is not constrained, and therefore the simulation model is simplistic and not applicable to the real life problem the authors look to solve. Also, little information is given about the relative efficiency of the method, against other more established search algorithms.

METAHEURISTICS

The majority of methods identified fall under the category of metaheuristics. These methods are summarised within this section. To define a metaheuristic, a definition of a heuristic is required.

A heuristic is simply a method of finding a solution to a particular problem. They are designed for a specific problem type, and exploit the particularities of a problem which enables a more efficient search for a solution. A heuristic may require the trading off between the time it takes to execute, and the accuracy of the solution found. Heuristics may be compared to exact methods. While an exact method may guarantee a proven optimal solution within a finite period time, in reality this time may be prohibitively large. Instead, heuristics are developed to find a 'good enough' solution within a 'small enough' period of time.²⁴¹

An example of a heuristic is the greedy algorithm, which simply selects the optimal local choice at each stage of the algorithm. For example, the classic travelling salesman problem (TSP) requires the shortest possible route between a network of cities to be found. Each city must be visited exactly once and the route must end where it begins (e.g. a closed loop). A greedy algorithm will select the optimal local choice, which in the case of a TSP is the nearest unvisited city. Because TSP is an NP-Complete problem*, the optimal solution for a relatively small problem is intractable. The greedy algorithm cannot guarantee to find the optimal solution, but by finding a 'good enough' solution, it is commonly applied in many situations.²⁴²

Metaheuristics are problem-independent techniques which may not guarantee an optimal solution. In contrast to heuristics, metaheuristics are not specific to any particular problem. Instead, they offer the capacity to be applied to a wide range of problem types. Metaheuristics are a higher-level procedure, meaning they *'...provide a set of guidelines or strategies to*

* NP-Complete is a computational complexity theory classification for particular decision problems. NP-Complete problems are both NP and NP-Hard, where NP refers to 'nondeterministic polynomial time'. A particular characteristic of these problems is that no fast solution is known, because as the input size increases, the algorithm time required is superpolynomial (for example, exponential).

*develop heuristic optimisation algorithms.*²⁴¹ Metaheuristics are therefore guidelines to follow when designing a search method to solve a particular optimisation problem.

Because metaheuristics look for a good enough solution within a relatively small period of time, they are not subject to combinatorial explosion, where the time required for an algorithm to find an optimal solution for an NP-hard problem increases exponentially with the problem size.

All metaheuristics can be defined on the basis of five components:

1. Representation
2. Evaluation function
3. Neighbourhood relation
4. Search process
5. Mechanism for escaping from local optima

Representation

The representation (or encoding) of a particular solution is fundamental to metaheuristic methods. Metaheuristics require a solution to be contained as an object within the computer program. It must be possible to manipulate the object using the different operators applied by the metaheuristic. Binary encoding, integer encoding, and permutation encoding are all common representation types for combinatorial problems, with different types being more appropriate for particular problems. For example, binary encoding may be more appropriate for knapsack and Boolean Satisfiability (SAT) problems, where a decision variable is yes/no, include/exclude. Integer encoding may be more appropriate for an assignment problem, where the value and order within a solution may represent tasks and resources applied for each task. Permutation encoding may be more appropriate for TSP and sequencing problems, where the order of elements (cities) is implied in the solution, and each element (city) is uniquely represented by an integer. The representation of a solution must retain the ability for the metaheuristic to change the solution, but also be coded efficiently.

Evaluation function

The evaluation function provides an indication of the quality of alternative solutions, and allows better and worse solutions to be distinguished. Many metaheuristics require the magnitude of change in competing solutions to be estimated (that is, precise estimates of the performance of two solutions to allow an accurate estimate of the difference), but some methods can work based on ordinal ranking of solutions (see the Ordinal Optimisation section). There may be single or multiple objectives within the optimisation problem which

each require evaluation. The evaluation function may be exact, or it may be approximated (e.g. approximation models) or estimated (e.g. simulation models).

Neighbourhood relation

The neighbourhood is a set of solutions that can be reached by a simple operator (often defined the ‘move operator’). If X is the solution space, and $x \in X$ a particular solution, then the neighbourhood function for $x \in X$ is denoted $N(x)$. Neighbour solutions are expected to provide similar solutions in terms of their performance. However, for discrete problems this is rarely the case, and therefore the neighbourhood relation requires particular consideration given the problem type and method of representation.

With a binary representation of a solution, a neighbourhood function may be flipping of one bit in the solution array (called a ‘bit flip’). The neighbourhood size is equal to the solution size. Representations and common move operators are demonstrated in Table 6.7, along with the corresponding neighbourhood size.

Table 6.7: Neighbourhood size and move operators

Representation		Example	Details	Neighbourhood size
Binary representation		10010 → 00010	Flip one bit in the solution vector	Binary vector of size n , neighbourhood size of n
Integer representation		57664 → 27664	A discrete value replaced by another character in the set	If set of size k , and vector size n , then neighbourhood size $(k-1)n$
Permutation	Adjacent pairwise interchange	51432 → 15432	Swap two adjacent elements	Permutation size n , neighbourhood size $n-1$
	Insertion operator	51432 → 54312	Select an element and insert in another position	Permutation size n , neighbourhood size $n(n-1)$
	Exchange operator	51432 → 51234	Two selected elements are swapped	Permutation size n , neighbourhood size $n-1$
	Inversion operator	51432 → 52341	Select two elements and exchange the sequence between them	Dependent on element selection distance

Search process

The search process within a metaheuristic represents the lower-level heuristic (a perturbation) chosen to determine a step to a neighbouring solution. Alternative processes include ‘best

improvement', where all neighbouring solutions are evaluated and the step with the best improvement (steepest descent/ascent within the neighbourhood) is accepted. Also commonly used is a 'first improvement' (gradient descent/ascent) process, where neighbouring solutions are evaluated and the first solution to offer an improvement is accepted. Finally, a random selection process can be used, where any neighbouring solution is selected at random, irrespective of if it is an improvement.

It is important to note that the best improvement search process will find the local minimum, but it requires all possible solutions in the neighbourhood to be evaluated before the algorithm can iterate to that identified best solution. For a metaheuristic, it is important to balance the search process between exploitation of the best solutions found (referred to as intensification around a local optima) and exploration of the whole search space (diversification). A search process which is extreme in terms of intensification will only accept an improving solution, whereas a process extreme in terms of diversification will accept any solution.

Mechanism for escaping from local optima

There are two broad approaches to avoid an algorithm being trapped within local optima. The first is to repeatedly restart the algorithm with alternative starting points/solutions. The second is to allow moves to inferior solutions to be accepted. Allowing non-improving moves allows the search to escape a local optima, and this is often achieved by introducing a stochastic process which allows a non-improving move to be accepted with a given probability; and these are therefore a subset of stochastic optimisation methods. A deterministic algorithm will always find the same solution with a given starting solution. However, a probabilistic/stochastic algorithm may report a different solution from an identical starting solution.

Memory and solutions

There is a distinction between metaheuristics which require a memory of previously visited solutions, and those which only require the current best solution. Tabu search and particle swarm optimisation are two methods with a memory based structure to the algorithm, and use the memory of previously visited solutions to guide the algorithm through the search space.

Also, in each iteration of a metaheuristic algorithm, there could be a single solution being considered, a population of solutions, or a set of local neighbouring solutions. These differences have important implications for the success of the algorithm, as well as the computational burden. Single solution searches include simulated annealing and tabu search,

population searches include genetic algorithms and particle swarm optimisation. Nested partitioning is a metaheuristic focussed on searching a current best neighbourhood/set. A table of methods classified by their memory structure and search type is provided in Table 6.8.

Table 6.8: Metaheuristics - memory structure and algorithm type

Algorithm type	Memory structure	
	No memory	Memory
Single	Simulated Annealing, Greedy randomised adaptive search procedure Iterative local search	Tabu Search
Population	Genetic and evolutionary algorithms	Particle Swarm optimisation Ant Colony optimisation
Set	Nested partitioning	-

Search results

Metaheuristics emerged from the search as the predominant class of method for simulation optimisation problems. 16 of the 28 (57%) identified studies were metaheuristic methods. Of these 16 studies, seven (44%) were simulated annealing methods, and five (31%) were genetic and evolutionary algorithms. See Table 6.5 for the full breakdown of studies and method types.

It is important to note that none of the metaheuristic studies identified for simulation optimisation of a combinatorial problem were specifically new methods designed especially for the problem. In fact, the majority of studies were applications of established metaheuristics, with only four adapting the method to fit the combinatorial or simulation aspect of the problem.^{219,220,222,230}

Genetic algorithms

Development

The original genetic algorithm as defined by Holland *et al.* is now known as the simple genetic algorithm (SGA).²⁴³ Alternative and newer GAs and evolutionary algorithms (EAs) use alternative operations within the algorithm, but the concept is broadly similar. The method was not specific to simulation optimisation, but was developed as a metaheuristic with broad applicability to a range of optimisation problems. Very quickly, GAs became widely researched and applied in a range of areas.^{244–246}

Practical applicability

Five studies reported the use of a genetic algorithm for a combinatorial simulation optimisation problem.^{224,226–229}

Ding et al. (2005) apply a genetic algorithm to a supply chain simulation model with a single objective.²²⁶ Binary encoding was used and elements of the array represented supplier utilisation, assignment weight and replenishment level. The algorithm used roulette wheel selection and two-point crossover. Fixed probabilities for mutation and crossover were applied and an elitist selection strategy incorporated. Two limitations of the approach were that the algorithm stopping rule was simply after a predetermined number of iterations (500), and penalty factors were used to account for constraints being broken. These constraints were not encoded out and therefore represent inefficiency within the algorithm. However, the simulation model only took 1 second to run, and therefore the algorithm terminated after a few minutes. Convergence was identified after 100 iterations in their particular problem.

Jun et al. (2010) report a modification to the SGA approach which incorporates an orthogonal quantized crossover operator.²²⁷ This method was originally designed for continuous optimisation, however it is discussed as applicable for combinatorial problems but it is not implemented in a simulation optimisation problem, and therefore the methodology proposed is theoretical at this stage.

Korytkowski et al. (2013) apply a genetic algorithm to a dispatching problem, with four independent objectives considered.²²⁸ Integer encoding was used and a combination of roulette wheel and tournament selection applied. Two point crossover and mutation were applied with predefined probabilities. Elitist selection was applied and the algorithm stopped when the best solution stabilised. The algorithm reached a stop condition after 10.5 hrs (~36,000 iterations). For its four criteria of interest, the algorithm converged and the authors conclude that a near optimal solution was found after an acceptable time.

Lacksonen et al. (2001) applied a genetic algorithm (along with three other optimisation algorithms) for four buffer-size problems.²²⁴ The study therefore provides a useful comparison of alternative methods across four different problems. However, because of this, full details about the genetic algorithm are not reported. The authors found that the size of the problem had a significant effect on the success of all methods (pattern search, simulated annealing, simplex method) apart from the genetic algorithm. The other three methods had poor results when the problem size increased, however the GA require significantly more replications to achieve the better result. The authors suggest that GAs have a clear trade-off between

accuracy and speed. Good solutions were generally found after 1000 replications; however this is problem-specific.

Yang et al. (2007) applied an evolutionary algorithm to a parallel-machine scheduling problem.²²⁹ Real encoding of the solution was used, and roulette wheel with elitism applied for selection. Two-point crossover and mutation was undertaken using a predefined probability. Stopping criteria was a fixed number of iterations. The algorithm was found to be robust for alternative tuning parameters (crossover and mutation rates) and starting populations. However, the authors raised concerns about the computational efficiency of the algorithm; the algorithm took 2 hours to execute.

None of the identified studies were concerned with a combinatorial sequencing problem, and therefore there was no permutation encoding of the solutions. Along with the encoding of the solution, the methods identified required tuning of the crossover and mutation parameters, the selection of the starting population, and the rules for terminating the algorithm. These user-defined aspects to the method introduce potential biases. However, there is an established body of literature outside of simulation-optimisation to ensure the GA is developed correctly and these parameters are unlikely to require special attention due to the evaluation of the objective function requiring a simulation.

None of the studies fully consider how constraints and infeasible solutions are accounted for. Ding et al. (2005) use penalty functions which are applied to the output of an infeasible solution.²²⁶ This is an obvious inefficiency and it is likely to be much quicker to encode out infeasible solutions rather than simulating a solution and penalising it.

Across most studies identified, convergence of the algorithm was identified after a few hundred iterations of the GA. It is not clear how relevant this may be to the treatment sequencing problem, but it provides an indication of the computational burden that GAs generally require. Lacksonen et al. (2001) and Yang et al. (2007) both report concerns about the efficiency of genetic algorithms.^{224,229}

The recognised strengths of GAs are that they are a good method for combinatorial problems, and the population approach balances exploration and exploitation, generally achieving convergence to near optima and avoiding early convergence. User-defined tuning of the algorithm is required, but there has been much research outside of simulation-optimisation to inform this.

The limitations identified include the concerns from two studies that the algorithm is slow. Also, no study used a permutation representation of the permutation, which is more complex

to apply mutation and crossover. In theory, the stochastic nature of a simulation model is likely to cause problems to the traditional stochastic selection process such as roulette-wheel. Although some authors used tournament selection, it was surprising that there was not more discussion regarding this within a simulation-optimisation context.

Simulated annealing

Development

Simulated annealing (SA) is a local search metaheuristic with the capacity to escape from local optima. The term 'simulated annealing' is an analogy to the process of annealing within crystalline solid. A solid is heated and then allowed to very slowly cool so that a crystalline structure of superior structural integrity remains.²⁴⁷ The stochastic acceptance mechanism of SA is a generalisation of the Metropolis algorithm, which is a method of sampling a Boltzmann distribution. The Metropolis algorithm is provided in Box 6.1.²⁴⁸

$$p(\delta E) = e^{(-\frac{\delta E}{kT})}$$

Where:

$p(\delta E)$ = probability of an increase in energy by δE

T = temperature

k = Boltzmann's constant (from the law of thermodynamics)

Box 6.1: Metropolis algorithm

The Metropolis algorithm requires a control parameter, called the 'temperature'. SA allows the temperature parameter to be slowly cooled as the algorithm iterates, and this cooling rate is a key determinant of the success of the SA algorithm. This method was applied to solving combinatorial optimisation problems by Kirkpatrick et al. (1983) and Černý et al. (1985).^{249,250} It has now become one of the most widely used metaheuristics in combinatorial optimisation. A key development by Belisle (1992) was the development of a generalised SA algorithm which contained a heuristic temperature cooling schedule.²⁵¹

The method has been mathematically proven to converge on a global optimum, even in multi-modal, discontinuous and noisy functions. A good discussion and summary of the proofs and conditions of convergence can be founded in Henderson et al. (2003).²⁴⁷

How does it work?

The original SA algorithm begins with an initial solution, typically determined at random. At each iteration of the algorithm a neighbouring solution is selected (either at random or some other low level heuristic). If the neighbour is better than or equal to the current solution, it is selected as the current best. If the neighbour is worse, then it is selected with a probability determined by the difference in objective values between the two competing solutions, and a

temperature parameter determined by the current number of algorithm iterations. The acceptance criterion is formally stated below in Box 6.2.

$$P\{\text{Accept } x' \text{ as next solution}\} = \begin{cases} \exp\left[-\frac{g(x') - g(x)}{t_k}\right], & \text{if } g(x') - g(x) < 0 \\ 1, & \text{if } g(x') - g(x) \geq 0 \end{cases}$$

Where:

x' is a solution in the neighbourhood of x

t_k is the temperature parameter at iteration k such that:

$t_k > 0$ for all k

Box 6.2: Simulated Annealing acceptance criterion

The probability of accepting a worse solution is based on Boltzmann’s law of thermodynamics, analogous to the annealing process that underpins the SA methodology. As with annealing, the SA algorithm begins with a high temperature (a higher probability of accepting a worsening move). This allows the algorithm to move out of local optima, especially in the early iterations. As the algorithm iterates, the temperature is gradually reduced based on a *cooling schedule*. This allows the algorithm to gradually focus on an area within the search space where a near optimum can be found.

A recognised advantage of the SA algorithm is the relatively small number of parameters which need to be modified (tuned). The initial solution and the neighbouring feasible solution can be randomly selected. The two key parameters within the algorithm are the initial temperature, and the temperature cooling schedule. The stopping criteria for the algorithm are often based on either a maximum number of iterations, or when an improvement is not seen for a number of iterations.

However, the tuning of the temperature value and cooling schedule is absolutely crucial to the success of a SA application. A large temperature value evaluates to a probability of one within the acceptance criteria for accepting an inferior move – resulting in a random search. However, a small temperature value results in a standard local hill climbing algorithm, because the probability of accepting an inferior move is zero. Therefore the tuning of this parameter is important to ensure a balance between exploration and exploitation.

Three main cooling schedules have been the focus of much research involving SA – the logarithmic, Cauchy and exponential cooling schedules. Proofs are available which show that SA will converge to a global optima when the logarithmic cooling schedule is used.^{252,253} A faster cooling schedule was developed by Cauchy, and proven to converge on a global optima.²⁵⁴ An even faster exponential schedule has been often used. Attempted proofs regarding its convergence are contentious, however there have been strong arguments for its convergence when variables are bounded, which supports its use for discrete and finite

optimisation problem.²⁵⁵ Although its theoretical capabilities are unproven, the speed of the algorithm is often required for efficient computation and therefore it has been applied for optimisation or large problems or those with an expensive simulation.

Practical applicability

Seven studies reported the use of simulated annealing for a combinatorial simulation optimisation problem.^{219–225}

Ahmed et al. (1997) applied a slightly modified SA method to three multi-echelon repairable item inventory systems. The neighbourhood function was defined as the set that could be reached via a single perturbation (one change to the solution). The approach deviated from the standard SA, in particular by using the White method to determine the initial temperature. This involves computing a number of transitions from a starting solution and estimating the standard deviation of the objective value. The initial temperature is then set to this value.²⁵⁶ A cooling rate of 0.9 was selected based on the theoretical research by Kirkpatrick et al. (1983).²⁵⁷ As per the White method, the stopping criterion for the algorithm was based on the final temperature value. There was no modification or consideration of the stochastic nature of the evaluation function and how that may influence the performance of the SA algorithm. Instead, the authors focus on the fact that their particular problem contains stochastic constraints within the search space. However, the results from the study suggest a good performance of their SA algorithm. They were able to generate analytical solutions to the Markovian system that their problem was related to. In all three test cases the SA algorithm found results very close to the analytically proven optima. The authors conclude that the algorithm was relatively efficient and viable for their simulation run time and problem size.

Interestingly, they highlight the problem of using the penalty approach to avoid an infeasible solution being the selected as near optima by the algorithm. During early configurations, the algorithm selected final solutions outside of the feasible region even when using penalty functions. Therefore they encoded the neighbourhood function to ensure that only feasibly transitions were accepted.

In another study by the same research group, Ahmed et al. (2002) present a modified SA algorithm which incorporates ranking and selection methods to solve discrete stochastic optimisation problems.²²⁰ This is called Simulated Annealing with a Ranking and Selection procedure (SARS). A mathematical proof for near-guaranteed convergence is provided and empirical estimates based on an inventory optimisation problem.

Ranking and selection (R&S) is a statistical method for selecting the best result from a comparison of stochastic simulation results with a specified level of confidence. Often R&S is used when a small set of comparisons are reported by a simulation model. However R&S is also able to consider, at each iteration of a simulation model, whether the extra unit of computation effort is worthwhile. Therefore R&S minimises the computation effort to reach this specified level of confidence. This allows a specification of the number of samples required to ensure a desired probability of selecting the best alternative, as well as determining when alternatives can be designated as inferior and the simulation terminated. For optimisation, this is a very useful method if the search algorithm has already identified some good solutions because it provides a method of discarding inferior solutions at an early stage of the simulation. Traditionally a frequentist procedure, Chick and Inoue (2001) developed R&S procedures based on Bayesian expected value of information methods,²⁵⁸ with good results shown in several studies.^{259,260}

Ahmed et al. (2002) prove that by combining R&S with SA, the sequence converges almost surely to the optimum as the simulation runs required to evaluate the solution objective goes to infinity.²²⁰ However, in practice a stopping criterion is required for the algorithm.

The algorithm was tested within an inventory problem, where an analytical solution was provided. Across eight different case test cases, the algorithm found solutions very close to the proven analytical solution. Little detail was given to the practical implications of these results and the run time required. The simulation used was relatively simplistic to enable comparison with an analytical solution.

Alrefaei et al. (1999) used a SA algorithm to solve a discrete stochastic optimisation problem. They modify the algorithm to use a constant temperature, rather than the standard SA approach using a decreasing temperature as the algorithm iterates. This allows the search to more freely move within the state. They find that under mild conditions the search converges. They also modify the approach to treat the most visited state as the current optimum, rather than the recorded 'current best' solution. The rationale for these changes is not clearly provided. They test the performance of their algorithm in two queuing problems and against two alternative SA algorithms. They conclude that their algorithm is superior, with better solutions identified, but the performance was dependent on tuning a number of parameters (temperature and neighbourhood structure in particular).

Ghiani et al. (2007) use a SA algorithm to solve a discrete stochastic optimisation problem. The focus of their study, rather than proving convergence, is the efficient applicability to numerous problems and the ability to parallelise (see later) the SA algorithm to ensure a consistently

good result is found in a reasonable amount of time. They also incorporate an indifference-zone R&S procedure, based on the method by Rinott.²⁶¹ This method requires an 'indifference zone width' δ parameter, which selects a solution with expected performance within δ units of the optimal performance with a confidence level of $1-\delta$. The procedure requires a reference configuration of samples to calculate the sample mean and marginal sample variance (first-stage). The second stage involves a number of additional samples of this reference configuration to calculate the overall sample means. The number of replications allocated to each configuration is proportional to the estimated sample variance in the first stage of the procedure, as well as the indifference zone width parameter.

As mentioned previously, R&S procedures are traditionally only relevant to small sized problems with full enumeration possible (although computationally expensive) and they require a decision-maker to be 'indifferent to very similar differences'. Within the Ghiani et al. (2007) study, the authors have incorporated the Rinott R&S procedure within the SA algorithm in an innovative approach to avoid the traditional R&S limitations.

The algorithm begins with estimating the Rinott procedure for the initial starting feasible solution (randomly chosen). The acceptance criteria are then determined by the estimated sample mean from the Rinott procedure for the new neighbouring solution. In practice, the process requires the interruption of the simulation model to estimate the first stage sample statistics, which informs the number of samples required for the second stage.

Parallelisation (parallel computing) is where an algorithm or computer process is written so that many calculations are carried out simultaneously (as opposed to sequentially).²⁶² If a large problem is divided into smaller ones, then these smaller problems can be allocated across numerous processors (from single computers, to clusters and clouds) and the total problem solved much more quickly. However, the concurrent nature of tasks running in parallel introduces complexity, and therefore parallel programming is often much more difficult than sequential programming.²⁶³ Also, the nature of the SA algorithm does not naturally offer up parallelisation solutions, because the standard Metropolis algorithm for acceptance *"...depends on one or more previous states plus one or more random variables. This serial nature of SA is an inherent distraction to parallelization."*²⁶⁴ However, some literature outside of simulation-optimisation has attempted to develop parallelised versions of SA, with acceptable speed-up gains found.²⁶⁵⁻²⁶⁷

Within this study, a master-slave single-thread parallelisation process was incorporated, which found significant speed-up gains (between ~ 2 (2 processors) and ~ 13 (16 processors) times faster than a sequential implementation). However, better solutions were not obtained. Multi-

thread parallelisation process found slightly superior results (less than 3% on average), but much longer computation times was required to achieve this relatively small improvement, compared to master-slave single-thread parallelisation. The balance between efficiency and effectiveness of search algorithms appears to be a major issue within the parallelisation literature, as well as the general metaheuristics literature.

Haddock et al (1992) used simulated annealing to solve a discrete manufacturing process problem. The problem size was small (120 combinations) and therefore it was possible to enumerate each solution. The authors found that a more expensive process of multi-restarting the algorithm was required to achieve a value which is optimal or near optimal. Otherwise, the algorithm was sensitive to different tuning parameter values (initial temperature, number of iterations at each temperature, final temperature value). The authors found that the SA algorithm identified the global optimum after evaluating approximately 30% of the total 120 input variable combinations.

The small problem size limit the conclusions of this study in terms of its relevance; however it does suggest that SA can be a reliable method to identify a near optimal solution, in particular if a multi-restart method is incorporated.

Lacksonen et al. (2001) applied a SA algorithm (along with three other optimisation algorithms) for four buffer-size problems.²²⁴ As mentioned in the genetic algorithm subsection, the study therefore provides a useful comparison of alternative methods across four different problems. The SA algorithm used was simplistic, and may account for its lack of performance compared to the genetic algorithm. The cooling schedule had little justification and it was not clear what tuning was applied to each parameter. For each problem size, the SA algorithm was found to be inferior to the GA, however the GA required almost double the amount of replications to solve. The authors explore the possibility that the problems with the SA algorithm arose from requiring different cooling schedules and starting points for each solution. It was not clear why this was not conducted to provide a fair comparison between each search method.

Rosen et al. (2005) applied a SA algorithm to solve a discrete manufacturing process problem. The algorithm is a modification of the standard Kirkpatrick et al. SA algorithm.²⁵⁷ The authors implement both algorithms to provide a comparison, and find that their modified SA algorithm identified either equivalent or superior results in all instances, and usually required a similar number of simulation runs. Significant was the finding that the new algorithm was easier to tune and in general required fewer simulation runs.

The modified algorithm incorporated a linear approximation method to identify an area where a solution of high quality is likely to be. These linear approximation methods are adaptations of the response surface methods by Box and Wilson.²⁶⁸ The assumption made is that other high quality solutions could exist in an adjacent neighbourhood, which is explored using SA. Repeatedly iterating this two-phase process using different starting solutions is undertaken to find a final best solution.

Tabu search

Development

Tabu search was developed by Glover (1989, 1990) to solve combinatorial optimisation problems.^{269,270} The method is an extension to standard local search hill climbing methods, and introduces memory to prevent reversing recently accepted moves. Therefore whenever a local optimum is encountered, non-improving moves are accepted because previously visited solutions recorded in the memory (the tabu-list) are forbidden.

Since its development in the late 1980s, the method has a reasonably stable history, with only minor modifications to the process of the algorithm and the rules regarding memory structure. In the first fifteen years from the method's development, TS has been applied to combinatorial optimisation problems in well over 100 published papers.²⁷¹

The development of TS was partially motivated by human behaviour. Glover discusses that inconsistent behaviour has often be observed even in similar circumstances and the tendency to deviate from a course might both be an error but also an unexpected gain.²⁶⁹ The TS does not deviate randomly, but instead supposes that there is no point accepting a poor solution unless it is to avoid an already investigated area.

Tabu search is an example of a computational algorithm with adaptive memory. Adaptive memory programming is a growing area of research within metaheuristics and an avenue of rich potential.²⁷²

Practical applicability

Two studies have reported the use of TS for a combinatorial simulation optimisation problem.^{230,231}

Azadeh et al. (2010) develop an integrated process of response surface methodology (RSM) and TS to optimise a TS to a discrete production system problem.²³⁰ A design of experiment analysis is undertaken using the RSM to provide a modelled estimate of the search space and objective function. The TS is then applied to these modelled estimates, rather than requiring

simulation runs to evaluate the objective function for each solution. There are a number of limitations to the study which should be noted. Firstly, the problem size is relatively small, and the RSM is fitted for 81 (3^4) observations, which allows enumeration rather than requiring a TS. The TS only required 26 replications to converge, and the neighbourhood structure for three competing design factors results in six possible neighbours for each solution. It is not fully clear how long the simulation model of the production process requires to achieve a steady state estimate of the objective function. A useful comparison would have been to see what efficiency saving was made by incorporating the RSM model and whether the results with and without (just a simulation-optimisation via TS) identified would differ. The RSM invariably adds a level of uncertainty and it is not fully detailed how the model was specified and validated. The authors conclude that the methods they follow can be applied for all types of discrete production system optimisation which is probably over optimistic given the problem size they consider. They also conclude that a global optimum is identified, but no proof is provided.

Yang et al. (2004) attempt to solve a flow shop with multiple processors (FSMP) problem using a TS simulation optimisation method. The FSMP problem involves sequencing jobs in a flow shop for processing by more than one identical processing machine. This is a commonly observed NP-hard problem in manufacturing and operational research. A discrete event simulation model is used to evaluate the performance of each solution, with tardiness the objective to be minimised. The TS is separated into two methods, TS1 and TS2. TS1 presents the basic implementation of TS as proposed by Glover. TS2 incorporates a long-term memory structure to inform the restart procedure of the algorithm. Moves from the current solution are undertaken using a pairwise-exchange/swap method. This is a commonly used method to construct a neighbouring solution in a permutation-represented problem.²⁷³ An aspiration criterion is applied so that tabu moves are (re)accepted if they are good solutions. Tabu list size is based on previous research which suggests that it is an integer between $n/3$ and $3n/2$ where n is the problem size.²⁷⁴ The search stops after a defined number of iterations, based on the problem size and number of possible swaps, $n(n-1)$.

The search is run for five different problem scenarios. The authors apply a steepest descent pairwise interchange (SDPI) heuristic to solve the problem and provide a benchmark solution. Both TS1 and TS2 were a superior search, taking a comparable computational time compared to SDPI and finding superior solutions in all five scenarios. In three scenarios, TS1 found a better solution compared to TS2. The study represents a good implementation of TS for a permutation problem. The high speed of the simulation model results in a final solution being found by the TS algorithm within 30 minutes, which represents a good method for solving a

practical optimisation problem. The authors do not identify any limitations with the TS method that they propose.

Nested partitions

Development

Nested Partitions (NP) is a randomised optimisation framework.^{233,275,276} The method was developed to be applicable to both deterministic and stochastic discrete optimisation problems. Immediately after development, the method was applied to two combinatorial problems, Shi et al. (1999) for a Travelling Salesman Problem, and Olafsson et al. (2000) for a Parallel-Machine Flexible-Resource Scheduling problem.^{276,277}

Practical applicability

One study has reported the use of Nested Partitions (NP) for a combinatorial simulation optimisation problem.²³³ Shi et al. (2000) use the stochastic nested partitions method to optimise a stochastic travelling salesman problem (TSP).²³³ The problem is stochastic because the time between each city is uncertain, and the total time for each route estimated as the average time across each replication of the simulation model. Test problems of size 51, 76 and 101 nodes were used. The percentage over the optimum for the three problems was 2.77%, 3.12% and 5.38% respectively, which required 300 iterations of the NP algorithm. This performance worsened by approximately 10% when variable noise was added to the performance measure of the problem. The authors suggest this is a reasonable amount and conclude that the algorithm is relatively robust to noise in the simulation. Convergence to a near optimum after 300 replications across three relatively large problem sizes is a positive result for the relatively new NP method.

Ordinal optimisation

Development

The development of ordinal optimisation (OO) for simulation optimisation arose because making a decision to move within the solution space only requires an ordinal comparison between two solutions ($g(x_1) < g(x_2)$). A precise estimate of the difference between two solutions is not required. As stated by Deng and Ho (1999), "*It is much easier to determine 'order' than 'value'.*"²⁷⁸ These comparisons are the focus of OO, which was developed by Ho et al. (1992).²⁷⁹ In particular, when uncertainty is present then the benefit of an ordinal approach over a cardinal approach is even more significant. There is an error possible in simulation optimisation because a solution chosen as superior could be inferior due to the uncertain

estimate of their performance. This error decreases in probability as the difference between two comparators increases.

It has been shown that ordinal comparison methods converge much more quickly compared to cardinal estimation, potentially at an exponential rate with respect to the number of algorithm iterations.²⁸⁰ The development of these methods was independent of any particular metaheuristic. This is because in all metaheuristics, moves are all based on comparisons between solutions and therefore the ordinal comparison approach is relevant to all metaheuristics.

Practical applicability

One study has reported the use of OO for a combinatorial simulation optimisation problem.²³² Ho et al. (2000) use OO for a buffer allocation problem. A stochastic resource allocation algorithm is used with ordinal comparison between alternatives applied. The problem was fixed so that the optimal solution was analytically possible to find. The algorithm required only approximately 20 iterations to converge on the analytical solution. Even when using the worse possible allocations as the starting selection, the algorithm very quickly converged.

However, there are some limitations to note. Firstly, the problem is simplistic and the simulation (6 users for 24 buffer slots) would be extremely fast to run. Secondly, very little detail is provided regarding the simulation model and the methods used to apply the OO algorithm.

Particle swarm optimisation

Development

Particle Swarm Optimisation (PSO) was developed by Eberhart & Kennedy (1995).²⁸¹ It is a population based metaheuristic with candidate solutions (particles) moving through the search space via formulae determining their position and velocity. These formulae incorporate information regarding the particle's own best solution, but also the best solutions of other particles in the 'swarm' which guide each particle towards good positions within the search space.

PSO was developed as a computer algorithm to simulate the movement of organisms, such as birds and fish. The algorithm's original objective was to simulate the choreography of swarms of organisms, which results in synchronous movement. With particles emulating the success of neighbouring particles, the algorithm was found to discover optimal regions within high dimensional search spaces. The method was refined and simplified when the objective altered

and became focussed specifically on the applicability of the particle swarm algorithm for optimisation.²⁸²

Practical applicability

One study has reported the use of PSO for a combinatorial simulation optimisation problem.²³⁴ Kuo et al. (2011) report using a PSO algorithm for an assembly line problem. They modify the standard PSO approach by incorporating a mutation operator which alters similar particles (Particle Swarm Optimisation algorithm with Mutation based on Similarity – PSOMS). The method also incorporates an inertia weight. They also incorporate two genetic algorithms to compare against the PSOMS method. Only very limited information about these genetic algorithms is provided which is why they are not formally included in the review. A standard PSO algorithm was implemented as well.

The results showed that PSOMS converged fastest and to the best solution when compared to PSO and two GAs. However, the success of the algorithm was dependent on the inertia weight parameters. Five evaluations were undertaken and the results varied dramatically by the inertia weight used. The PSOMS was also very sensitive to the number of particles used. In general, the algorithm improved with higher numbers of particles. The results were also sensitive to the learning factor parameters and the mutation parameters. On the whole, the PSOMS appears to require the tuning of several parameters and therefore the appropriateness of the method must be considered.

HYBRID AND OTHER METHODS

Not all of the identified methods for a combinatorial simulation optimisation problem can be easily classified. Many of the methods that have been reviewed already are themselves modifications of existing methods. Often, particular elements from a method are revised or used within a different method, to provide a more appropriate solution to a given problem.

However, the two methods reviewed in this section, averaging for simulated annealing, and empirical stochastic branch-and-bound, are truly hybrid methods which look to combine heuristic algorithms with statistical methods so that they resolve the unique issues which arise when attempting to optimise a simulation model.

Also included in this section is a discussion regarding hyperheuristic methods. Although hyperheuristics methods for simulation optimisation were not found within this review, discussion with experts at a National Taught Course Centre in Operational Research (NATCOR) metaheuristics course highlighted that hyperheuristics is a field of research which may have

potential applicability to our sequencing problem. Therefore a discussion and review of the key literature is provided.

Hybrid methods

Development

Prudius et al. (2012) developed an averaging framework for simulated annealing.²³⁵ The method is a hybrid application of random search methods as well as averaging methods. The method is recently developed, however it draws upon the whole spectrum of simulation optimisation methods as reviewed within this chapter.

Xu et al. (2013) developed an empirical stochastic branch-and-bound method for simulation optimisation.²³⁶ This is a hybrid method which combines nested partitions and stochastic branch-and-bound.

Practical applicability

Prudius et al. (2012) apply their average framework for simulated annealing method to a three-stage buffer allocation problem.²³⁵ They do not provide a great deal of information about the practical application of their method. The global algorithms incorporating averaging perform better than local algorithms, however this is not unexpected. They do identify that averaging does not always benefit the algorithm it is applied to. They apply two variants of simulated annealing and demonstrated that averaging alone may either help or hurt the performance of the algorithm, compared with no averaging. They do find that averaging within the algorithm as well as an adaptiveness component to avoid unnecessary simulation does appear to be an effective strategy.

Xu et al. (2013) apply their empirical stochastic branch and bound method to a three-stage flow line with finite buffer storage problem. They find that the algorithm converges asymptotically to the global optimum, and they empirically show that their method outperforms the standard nested partitioning approach. The advantages of the method are maximised when the problem is noisy or there are significant interactions between the decision variables. Little other information regarding performance was given, due to the study being focussed on the methodological development and proof of convergence.

Hyperheuristics

The current heuristic approaches for searching and optimisation still tend to focus on bespoke systems and solutions, even given their flexibility. In general, these solutions and algorithms have been found to be expensive to develop and run. However, despite this they have provided successful results for real world problems.

Unfortunately, the application of a metaheuristic to a new problem domain (or even a new problem instance) still tends to require expert involvement and often an expensive period of research. While a metaheuristic may generate good performance in an alternative problem instance, it is often observed that an alternative may also generate good or superior performance.^{283,284}

The goal of much current research is to raise the generalisability of search methodologies through learning and adaptation so that these methods are applicable across a broad range of problems without requiring any human expert intervention. This area is challenging, because there is a lack of theoretical understanding of how to build a meta-level intelligent search method which is capable of automatically tuning, selecting and generating a search system.²⁸⁵

A hyperheuristic is a search method or learning mechanism for selecting or generating heuristics to solve computationally difficult problems. The hyperheuristic operates on the high-level search space of heuristics, rather than on the low level search space of solutions. Or more simply, a hyperheuristic is optimising the selection of a lower level heuristic, which in turn derives a solution for the problem.

Although a cutting edge area of current research, the origin of hyperheuristics can be traced back to the early 1960s.^{286,287} Recent research has involved the benchmarking of alternative hyperheuristics (HyFlex research project), with the AdapHH heuristic showing clear superiority.^{288,289} However, it in itself requires the tuning of approximately 20 parameters. It is still not clear if it really offers a generalisable framework for applying heuristics to solve a computationally difficult problem.

The cutting edge nature of hyperheuristics contrasts the more established background of metaheuristics, many of which have a proven theoretical basis. For our particular sequencing problem, it does not currently seem appropriate to take advantage of hyperheuristics and automate the process of heuristic selection. Instead it seems more appropriate to apply one or more established metaheuristic methods to attempt to find a good solution and method, and then in the future to maybe consider the applicability of a hyperheuristic.

6.9 DISCUSSION AND CONCLUSIONS

The aim of this chapter was to systematically search published literature to identify relevant methods for finding an optimal or near-optimal sequence of treatments in an economic evaluation using discrete event simulation.

The treatment sequencing problem was formulated as a combinatorial discrete simulation optimisation problem. Therefore methods were searched for which were either developed for, or addressed, problems of this type. To facilitate this review, a citation pearl growing search was undertaken to identify all relevant methods across a disparate area of research. From the results of this search, a narrative review and synthesis was undertaken to draw conclusions regarding the applicability of each main method type. Methods were assessed using a bespoke framework which addressed their theoretical basis, practical applicability, and relevance to the treatment sequencing problem.

The methods identified could be grouped into two broad categories: random search methods; and metaheuristic methods. The majority of methods identified were metaheuristics, which are generalised methods with the objective of being applicable to a range of problem types. This is an attractive property, because they provide a theoretical basis for their applicability, and the review found that they have been applied to a wide range of problems. This supports their use for a combinatorial discrete simulation optimisation problem, although no instances were found where a metaheuristic was found for a problem with a permutation encoding. This is most likely to be the superior way to represent the treatment sequencing problem. Little evidence was found about the relative performance of each metaheuristic. Lacksonen et al. (2001) applied SA and GA algorithms for four buffer-size problems.²²⁴ The SA algorithm used was simplistic, and may account for its perceived lack of performance compared to the GA. Also, the detail provided in the article was limited, and therefore robust conclusions cannot be drawn.

Statistical methods such as metamodeling, adaptive sampling, and approximation, were excluded from the review. These methods are concerned with how simulation runs can be efficiently 'selected', and how superiority can be proven in the presence of noise. These methods have been shown to work well in combination with a range of search algorithms, and may provide a significant 'speed up' to the ability of a search algorithm to identify a near optimal solution. However, some of the methods were only relevant for relatively small problems. Additionally, some required a continual evaluation of the output of the simulation model while it was simulated.

Some important trade-offs have been identified, and some methods explicitly provide the ability to 'balance' these. In particular, there is a trade-off between exploitation and exploration – ensuring the global space is searched but also looking for the local optima within a particular area. Furthermore, there is a trade-off between the ability to search within the

space, and ensuring that precision is obtained in each simulation of a solutions objective function.

Some limitations of the systematic review should be noted. Firstly, the use of a bespoke framework is required because there is no gold standard method against which to compare. This is open to reporting bias, in particular due to one researcher working on this review. However, the use of a framework to assess the relevance and merits of alternative methodologies has been recommended in the methodological systematic review literature.²⁰²

Also, the review of methods fell well outside of health economics and the regular area of research and expertise – into engineering, computing and mathematics. This may provide the potential for errors in interpreting and understanding particular aspects of a method or problem. Intensive training was undertaken alongside this PhD research, in particular on metaheuristic methods and simulation modelling, to ensure the skills were available to assess and apply alternative methods in this area. The objective of the reviews was not to inform the theoretical design of a new optimisation method, but to understand existing methods to an appropriate level which would allow a judgement regarding their appropriateness to the treatment sequencing problem, and subsequently the ability to implement the method using the RA DES model.

Finally, the search was specific to combinatorial and simulation optimisation problems, however methods that are not specific to these may still be relevant and applicable. It should be noted that very few sequencing or permutation problems were identified, which are of particular applicability to the treatment sequencing problem. These problem types have particular issues for search algorithms, in particular neighbourhood functions and move operators. Further reviewing for approaches to resolve these may be required when selecting, implementing and evaluating a method in the next stage of this thesis.

(Blank page)

CHAPTER 7: APPLICATION OF SIMULATION OPTIMISATION VIA SIMULATED ANNEALING FOR AN ECONOMIC EVALUATION OF TREATMENT SEQUENCES FOR RHEUMATOID ARTHRITIS

7.1 CHAPTER OVERVIEW

This chapter contains the pivotal analysis within this thesis. It draws from Chapter 5, where a discrete event simulation (DES) model is developed for estimating the cost-effectiveness of RA DMARDs. It also draws from Chapter 6, where a systematic review to identify methods which are relevant for the treatment sequencing problem is conducted.

In this chapter, simulated annealing (SA) with a memory function is the selected methodology to enable the simulation optimisation (SO) of the DES model developed in Chapter 5. The whole method is referred to as SOSA (Simulation Optimisation via Simulated Annealing).

The reasons for selecting SA are fully explained, and the methodology is reported for transparency. Modifications to the DES model are required to enable it to be operated by the simulation optimisation algorithm. Justifications for these modifications are provided and the model is validated.

The SA algorithm is comprehensively 'tuned', where key control parameters are altered to obtain good performance of the algorithm. The algorithm is then used to identify a good solution to the problem, and sensitivity analysis conducted to test the robustness of the results. The identified results are considered in terms of their potential implications for health resource allocation policy.

The overall performance of the SA algorithm is considered in the final discussion section of the chapter.

7.2 INTRODUCTION

The aim of this chapter is to undertake a simulation optimisation of a DES model, to enable the identification of an optimal or near-optimal DMARD treatment sequence for patients with RA.

Section 7.3 reports the methods of the analysis. This includes a justification for the chosen SO method, a SA algorithm with a memory structure. Modifications are required to the existing

DES model to enable it to be operated by the SA algorithm, and these are fully reported. The SOSA method and the process undertaken to tune its performance are detailed.

Section 7.4 presents the results of the SOSA analysis, along with sensitivity analyses. Section 7.5 provides a discussion on the analysis, and draw conclusions.

7.3 METHODS

SIMULATION OPTIMISATION

As reviewed and discussed in Chapter 6, simulation-optimisation (SO) methods are concerned with finding an optimal or near-optimal solution for a problem when potential solutions are evaluated using a simulation model. SO has two unique aspects, compared with standard optimisation methods. Firstly, the simulation model may be stochastic, with an uncertain estimate of the objective function. Secondly, the simulation model may take a significant period of time to evaluate a solution. SO methods must therefore overcome these challenges.

Rational for Simulation-Optimisation using Simulated Annealing (SOSA)

On the basis of the review and appraisal reported in Chapter 6, SA has been selected as the metaheuristic SO method to address the RA treatment sequencing problem outlined in earlier chapters. The rationale for this decision is based on the following four factors.

1. SA was found to be the most commonly used method in the systematic review reported in Chapter 6. Seven of the 28 (25%) of studies identified were studies where SA had been applied to a combinatorial SO problem. These studies on the whole reported that SA performed well for their particular SO problem.
2. Only one study has been identified which compares the performance of SA against other metaheuristics.²²⁴ Lacksonen et al. (2001) found in their case that SA might not perform as well as a genetic algorithm (GA), however neither method was extensively 'tuned' (the process of optimising the control parameters of the algorithms). On one occasion the SA method found a solution in much fewer replications compared to GA, which indicates that no definitive conclusions were found. There are several limitations of the Lacksonen et al. (2001) study which are detailed in Chapter 6.
3. SA is relatively easy to implement and programme, compared with more complex metaheuristics like Tabu Search, Particle Swarm and Genetic Algorithms.

4. Genetic Algorithm, another potential method, has in general been found to be robust, but also very slow to implement and run. Therefore a trade-off between speed and precision is required.

These four factors have led to the decision to select SA as the method to implement. Unfortunately the time taken to implement, validate and run multiple SO methods means that it is not possible to implement more than one SO method and judge their relative performance within this thesis.

Simulation Optimisation algorithm

A Simulation Optimisation (SO) algorithm has been developed to allow the simulation optimisation of the RA treatment sequencing problem. The algorithm is detailed in Figure 7.1. The SO algorithm allows an optimisation algorithm (in this case, a metaheuristic) to evaluate a particular potential solution by running a simulation model. The optimisation algorithm then considers the performance of that particular potential solution (the output of the simulation model) before generating a new potential solution for the simulation model to evaluate, or stopping the SO algorithm.

The algorithm begins by initialising both the simulation model and the optimisation algorithm. User-defined parameters govern both parts of the SO algorithm and need to be selected before starting the procedure. Also, a patient-level dataset is required for the simulation model. This contains patient characteristics as well as probabilistic parameter values. This dataset is generated at the start of the SO algorithm and passed to the patient simulation model. Alternatively, to ensure consistent results, a previously generated dataset can be loaded into the simulation model.

A starting treatment sequence is then generated (either at random, or a user-defined sequence) and this is evaluated by the simulation model. The results from this evaluation (the net monetary benefit (NMB) given a specified cost-effectiveness threshold) are then passed to the optimisation algorithm (in this case, a simulated annealing metaheuristic algorithm) which generates a neighbouring solution and loads this into the simulation model. The process of evaluation, results processing, next sequence generation and re-evaluation is fully automated once the algorithm is begun. The algorithm stops when a stopping criterion is met. The following sections within this chapter report on the methods utilised to contribute to this SO algorithm.

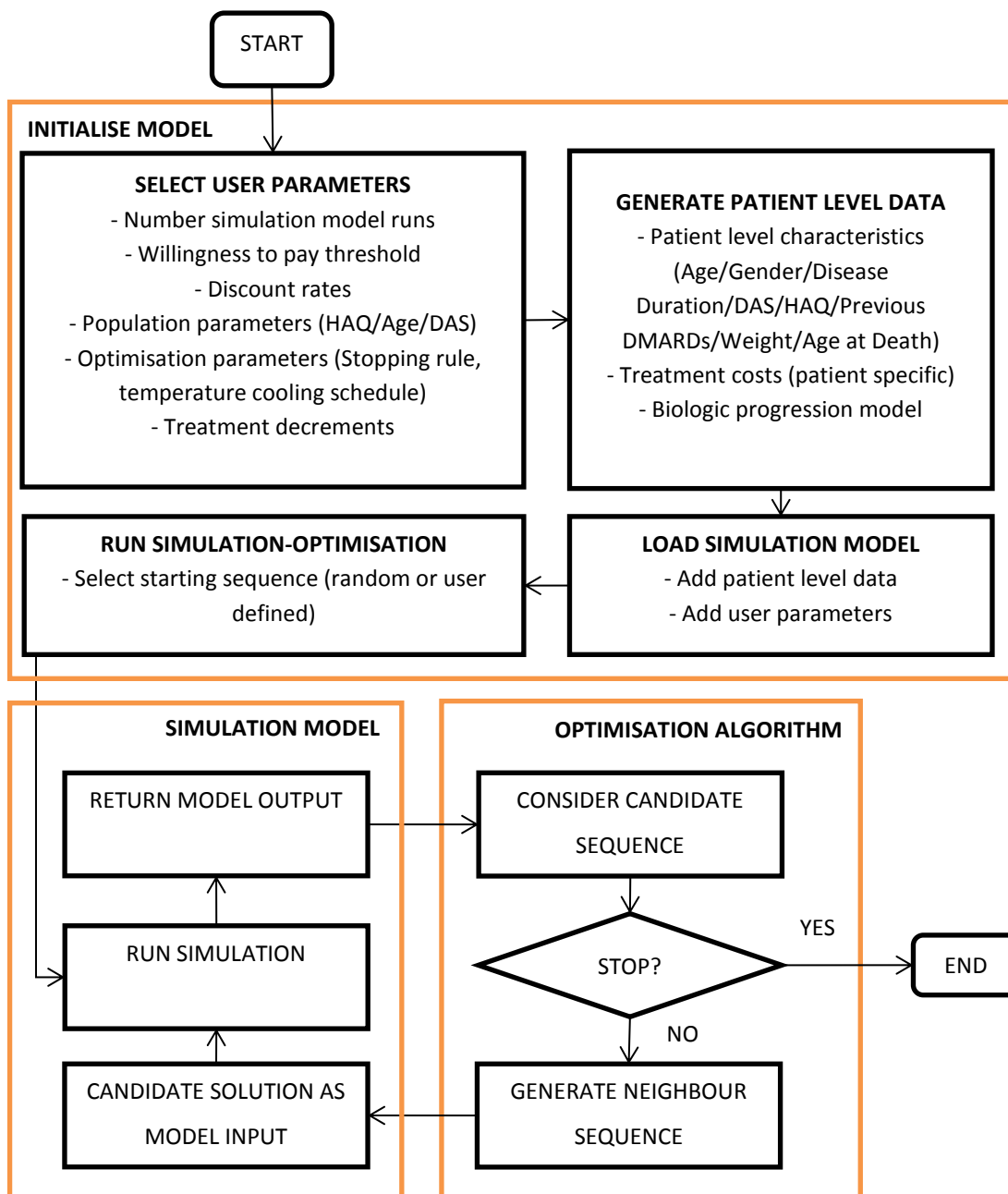


Figure 7.1: Simulation Optimisation algorithm

SIMULATION MODEL

Simul8

Simul8[®] is a simulation modelling software package widely used to develop discrete event simulation models. In recent years it has been increasingly used to develop patient level simulation models for cost-effectiveness analyses.²⁹⁰⁻²⁹³ The advantages of Simul8 for developing a patient level simulation model include: being faster than similar models implemented in MS Excel; having an intuitive graphical user interface (GUI) which allow easy

debugging and validation; and possessing a relatively powerful Visual Logic (VL) language which allows user defined functions and code to be run when events occur in the model.

Simul8 was chosen as the software package to implement the RA model because, from experience, it is much faster than MS Excel when running similar patient level simulation models. However, Simul8 is not routinely supported by NICE for submissions to its NICE Technology Appraisal process (because it is a relatively expensive bespoke simulation software package) and therefore the original RA model as reported in Chapter 5 was developed in MS Excel. It was decided the investment in time to rebuild the model in Simul8 was worthwhile to gain a much faster patient level model.

Modifications from NICE model

Where possible, the Simul8 version of the NICE model is an exact copy of the original NICE model developed in MS Excel. All analysis is in a severe RA population (Population 1 and 2). Identical parameters, evidence sources and costs are used (including confidential 'Patient Access Schemes' which reduce the cost of selected bDMARD treatments). The following sections report the particular areas of the current Simul8 model which deviate from the original NICE RA model.

Treatment decrement parameter

A treatment decrement parameter has been included within the Simul8 model. This parameter was not included within the NICE RA model, because of time constraints within the appraisal. Treatment decrement parameters have been used in previous NICE RA appraisals to consider the impact on the efficacy of a treatment of having several previous failures on treatments of the same general class. It is not seen as clinically plausible that the efficacy of a bDMARD after no previous bDMARDs is going to be the same as after 3 or 4 previous bDMARDs. Therefore for both bDMARDs and cDMARDs, a treatment decrement parameter was incorporated, broadly in line with the approach requested by NICE for the TA198 (now TA247) appraisal of TCZ for RA.* In this appraisal, the manufacturer undertook trial subgroup analyses specific to TCZ, ETN and RTX to adjust the efficacy data used in the economic model (efficacy reported using ACR 20/50/70 criteria). Their subgroup analysis produced inconsistent results in places, with improved efficacy seen in some levels of response, however it was accepted by the NICE appraisal committee as being valid in the absence of better data. These data are reported in Table 7.1.

* <http://www.nice.org.uk/guidance/ta247> - Accessed June 2015

Table 7.1: TA198 Treatment decrement estimates

Treatment		ACR 20/50/70 (%)
Tocilizumab (TCZ)	First line efficacy	62/31/12
	'Degraded' efficacy	50/31/15
Etanercept (ETN)	First line efficacy	62/38/16
	'Degraded' efficacy	49/26/7
Rituximab (RTX)	First line efficacy	46/23/14
	'Degraded' efficacy	42/22/10

It has been assumed that after two prior bDMARD or cDMARD treatments in the sequence, the decrement is applied. In the basecase, this treatment decrement is set at 10%, and will be varied in scenario analyses.

Meta model and curve-fitting

The NICE RA patient-level simulation uses the estimated coefficients from three econometric models to provide patient-specific model parameters.

Specifically, these models are:

- A generalised gamma regression model used to determine the length of treatment for a patient with a moderate EULAR treatment response
- A log-normal regression model used to determine the length of treatment for a patient with a good EULAR treatment response
- An Adjusted Limited Dependent Variable Mixture Model (ALDVMM) use to estimate health related utility values (HRUVs).

Table 7.2 provides the independent (patient specific) variables used throughout the patient simulation, which when combined with the estimated coefficients from the models, determine time on treatment (given a moderate or good EULAR response) and HRUV.

In MS Excel, these models can be easily incorporated to obtain a patient level prediction within the simulation model, by using Excel's inbuilt statistical functions (specifically, probability and cumulative distribution functions for normal, log-normal and gamma distributions). However, within Simul8 these statistical functions are not included.

From experience, Simul8 provides a significant improvement in simulation speed compared to MS Excel, and therefore undertaking a simulation-optimisation process is not likely to be feasible just in MS Excel. However, these econometric models represent the best use of available evidence to provide patient-level parameters in the simulation model. Therefore,

metamodels of these econometric models were constructed to enable their inclusion in the Simul8 model.

Table 7.2: NICE RA econometric models for patient level parameters

Model	Log normal	Generalised Gamma	ALDVMM
Dependent variable	Time on treatment (Good EULAR response)	Time on treatment (Moderate EULAR response)	Health Related Utility Value
Independent variables		Age Age ² Gender Disease Duration Disease Duration ² DAS Number of previous DMARDs HAQ	HAQ HAQ ² Age Age ²
Prediction parameters	Constant term Independent variable coefficients Sigma	Constant term Independent variable coefficients Sigma Kappa	-
Method to account for uncertainty	Multivariate normal distributions across all parameters		

Health Related Utility Value (HRUV) meta model

For the HRUV econometric model, a pragmatic solution was to fit a meta-model (essentially a regression model of a model) to allow the prediction of the dependent variables given specific patient characteristics, which can then be easily incorporated into the Simul8 simulation model.

To fit the metamodel, a dataset of the dependent and corresponding independent variables from the econometric model was required. Therefore, a dataset was generated using the simulation model in Excel, which returned both the dependent and independent variables. A dataset of 2000 simulation observations was generated for the utility value and corresponding independent variables. For the utility model, the independent variables are simply Age, Age², HAQ, HAQ². Therefore these independent variables were fitted to the meta model.

An OLS regression model was therefore fitted with the following functional form:

$$Utility = \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_3 HAQ + \beta_4 HAQ^2 + \varepsilon \quad [7.1]$$

The results of the regression are provided in Box 7.1.

Source	SS	df	MS	Number of obs = 2000		
Model	41.0386948	4	10.2596737	F(4, 1995)	=	4262.72
Residual	4.80164569	1995	.00240684	Prob > F	=	0.0000
-----				R-squared	=	0.8953
-----				Adj R-squared	=	0.8950
Total	45.8403405	1999	.022931636	Root MSE	=	.04906

util	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	-.000971	.0006111	-1.59	0.112	-.0021695	.0002275
age2	.0000105	5.11e-06	2.05	0.040	4.67e-07	.0000205
haq	-.2041815	.0074534	-27.39	0.000	-.2187988	-.1895643
haq2	-.0129101	.0024876	-5.19	0.000	-.0177887	-.0080315
_cons	.9100248	.0183569	49.57	0.000	.8740241	.9460255

Box 7.1: HRUV meta model regression results

The residuals showed some slight correlation (Figure 7.2), and alternative functional forms and other independent variables (such as a cubic term) may have improved the fit of the model, however within the timeframe of the thesis this was not possible. The predicted values appeared to fit the data (Figure 7.3).

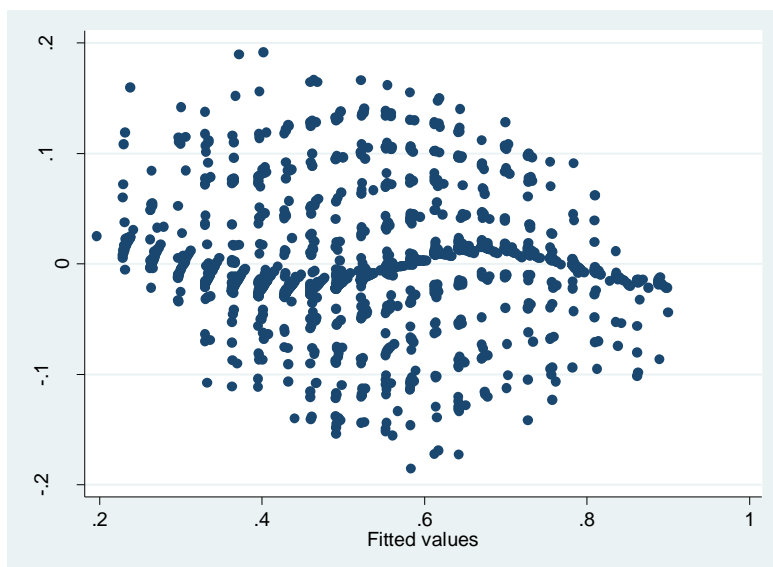


Figure 7.2: HRUV meta model regression residuals

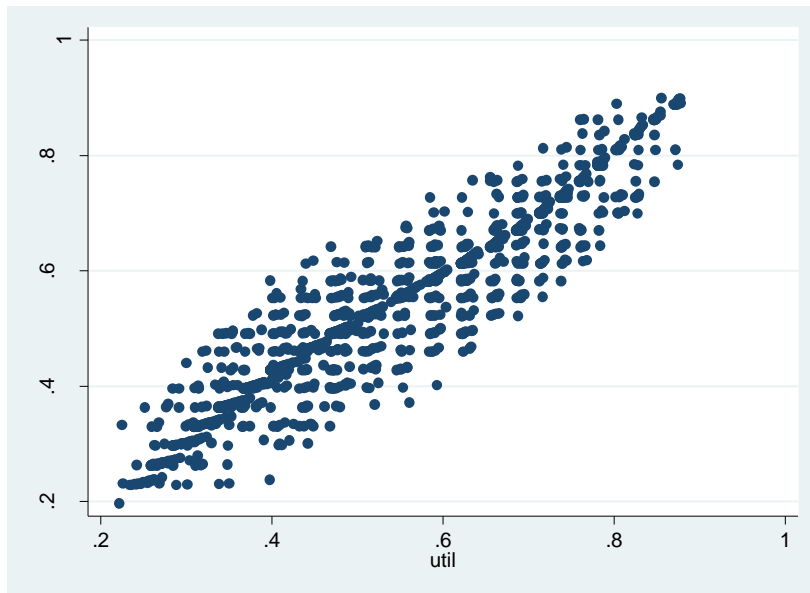


Figure 7.3: HRUV meta model regression fit

Time-on-treatment meta model

For the two time-on-treatment models, the output from the econometric model was a survivor function and corresponding survival curve. The survival function could not be implemented in Simul8, and therefore the two curves (good and moderate EULAR response) for a patient with mean characteristics from the analysis were digitised to allow a parametric distribution to be fitted to the digitised data. This is a commonly conducted method to allow a parametric model to be fitted to a published aggregate survival curve. In this case it represented the most effective way of incorporating the survival model data into the Simul8 model, however it is at the expense of reduced accuracy, in particular because the time-on-treatment parameters are now essentially global parameters, rather than being patient specific.

Parametric distributions were fitted to the data and the distribution parameters optimised by minimising the mean squared error using the MS Excel Solver add-in. Log-normal distributions were selected for both models (the EULAR Good Response time-on-treatment model is a log-normal, however the EULAR Moderate Response model is a generalised gamma) and were found to fit very closely to the original data (Figure 7.4, Figure 7.5).

EULAR Moderate

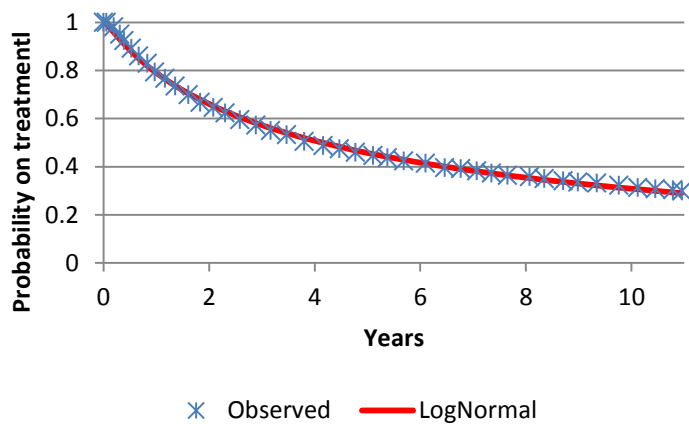


Figure 7.4: Time-on-treatment (Moderate EULAR response)

EULAR Good

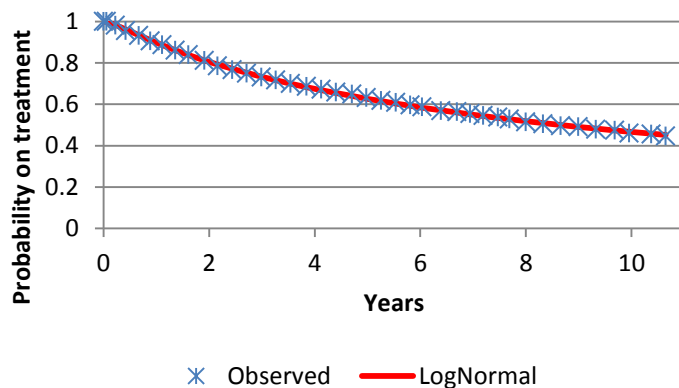


Figure 7.5: Time-on-treatment (Good EULAR response)

Treatments and decision space size

As highlighted in the systematic review of economic evaluations in Chapter 4, there is a large set of DMARDs for patients with rheumatoid arthritis.

For this analysis, all potential competing DMARDs currently used regularly in the NHS were selected for the pool of treatments from which to derive each treatment sequence. There are several cDMARD therapies (leflunomide, aziathaprine, ciclosporin, penicillamine, injectable gold) which are less commonly used and have therefore been omitted. This decision was based on two factors. Firstly, to ensure that resulting sequences are clinically meaningful and appropriate. Secondly, there is less comparative evidence to provide estimates of the relative efficacy of these older and less-commonly used treatments compared to more commonly used cDMARD treatments and the new biologic DMARDs (bDMARDs).

The set of treatments within the analysis is provided in Table 7.3. The treatments are grouped by their general classification, and include the current licenced indication for each treatment. Treatments can either be used first line, after a previous cDMARD, or after a previous bDMARD (rituximab).

Table 7.3: Treatments included in analysis

Classification	Treatment	Abbreviation	First possible use*		
			Pre cDMARD	Post cDMARD	Post bDMARD
Biological DMARDs (bDMARDs) [†]	Abatacept	ABT		X	
	Abatacept (subcut.)	ABTS		X	
	Adalimumab	ADA	X		
	Certolizumab Pegol	CTZ		X	
	Etanercept	ETN	X		
	Golimumab	GOL	X		
	Infliximab	IFX		X	
	Rituximab	RTX			X
	Tocilizumab	TCZ		X	
Conventional DMARDs (bDMARDs)	Hydroxychloroquine	HCQ	X		
	Methotrexate	MTX	X		
	Sulfasalazine	SSZ	X		
	TICORA [‡]	TICORA	X		
Sequence end treatment [§]	Best supportive care	BSC	X		

Therefore there are 14 treatments (including Best Supportive Care (BSC)) in the pool from which sequences can be generated and evaluated using the simulation model. There is a maximum possible sequence length of 13, due to it not being clinically appropriate to have a sequence including both abatacept and subcutaneous abatacept. The size of the decision space (S) is very complex to formally estimate given the various rules regarding the potential position of each treatment (e.g. rituximab cannot start a sequence). However, an upper bound (S_{\max}) can be derived using the following formula for k -permutations of n objects, where k is

* Based on current EMA licensed indication for adult patients with established moderate-severe rheumatoid arthritis

[†] All bDMARDs administered with concomitant MTX

[‡] TICORA in an intensive combination cDMARD strategy based on the TICORA study and treatment protocol.

[§] Always provided at the end of a sequence of active treatments for symptomatic relief

the length of the sequence (up to a maximum sequence length L) and n is the total number of treatments in a set.*

$$\sum_{k=1}^L \frac{n!}{(n-k)!} \quad [7.2]$$

In our case, $L = 12$ and $n = 13$. An upper bound is therefore:

$$S_{max} = \sum_{k=1}^{12} \frac{13!}{(13-k)!} \quad [7.3]$$

The maximum size of the decision space is provided in Table 7.4.

Table 7.4: Maximum size of the decision space

Sequence length (k)	Number of sequences (S)
0 (only BSC)	1
1	13
2	156
3	1,716
4	17,160
5	154,440
6	1,235,520
7	8,648,640
8	51,891,840
9	259,459,200
10	1,037,836,800
11	3,113,510,400
12	6,227,020,800
TOTAL (S_{max})	10,699,776,685

S_{max} is therefore over 10 billion sequences. This is far beyond the capacity of even a highly efficient patient level simulation to evaluate within a reasonable period of time, and likely to be far larger than any decision space that has been formally evaluated within a health economic evaluation.

Validation

The Simul8 RA sequence model was carefully and fully validated. The optimisation algorithm was programmed in Excel to allow defined sequences to be run as a group, which allows a standard fully incremental analysis of competing sequences to be estimated. The paper by Tappenden & Chilcott (2014) provides a list of practical procedures to follow when validating a health economic model.²⁹⁴ These are reported below in Table 7.5, and represent just the main ways in which the model was validated.

* (a selection of k objects from a list of n , where $k \leq n$), and where the order of selection matters and selections cannot be repeated.

Table 7.5: Validation tests

Area of model	Test applied
QALYs calculated properly	If utility is fixed to 1, then total QALYs equal total life years gained (LYG)
Model is valid	Set equal efficacy for all treatments – should result in optimal treatment sequence in ascending order of cost
Checked user defined formulae and functions within Simul8	Tested externally in MS Excel, and Simul8 ‘stepped into’ when run to check intermediate values were correct
All treatments set to the same efficacy	Each treatment run in a ‘1 treatment’ sequence and tested for equal QALYs and as expected total cost
All treatments set to equal efficacy and cost	Every evaluated sequence has identical costs and QALYs
Check ‘eligibility’ code	Generated 100 random sequences and manually checked the ‘error flag’ to ensure ineligible sequences had be correctly identified
HAQ progression modelling	Simul8 model built with ‘debug code’ so that intermediate values over time were output by the model. This allowed a visual debug of the progression of HAQ over time
Model breakpoints	Breakpoints were added to the model, so if impossible parameters/values or situations occurred, the model would break and an error message reported (e.g. negative costs, HAQ outside feasible range (0-3), utility over 1 etc.

The external validity of the results was tested by comparing to the original NICE RA model

Number of patient simulations

In a patient level simulation model, there is sampling error due to the variability between each simulated patient, often called first-order uncertainty.²⁹⁵ For an identical simulated patient, there are random events which will occur, meaning one simulated patient may respond to a particular treatment (for example), but an identical patient in the next simulation may not. By increasing the number of patients simulated, this first-order uncertainty can be reduced, and a better estimate of the true NMB can be derived. However, the reduction in uncertainty is at the expense of increased computational time. Therefore, identifying the appropriate number of patient simulations is of crucial importance.

When undertaking a cost-effectiveness analysis using a patient simulation model, the current recommendation by the NICE Decision Support Unit is to clearly justify the number of patients simulated, normally by using a graphical representation of the model output along with an estimate of the cumulative mean and its standard error.²⁹⁵ It is also important to ensure that the pseudorandom number generator (PNG) used within the simulation has no effect on the

results. Many software packages (including Simul8) allow a seed number to be used to initialise the PNG. When the same seed number is used, the PNG generates an identical string of pseudorandom numbers, which allows a model with the same parameters and seed number to generate identical output when run repeatedly. This functionality is useful for debugging a simulation model. When changing the PNG seed number and re-running the model, there should be negligible discrepancy in the results if sufficient individual patients have been simulated.

For the RA model, the output of interest is the Net Monetary Benefit (NMB). This enables a continuous numerical output from the model to act as the estimate of the objective function within the optimisation problem. Using an ICER would be problematic, because improving solutions cannot always be inferred by an improved (lower ICER) due to its ratio properties, and therefore a multi-criteria optimisation method would be required to incorporate cost and QALY output.

When comparing two solutions (sequences), the incremental NMB (iNMB) is simply the difference in the estimated NMB of the two solutions. Ideally, the simulation model should therefore be run with enough patients to ensure that an accurate estimation of the iNMB can be calculated between two similar solutions, and that the standard error of the mean iNMB is low enough for it not to affect the results.

Figure 7.6 to Figure 7.8 provide graphic output of running 50,000 patients in a simulation (approximately 4 minutes simulation time for each comparison). This is repeated across three sets of different PNG seeds. Reported are the cumulative mean iNMB, its standard error, and the 95% confidence interval. The treatment sequences compared were similar*, to ensure sufficient numbers of patients were selected to enable an accurate estimate of low values of iNMB (or, to rank sequences with similar NMB).

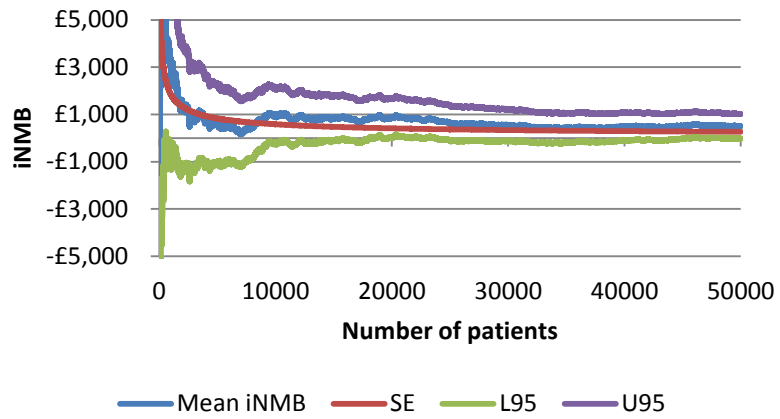
Each set of figures relates to a particular setting of uncertainty in the model. In Figure 7.6, the model has no probabilistic uncertainty, and uses only point estimates for each of these model parameters. Also the patients sampled are all identical, with no heterogeneity. This setting therefore corresponds to a deterministic analysis of a particular patient cohort (in this case, the mean characteristics of the patient population as taken from the BSRBR study (see earlier section). With probabilistic uncertainty accounted for, the estimated iNMB is stable, as is the corresponding confidence interval. However, this interval spans 0 in all three random number sets, and at 50,000 patients. Therefore the underlying patient variability is such that using

* {TICORA, ETN, RTX, TCZ, MTX, BSC} vs {TICORA, CTZ, RTX, TCZ, MTX, BSC}

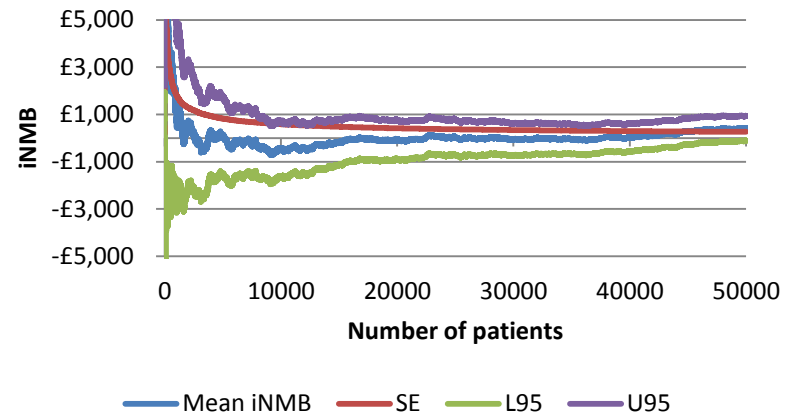
50,000 simulated patients is not sufficient to accurately confirm which sequence has the higher NMB, at 95% confidence.

The model used to generate Figure 7.7 also has no PSA, but patient level heterogeneity is reintroduced. With Figure 7.8, the model includes both patient heterogeneity and probabilistically sampled parameters. This model therefore is equivalent to the standard basecase model. Also provided in Table 7.7 to Table 7.9 are the results for different numbers of patients simulated.

Random Number Set 1



Random Number Set 2



Random Number Set 3

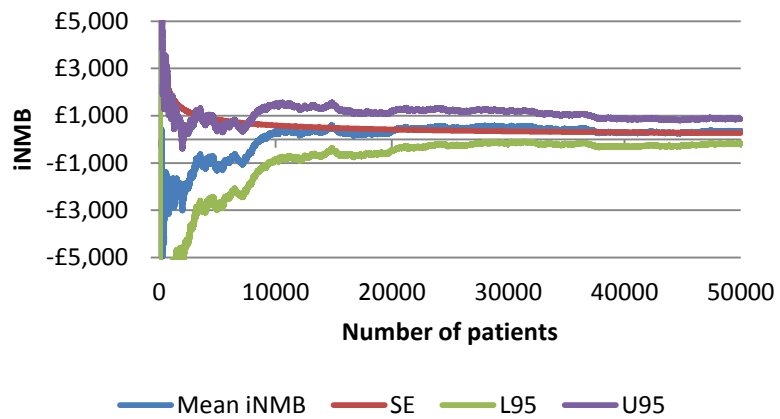
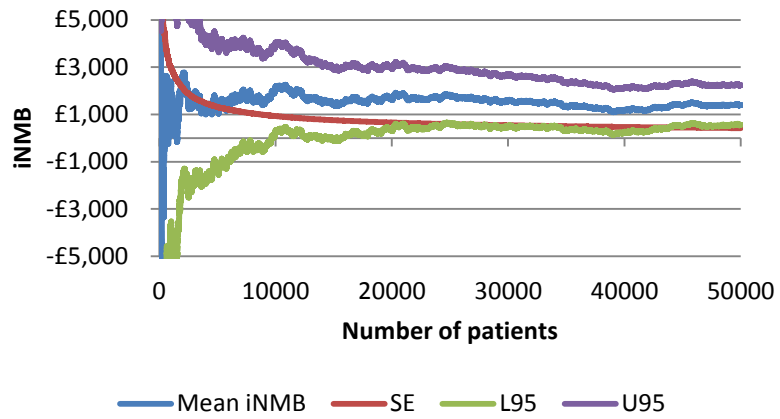
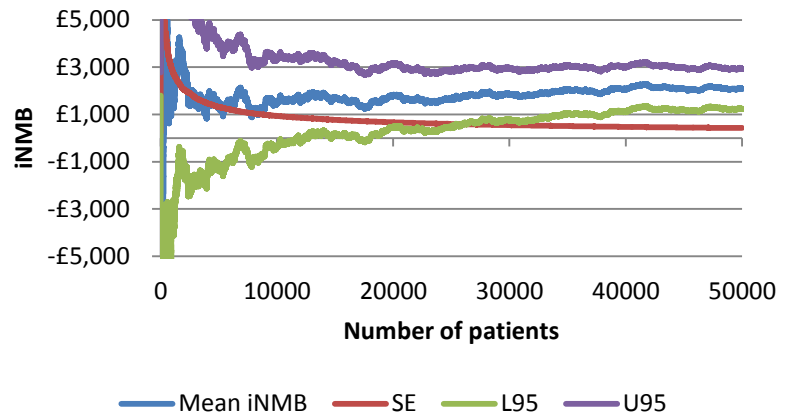


Figure 7.6: No patient heterogeneity. No PSA (Deterministic)

Random Number Set 1



Random Number Set 2



Random Number Set 3

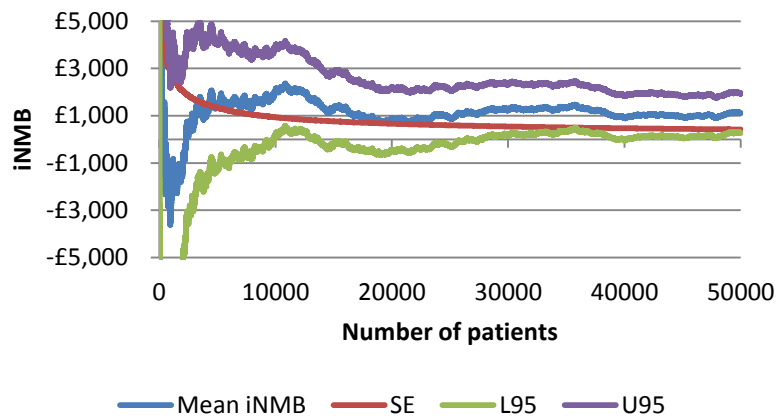
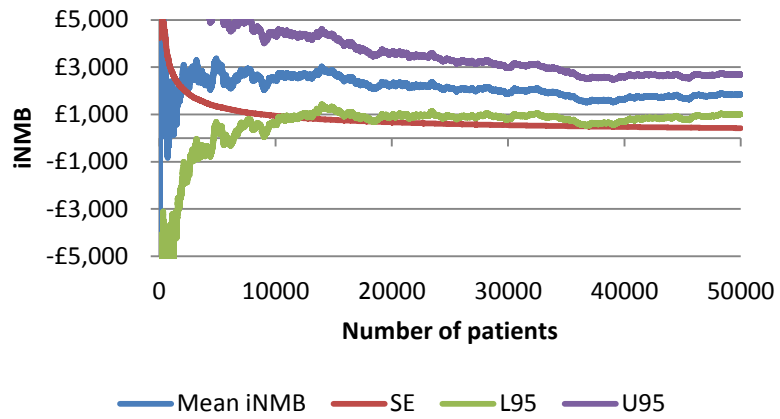
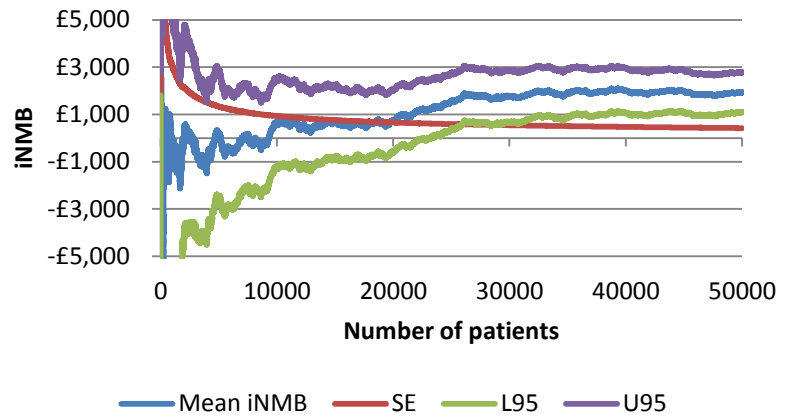


Figure 7.7: Patient heterogeneity. No PSA

Random Number Set 1



Random Number Set 2



Random Number Set 3

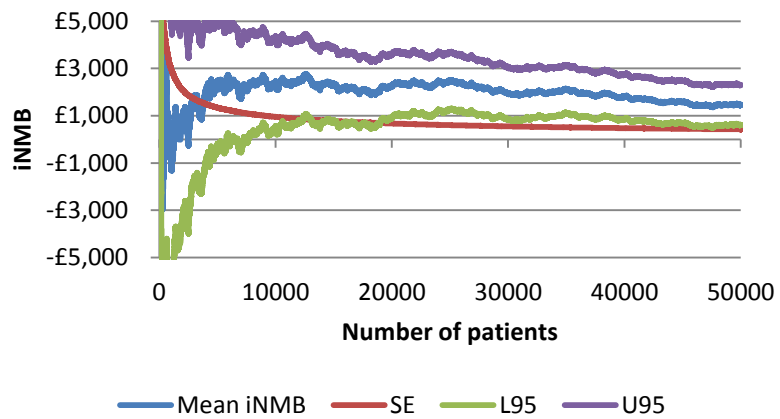


Figure 7.8: Patient heterogeneity. PSA (Basecase model)

Table 7.6: No patient heterogeneity. No PSA (Deterministic)

Patient Simulations	iNMB: Random Seed Set 1				iNMB: Random Seed Set 2				iNMB: Random Seed Set 3			
	Mean	SE	L95	U95	Mean	SE	L95	U95	Mean	SE	L95	U95
1,000	£2,572	£1,844	-£1,117	£6,260	£1,400	£1,847	-£2,294	£5,094	-£2,468	£1,880	-£6,228	£1,292
10,000	£992	£584	-£176	£2,160	-£463	£586	-£1,635	£710	£299	£588	-£877	£1,475
20,000	£860	£412	£36	£1,684	-£74	£414	-£902	£754	£347	£416	-£484	£1,179
30,000	£542	£336	-£129	£1,214	-£56	£339	-£735	£622	£482	£338	-£195	£1,158
40,000	£482	£292	-£102	£1,066	£13	£294	-£574	£601	£289	£293	-£298	£875
50,000	£500	£261	-£21	£1,022	£413	£262	-£111	£938	£344	£262	-£181	£868

Table 7.7: Patient heterogeneity. No PSA

Patient Simulations	iNMB: Random Seed Set 1				iNMB: Random Seed Set 2				iNMB: Random Seed Set 3			
	Mean	SE	L95	U95	Mean	SE	L95	U95	Mean	SE	L95	U95
1,000	£2,189	£3,005	-£3,820	£8,198	£2,167	£2,882	-£3,597	£7,931	-£2,745	£2,823	-£8,390	£2,900
10,000	£2,117	£928	£260	£3,974	£1,507	£935	-£363	£3,377	£2,063	£933	£197	£3,929
20,000	£1,628	£658	£312	£2,943	£1,804	£667	£469	£3,139	£876	£661	-£446	£2,198
30,000	£1,564	£543	£479	£2,650	£1,840	£546	£747	£2,933	£1,278	£543	£192	£2,364
40,000	£1,174	£470	£234	£2,114	£2,047	£473	£1,101	£2,994	£903	£469	-£36	£1,841
50,000	£1,398	£419	£559	£2,236	£2,077	£423	£1,230	£2,924	£1,104	£419	£267	£1,941

Table 7.8: Patient Heterogeneity. PSA (Basecase model)

Patient Simulations	iNMB: Random Seed Set 1				iNMB: Random Seed Set 2				iNMB: Random Seed Set 3			
	Mean	SE	L95	U95	Mean	SE	L95	U95	Mean	SE	L95	U95
1,000	£1,318	£2,919	-£4,521	£7,157	-£880	£3,077	-£7,033	£5,274	-£948	£3,047	-£7,041	£5,145
10,000	£2,523	£940	£643	£4,404	£684	£936	-£1,188	£2,557	£2,451	£951	£549	£4,354
20,000	£2,262	£667	£928	£3,595	£724	£657	-£591	£2,039	£2,249	£670	£909	£3,590
30,000	£1,898	£545	£808	£2,988	£1,760	£545	£671	£2,849	£1,918	£549	£820	£3,016
40,000	£1,631	£470	£690	£2,572	£1,998	£472	£1,054	£2,942	£1,817	£473	£871	£2,764
50,000	£1,853	£421	£1,011	£2,694	£1,929	£421	£1,087	£2,771	£1,445	£423	£599	£2,291

In Figure 7.6, the standard error is low compared with the mean iNMB, however the 95% confidence interval spans zero, which confirms that the patient variability in the model is too severe to allow a confident distinction between two very similar sequences. Certolizumab pegol and etanercept were specifically chosen as comparator treatments within a fixed sequence. Their efficacy and costs are not overly dissimilar, as shown in Table 7.9. The downstream sequences after these treatments are identical.

Table 7.9: Comparing certolizumab pegol to etanercept

Parameter		Certolizumab	Etanercept
Probability of a EULAR Response	None	0.400	0.274
	Moderate	0.274	0.318
	Good	0.325	0.409
Treatment cost ¹	Response period	£5,440.59	£4,663.36
	Subsequent annual cost	£9,326.73	£9,326.73

¹Excluding Patient Access Scheme (PAS) discounts

Adding patient heterogeneity and parameter uncertainty increases the standard error of the mean iNMB. It also results in a different estimate of iNMB compared with the deterministic results. This justifies the use of a probabilistic patient simulation model, because there is non-linearity between the input parameters and the NMB output. A further explanation is that, when patient heterogeneity is removed, only a female cohort of patients is run (females are significantly more likely than men to have RA). Therefore life expectancy and patient level predictions which include age (HAQ progression) lead to different results, compared with when patient heterogeneity is introduced and men are simulated.

From 40,000 patient simulations, the standard error is under £500 for all three random number sets, compared to an iNMB output of over £1000. The lower 95% confidence intervals are all above zero (Table 7.8). It was decided that 40,000 patient was a pragmatic number of patients to run. The simulation model takes approximately 3 minutes to run each sequence, and the desire is for each optimisation run to take days, rather than months to complete. The average standard error (£472) for this comparison will be used as an indicator of ‘indifference’ between compared sequences when their iNMB is within 2 standard errors (£944).

There was not believed to be any significant marginal benefit in running more patient simulations, at the expense of significant increases in computational time. For the purposes of tuning the algorithm, 2,000 patient simulations was selected as an adequate number for each tuning run of the optimisation algorithm, because the accuracy of the model output is less important for this objective. This lower number of patient simulations allowed a greater degree of tuning to be undertaken given the fixed time available.

Linkage to MS Excel

For the SO algorithm to be fully automated, a data link between the optimisation algorithm and the simulation model is required. Treatment sequence information must be passed to the patient simulation model, and the evaluated results from the patient simulation model passed back to the optimisation algorithm. Within one software package or computing language, this linkage would be fully integrated. However, although Simul8 is an ideal software package for discrete event simulation, its inbuilt Visual Logic language is not flexible enough to enable some of the statistical requirements of the patient simulation model. Also, it is not possible to generate and manipulate array variables, which are the ideal variable structure for dealing with treatment sequence data and applying the low level heuristics to implement the neighbourhood function.

Therefore, a decision was made to link Simul8 with Microsoft Excel. Excel provides standard spreadsheet functions, as well as access to the powerful Visual Basic for Applications (VBA) language. Excel can be linked to Simul8 via Windows COM, the interface standard for Microsoft software. This allows the full automation and manipulation of a Simul8 model from MS Excel, including within an Excel VBA macro. Therefore this allows the required data linkage between Excel and Simul8 (see Figure 7.9). An Excel VBA macro can theoretically provide an infinite cycle of data flow between Excel and Simul8, which is required for an SO algorithm.

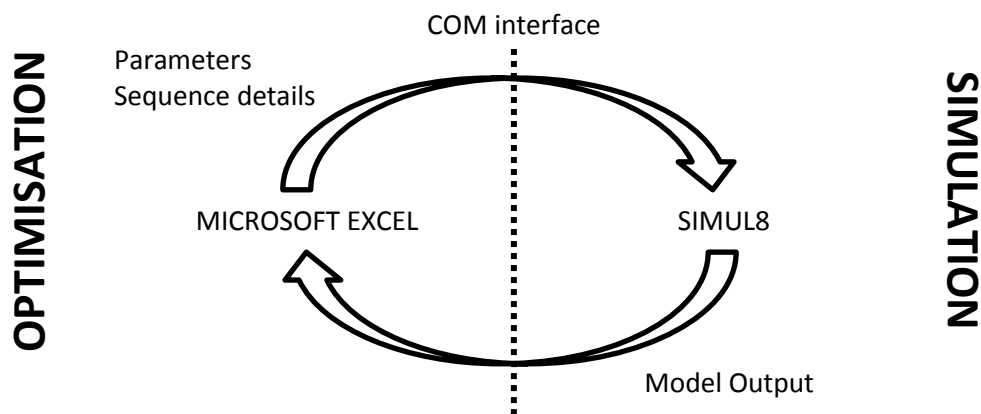


Figure 7.9: COM Interface

The linkage of two software packages via COM has introduced an element of inefficiency into the overall optimisation process. Each time the simulation model is updated with new sequence information, it has to be saved, and each time the Excel macro calls the simulation model to run, Simul8 has to open the saved simulation model, run the model, save the model and shut it down. It has been estimated that this process costs approximately five seconds per iteration of the optimisation algorithm.

CACHE SEARCH

The use of Excel for the optimisation algorithm allows evaluated sequences and their results to be stored within a spreadsheet (a cache). Due to common random numbers being used for each iteration of the patient simulation model, the repeated simulation of the same treatment sequence will result in identical results for the same patients.

Using a VBA search algorithm, the current sequence can be searched for in this cache to see if it has already been evaluated by the simulation model. If it has, then there is no need to run the simulation model, and instead the cached result can be inserted. This represents a form of 'simulated annealing with memory', and borrows principles from Tabu search, which is an alternative metaheuristic which incorporates memory. Unlike Tabu Search, the memory does not guide the search heuristic, but instead is just to avoid inefficient re-evaluation of sequences. The cache search heuristic was tested and even when the cache contained 10,000 evaluated sequences and their results, the cache took under 1 second to search, compared to at least 8 seconds to evaluate the sequence in the simulated model (assuming 2000 patient simulations). A simple process could be added to turn off the cache search when the cache search time was expected to take longer than the evaluation time, but it was not seen as necessary in this instance.

SIMULATED ANNEALING

Simulated annealing (SA) is a local search algorithm with the capability to escape from a local optima.²⁴⁷ It is a probabilistic metaheuristic which is frequently applied to global optimisation problems. It is so named due to its analogy to the physical annealing process undertaken by a crystalline solid. SA is most commonly applied to discrete optimisation problems, although it has the capability to optimise continuous problems. A more complete introduction to SA is provided in the methodological review of simulation optimisation methods reported in Chapter 6.

SA requires the decision problem to be encoded in a way which enables the decision space to be searched. Common methods of encoding include integer encoding, bit encoding and permutation encoding. These are detailed in Chapter 6. For this particular decision problem, permutation encoding is the more suitable, because each treatment within a sequence is unique and can only be used once only.

SA requires a neighbourhood function to be designed, which determines the movement of the algorithm from one solution to another 'nearby' solution. The neighbourhood function is a formal statement of what 'nearby' means in a problem specific context. The neighbourhood function therefore allows the movement of search to evaluate nearby solutions and find local

optima. At each iteration of the algorithm, two potential solutions are compared. These are the current solution in the algorithm (x), and a newly selected neighbour of the current solution (x'). If the newly selected neighbour is an improvement compared with the current solution ($g(x') - g(x) > 0$) then the new solution is accepted and it becomes the current solution (x). The algorithm then takes a neighbour from that new solution and repeats the process. However, a proportion of non-improving (inferior) solutions are also accepted based on an acceptance criterion. The acceptance of inferior solutions allows the possibility of escaping local optima. The acceptance criterion is a function of the temperature in the algorithm, which is a parameter which decreases in value as the algorithm iterates. The probability of accepting an inferior solution x' as the next solution is based on the Metropolis acceptance criterion, which is reported in Box 7.2: Metropolis acceptance criterion.²⁴⁸

$$P\{\text{Accept } x' \text{ as next solution}\} = \begin{cases} \exp\left[\frac{g(x') - g(x)}{t_k}\right], & \text{if } g(x') - g(x) \leq 0 \\ 1, & \text{if } g(x') - g(x) > 0 \end{cases}$$

t_i is the temperature parameter at iteration i such that:
 $t_i > 0$ for all i

Box 7.2: Metropolis acceptance criterion

Therefore the probability of accepting an inferior solution is a function of the current temperature, and the magnitude of the difference between the solutions. The probability decreases as the temperature decreases (moving the search from a random search to a local search as the algorithm iterates), and the probability increases as the difference between solutions decreases. As the algorithm iterates, the temperature is reduced. It can be reduced at every iteration (i), or after a fixed number of iterations called a temperature step length (n), where $n > i$. The number of steps that have occurred in the algorithm is step count (k). A graphical representation of the relationship between temperature, algorithm iterations and steps is provided in Figure 7.10

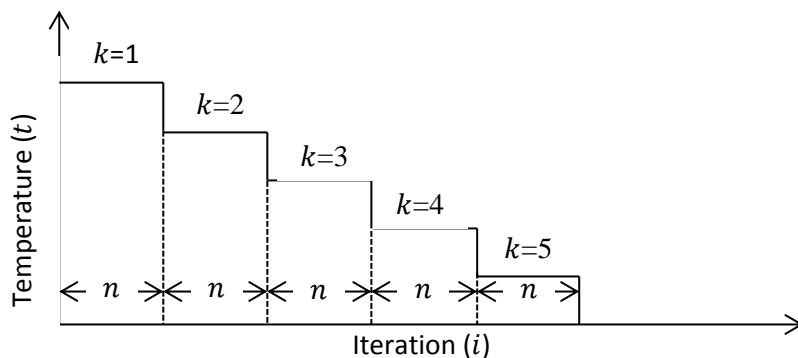


Figure 7.10: Temperature cooling steps and iterations

The generalised simulated annealing algorithm to maximise a discrete problem is provided in Box 7.3, and provides the basis for the applied algorithm.

<p>Parameters</p> <p>Current solution x</p> <p>Simulated estimate of objective function value (NMB) of solution $x = g(x)$</p> <p>Temperature t</p> <p>Best solution x_{best}</p> <p>Iteration i</p> <p>Temperature step count k</p> <p>Cooling function $\varphi(k)$</p> <p>Length of temperature step n</p> <p>Neighbourhood definition N</p> <p>Initialisation</p> <p>Initial solution x_0</p> <p>Initial Temperature t_0</p> <p>Starting iteration $i_0 = i = 1$</p> <p>Temperature step count $k = 1$</p> <p>Repeat</p> <p style="padding-left: 20px;">Repeat</p> <p style="padding-left: 40px;">Randomly select $x' \in N(x)$;</p> <p style="padding-left: 40px;">$\delta = g(x') - g(x)$;</p> <p style="padding-left: 40px;">If $\delta > 0$</p> <p style="padding-left: 60px;">$x = x'$;</p> <p style="padding-left: 40px;">Else</p> <p style="padding-left: 60px;">If $\exp\left(\frac{\delta}{t}\right) > \text{random}(0,1)$</p> <p style="padding-left: 80px;">$x = x'$;</p> <p style="padding-left: 60px;">Endif</p> <p style="padding-left: 40px;">Endif</p> <p style="padding-left: 40px;">If $g(x) > g(x_{best})$</p> <p style="padding-left: 60px;">$x_{best} = x$;</p> <p style="padding-left: 40px;">Endif</p> <p style="padding-left: 40px;">$i = i + 1$</p> <p style="padding-left: 20px;">Until $i = n$</p> <p style="padding-left: 20px;">Set $t = \varphi(k)$;</p> <p style="padding-left: 20px;">Set $i = 1$</p> <p>Until stopping conditions are met</p> <p>Output x_{best} as best solution found</p>
--

Box 7.3: Simulated annealing algorithm

It is common for implementations of SA to incorporate a restart process into the algorithm, and so this has been incorporated (but is not illustrated in Box 7.3). This allows the algorithm to automatically repeat once a stopping criterion has been met.²⁹⁶ This is useful to ensure that

the algorithm has converged on the best solution possible within that particular run (with multiple restarts) of the algorithm. The restart process normally selects the best solution from the prior iteration of the algorithm as the starting solution. This is analogous to the algorithm 'reheating' at a converged solution. Each run of the algorithm, and each consecutive restart of the algorithm, is called a 'round'. After multiple restarts, the overall best solution found is then used as a starting point for short local search, to ensure the overall best solution is the local optima. For tuning purposes, the algorithm can restart with a randomly generated starting solution to provide repeated tests of the performance of a particular tuning parameter.

Tuning the algorithm

The algorithm is contingent on the following parameters and components:

- Initial temperature t_0
- Cooling function (change in temperature) $\varphi(t)$
- Number of iterations at each temperature level k
- The construction of the neighbourhood of the problem N
- The rule(s) to determine when the algorithm stops

The appropriate specification of these parameters of the algorithm can be critical. In general, the algorithm is sensitive to all parameters and therefore successfully implementing the algorithm is contingent on finding appropriate parameter values. In this case, "success" is defined as the SA algorithm finding a good solution within a reasonable amount of time. A good solution is defined as having face validity, being a true local optimum, and not being improved upon when subjected to multiple restarts of the SA algorithm.

There are no hard rules which determine how these parameters are selected, and instead one relies on the user 'tuning' the algorithm to optimise its performance. Historically, parameterising an SA algorithm was a process of trial and error alongside low level heuristics and experience. However, newer methods have looked to automate the tuning process, and even newer hyperheuristic methods allow fully automated tuning of the metaheuristic. This includes selecting between SA and alternative metaheuristics. However, these methods are beyond the scope of this research problem, in particular because these methods are very computationally expensive (hundreds of iterations of the metaheuristic) and for a SO problem they are not likely to be feasible.

Each specification of the algorithm when tuning parameters represents a particular experiment. It would be possible to use a factorial design, to ensure that every possible combination of tuning parameter values is attempted. However, in this case this is not feasible.

Although SA has relatively few parameters compared to other metaheuristics, the range of potential parameter values and the time required to run the SA algorithm mean that this level of tuning is not feasible. Instead, each parameter will be tuned in turn, and then a full validation of the final set of parameter values to ensure the final specification is robust.

In reality, SA cannot guarantee that the true global optima will be identified within a feasible amount of time. However, re-running and restarting the process can build confidence that the best potential solution found by the algorithm is in fact the true global optimum. Also, running a local search at the end of the algorithm process allows verification that the best potential solution at the end of the SA algorithm is in fact at least a local optimum. As tuning is conducted, a 'current best' solution is likely to emerge. Therefore all references during the tuning process to the performance of the algorithm will be made to the 'current best' overall solution. The frequency in which it is found will be a measure of confidence that it is the true global optimum (the best sequence in the decision space), and if a particular specification of the algorithm does not perform well, then this judgement is based on its inability to efficiently find the 'current best'. The following sections provide in depth detail about the tuning of each component of the algorithm.

Neighbourhood function

The definition used for the neighbourhood function is crucial to the efficiency of any SA algorithm.²⁴⁷ When designing the neighbourhood function, there are two specific rules which must be enforced to guarantee convergence of the algorithm. Firstly, all potential solutions must be reachable (in a finite number of steps). Secondly, the neighbourhood must be symmetrical, and therefore backwards moves must be possible.

Much literature is devoted to highlighting how a neighbourhood function is required that imposes a smooth topology to the search space. In reality, that means a neighbourhood function with small changes to the solution, resulting in a relatively small change in the objective function. This also means that there are a large number of possible neighbourhood moves (the 'neighbourhood size'). However, for the RA treatment sequencing problem, any change to a solution (for example the addition or removal of a treatment, or swapping the position of two treatments) could in reality have a large impact on the objective function.

Research is ongoing about the size of the neighbourhood (e.g. how many neighbouring solutions can be drawn from for a given solution within the decision space).²⁴⁷ While some researchers believe a smaller neighbourhood to be favourable,²⁹⁷ others propose that the neighbourhood is as large as possible.²⁹⁸ A small neighbourhood means the algorithm takes a long time to search through the solution space, whereas a large neighbourhood corresponds

to random sampling within the solution space.²⁹⁹ The only area of agreement in this particular area of SA research is that the neighbourhood function tends to be a problem-specific issue.²⁴⁷

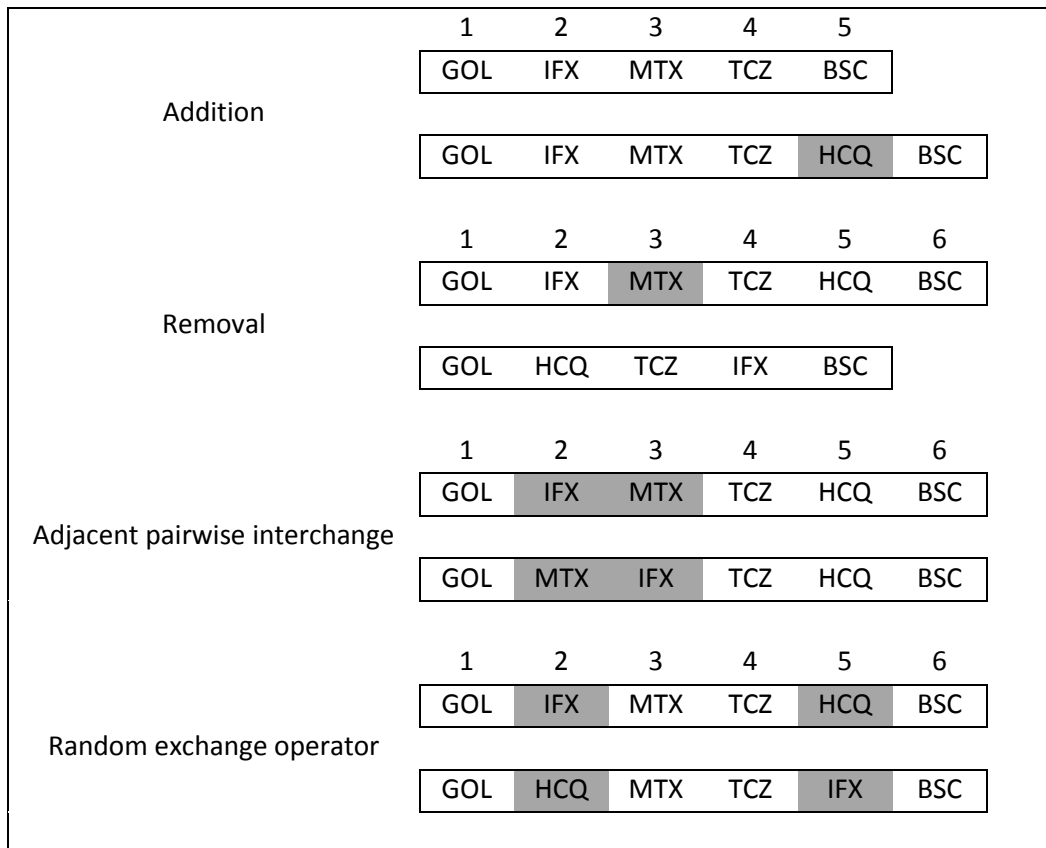
For this application of SA, a neighbourhood function was designed based on three possible changes (move operators) to the current solution: a treatment addition; a treatment removal; and a swap of two selected elements. These operators borrow concepts from the Genetic Algorithm literature regarding encoding a permutation problem. The order of the sequence is important, as well as the fact that a treatment cannot be used twice. These move operators respect those problem-specific conditions, in that they will not invalidate the permutation encoding of the problem. The first two change operators were relatively simple to implement. A list of unused treatments is maintained by the algorithm. For the treatment addition operator, a random treatment from this list is selected and inserted into the sequence at a randomly selected position. The downstream treatments all shift down by one position to accommodate the new intervention. For the treatment removal operator, a randomly selected treatment is removed from the sequence.

The swap operator was programmed with two options. Firstly, an adjacent pairwise interchange operator, with two adjacent treatments within the sequence swapped. Secondly, a random exchange operator, with two randomly selected treatments swapped. Box 7.4 provides a representation of the move operators for the neighbourhood function.

When problems can have an infeasible solution (for example, a treatment sequence with a treatment in a non-licensed position) then traditionally in optimisation these solutions are evaluated and then penalised. This process works well when both the evaluation and penalising of a solution can be done in a trivial amount of time. However, in SO, evaluating infeasible solutions can be very inefficient, especially when there are numerous rules regarding the eligibility of treatment in various positions. Therefore for this problem, VBA code was written to check that each sequence generated is eligible and feasible. Each time a sequence is generated, or a neighbour sequence generated, the code is run to check that a sequence is eligible. If necessary the move operation is re-run until an eligible sequence is generated. The eligibility check and re-running of a move operator proved to be many times more efficient than running the simulation model, evaluating the solution and penalising an infeasible solution.*

* Over 50 eligible sequences can be generated per second in MS Excel VBA. A simulation of 40,000 patients in the Simul8 model took over 2 minutes, with additional time required for data transfer between MS Excel and Simul8 for each algorithm iteration.

The addition, removal and adjacent pairwise interchange are all competing move operators when a neighbouring solution is required in the SA algorithm. Therefore, user defined probabilities are required so that the algorithm randomly selects a move operation. When a sequence has a length of 1 or 13 (the minimum and maximum possible sequence length) then the probability of a treatment addition or removal operator being selected is set to zero (as appropriate).



Box 7.4: Neighbourhood move operators

Three different probability values for the competing move operators were assigned to the algorithm and tested. Table 7.10 provides the results of tuning the neighbourhood function.

Each experiment for tuning was undertaken with a random starting sequence for each round (as opposed to using the best sequence identified in a previous round, which is the case for generating the final results). The stopping criteria applied were: when 2400 runs were completed; when the temperature was below 100; or 50 consecutive solutions were rejected. The initial temperature was set at 15,000. A 50 repetition geometric cooling schedule with a rate of 0.9 was used to decrease the temperature as the algorithm iterated.* These criteria and parameter values were selected based on some rapid experiments to establish settings for an initial algorithm that performed well.

* Further details regarding these particular cooling schedules is provided later in the chapter.

Neighbourhood function experiment 2 (see Table 7.10), with a medium probability of adding or removing a treatment to the sequence (20% for each operator) along with an adjacent pairwise interchange operator (compared to a random exchange operator), converged on the best performing sequence (across all tuning trials) and converged on this solution for each of the five rounds.

This set of neighbourhood function parameters (20% probability of add/remove and a 60% probability of an adjacent pairwise interchange) were taken forward when tuning the other components of the SA algorithm.

Table 7.10: Tuning the neighbourhood function

Neighbourhood function 1: 5% Addition, 5% Removal, 90% Adjacent pairwise interchange																
Round	Sequence (line)											NMB*	Costs*	QALYs*	Iteration when best solution found	
	1	2	3	4	5	6	7	8	9	10	11					
1	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1824
2	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	662
3	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1550
4	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	97
5	MTX	SSZ	HCQ	TICORA	GOL	RTX	CTZ	ETN	IFX	ADA	BSC		£140,657	£59,034	6.656	1586
Neighbourhood function 2: 20% Addition, 20% Removal, 60% Adjacent pairwise interchange																
1	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	845
2	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	796
3	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	798
4	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	315
5	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	813
Neighbourhood function 3: 40% Addition, 40% Removal, 20% Adjacent pairwise interchange																
1	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1019
2	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1125
3	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1148
4	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1284
5	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1029
Neighbourhood function 4: 20% Addition, 20% Removal, 60% Adjacent pairwise interchange																
1	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1096
2	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1074
3	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1277
4	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1496
5	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	499

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution identified by algorithm. bDMARDs in **bold**
 Algorithm settings: Initial temperature 15,000, 50 repetition schedule and geometric cooling at rate 0.9. Random initial solution for every round. Stop when 2400 runs completed (temperature < 100) or 50 consecutive solutions rejected. 2,000 simulation model runs for tuning.

Initial temperature

The initial temperature parameter value is of critical importance to the performance of the SA algorithm. If the initial temperature is low, then the search may be restricted to the region of the search space around the initial starting point, because the algorithm is less likely to escape a local optimum. If the initial temperature is too high, the SA algorithm will initially perform a ‘random walk’, with many inferior solutions accepted. This is inefficient, and it can lead to the algorithm terminating early if the total number of algorithm iterations is fixed.

The initial temperature parameter must be defined so almost any feasible solution is accepted during the first run of the algorithm. *A priori* information regarding the problem and the neighbourhood design can inform analytically derived values for the initial temperature. However, it is common to use the SA algorithm itself to provide an estimate of an appropriate initial temperature parameter value, in particular when information regarding the problem and neighbourhood design is difficult to derive. Both methods rely on the concept of an ‘acceptance ratio’ or ‘acceptance probability’, This is the number of worse solutions accepted at a given temperature, divided by the total number of worse solutions proposed by the algorithm. An acceptance ratio of 0.8 is suggested by Kirkpatrick, although he highlights that this parameter is problem-specific.²⁵⁷ The initial temperature parameter is a highly researched and contentious issue and beyond the scope of this thesis, however an alternative acceptance ratio of 0.6 will be tested, to see if the initial temperature it derives performs well compared to initial temperature derived by the established ratio of 0.8.

Based on a pre-specified acceptance ratio, two methods for deriving the initial temperature are commonly used; the van Laarhoven formulae, and a rapid warm-up algorithm.

Van Laarhoven formulae

Van Laarhoven et al. (1988) proposed a formulae to determine the initial temperature t_0 , where Ξ is the target acceptance ratio, and $|\Delta G_m|$ is the mean absolute change in the objective function of every proposed move from a set of iterations, m .³⁰⁰ This method of determining the initial temperature is well established.³⁰¹

$$t_0 = \frac{|\Delta G_m|}{\ln(\Xi^{-1})} \quad [7.4]$$

After running a random search algorithm, the mean change in objective function (NMB) after $m = 500$ iterations was £4,653. This resulted in estimates of initial temperature of 9,110 from a target acceptance ratio of 0.6, and 20,855 from a target acceptance ratio of 0.8 (Table 7.11).

Table 7.11: van Laarhoven formulae estimates of initial temperature parameter

Target acceptance ratio	Estimated initial temperature value (t_0)
0.6	9,110
0.8	20,855

Rapid warm-up algorithm

An alternative method to the van Laarhoven formula is to start the simulated annealing algorithm at a low initial temperature, and periodically increase the temperature until the acceptance ratio is met. This was undertaken by setting the initial temperature at 1,000. The algorithm then increased the temperature in increments of 1,000 every time the algorithm had random drawn 20 inferior potential solutions. Therefore the algorithm begins with a low temperature, and therefore a low likelihood of accepting a worse move. As the temperature is increased and worse moves begin to be accepted, the probability of accepting a worse move increases, until the predefined acceptance ratio target is achieved.

This was conducted over 10 rounds of the algorithm for both 0.6 and 0.8 acceptance ratio targets. The results are provided in Table 7.12.

Table 7.12: Rapid warm-up algorithm for initial temperature parameter

Round	Estimated initial temperature value (T_0)	
	Acceptance ratio target = 0.6	Acceptance ratio target = 0.8
1	5,000	11,000
2	6,000	10,000
3	7,000	8,000
4	6,000	7,000
5	4,000	9,000
6	5,000	11,000
7	5,000	11,000
8	4,000	7,000
9	7,000	11,000
10	5,000	7,000
	Maximum = 7,000	Maximum = 11,000
	Mean = 5,400	Mean = 9,200

Temperature increased in steps of 1,000

20 worse moves required at each temperature iteration

Each round uses a random initial solution

With an acceptance ratio of 0.6, the estimated initial temperature generated by the rapid warm-up algorithm (mean = 5,400) is smaller than the estimated initial temperature from the

van Laarhoven formula (9,110). This difference is also found with a higher acceptance ratio of 0.8, with a mean t_0 of 9,200 estimated by the rapid warm-up algorithm, compared to 20,855 as estimated by the van Laarhoven formula.

The algorithm was run with four alternative initial temperatures, which spanned the range of temperatures identified via the two methods and two acceptance criteria. The results are provided in Table 7.13. The results show that, in this decision problem, the algorithm was not sensitive to the alternative initial temperature values. The current best sequence found from any tuning run {MTX, SSZ, HCQ, TICORA, BSC} was found by all experiments and on every restart. To reduce the computational time required, the lowest value tested (7,000) was taken forward for the next stage of algorithm tuning.

It is common for the initial temperature to be tuned as part of the overall annealing component of the algorithm (including the cooling schedule and stopping rule). However, it was not feasible to do a full factorial experimental design to tune all possible combinations of all parameters, which is a common problem for SA, even though it has relatively few parameters. Therefore a decision was made to tune individual components and then test the final set of parameters for robustness. With several rounds and full sensitivity analysis conducted, any invalid solutions should be able to be identified and algorithms re-run if required.

Table 7.13: Tuning the initial temperature parameter

Initial temperature $T_0 = 7,000$ (Maximum possible runs per round = 1400)									
Round	Sequence (line)					NMB*	Costs*	QALYs*	Iteration when best solution found
	1	2	3	4	5				
1	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	821
2	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	893
3	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	640
4	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1073
5	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1052
Initial temperature $T_0 = 11,000$ (Maximum possible runs per round = 2150)									
1	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1029
2	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	347
3	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	757
4	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	575
5	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	850
Initial temperature $T_0 = 15,000$ (Maximum possible runs per round = 2400)									
1	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	845
2	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	796
3	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	798
4	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	315
5	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	813
Initial temperature $T_0 = 20,000$ (Maximum possible runs per round = 2550)									
1	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	470
2	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1180
3	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1393
4	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1243
5	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1010

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**

Algorithm settings: 50 repetition schedule and geometric cooling at rate 0.9. Random initial solution for every round. Stop when temperature < 100 or 50 consecutive solutions rejected. 2000 simulation model runs for tuning.

Cooling schedule

The strategy for cooling (reducing) the temperature as the algorithm iterates is of critical importance. When the temperature is very high, the algorithm operates as a random walk (all inferior solutions accepted), and when the temperature is very low, the algorithm operates as a local search (only improving solutions accepted). Therefore the rate of cooling acts as formal parameter to balance the exploration and exploitation of the search space by the algorithm.

Many cooling schedules exist for SA. Static cooling schedules are only dependent on globally defined parameters T_0 and the current number of algorithm iterations and temperature. Five commonly used static cooling schedules are provided in Table 7.14.

The algorithm iterates (i) from $1, \dots, l$ where l is a maximum number of iterations. The algorithm temperature steps down after a set of iterations, n . The cooling schedule of the algorithm is generally comprised of three parameters: the number of temperature steps, k , the number of iterations, i , and the temperature change parameter. With these three

parameters, the temperature at iteration i , $t(i)$, can be estimated. The cooling schedule can be dependent or independent of the current temperature parameter. A commonly used temperature dependent cooling schedule is a geometric cooling function, where the change in temperature is a function of both the temperature in the algorithm, and the temperature change parameter. Another commonly used independent cooling schedule is a linear cooling function, where the magnitude of the change in temperature is the same at every step, irrespective of the current temperature parameter.

Table 7.14: Static cooling schedules for simulated annealing

Type	Temperature at iteration i	Temperature change parameter(s)
Geometric	$t(k_i) = t_0 \alpha^k$	$0 \ll \alpha < 1$
Linear	$t(k_i) = t_0 - \eta k$	$\eta > 0$
Lundy and Mees	$t(k_i) = t(k) = t_{i-1}/(1 + \beta t_{i-1})$	$0 < \beta \ll 1$
Exponential	$t(k_i) = t_0 \exp(-\alpha tk^{\frac{1}{N}})$	$0 \ll \alpha < 1$ $N = \text{Neighbourhood size}$
Logarithmic	$t(k_i) = \frac{c}{\log(k + d)}$	$c \geq \text{Energy barrier}^*$ Typically $d = 1$

Where:

$t(k)$ is the current temperature at cooling 'step' k

$t(i)$ is the current temperature at algorithm iteration i

t_0 is the initial temperature

*Energy Barrier is the largest possible difference in the objective function between neighbouring solutions

A unique one-parameter cooling schedule was proposed by Lundy and Mees, with one iteration performed at each temperature step ($n = 1$, therefore $k = i$).³⁰² The temperature is reduced according to $t \rightarrow t/(1 + \beta t)$. Exponential and Logarithmic cooling schedules are also used, however they require further parameterisation which is specific to the problem and neighbourhood function. The logarithmic cooling schedule is analytically proven to converge on the true global optima, but only within a very long period of time, and for all purposes is not practical.

More advanced cooling schedules are adaptive, and incorporate learning and feedback from the decision space at each temperature level. Reheating of the temperature is sometimes incorporated when good solutions are not being found. Our particular implementation of the algorithm incorporates a restart procedure once a stopping rule has been met, which is a simple incorporation of a reheating process. A true reheating process was not applied due to time constraints.

The geometric, linear and Lundy & Mees cooling schedules are all commonly performed and well tested for their performance. An appropriate cooling schedule is problem-specific, and so tuning is required to identify those which perform well. These three cooling schedules will be tested, and if none perform well then the exponential will be tested.

As can be seen from Figure 7.11, the three selected cooling schedules all offer quite different performance when moving through the decision space. Given a fixed number of iterations to reach a target temperature, they all move from explorative searching at a high temperature to exploitative search at low temperature. The Lundy schedule is aggressive, moving quickly from a high temperature down to a relatively low temperature, where it spends a significant amount of time as an exploitative search. The Linear schedule is much slower at reducing the temperature. The geometric schedule operates in-between the two more extreme schedules.

The tuning of the cooling schedule will be constrained by the total running time that can be afforded for each experiment. The exact number of replications and temperature changes will be estimated on this basis, using data from previous tuning experience to identify the expected running time.

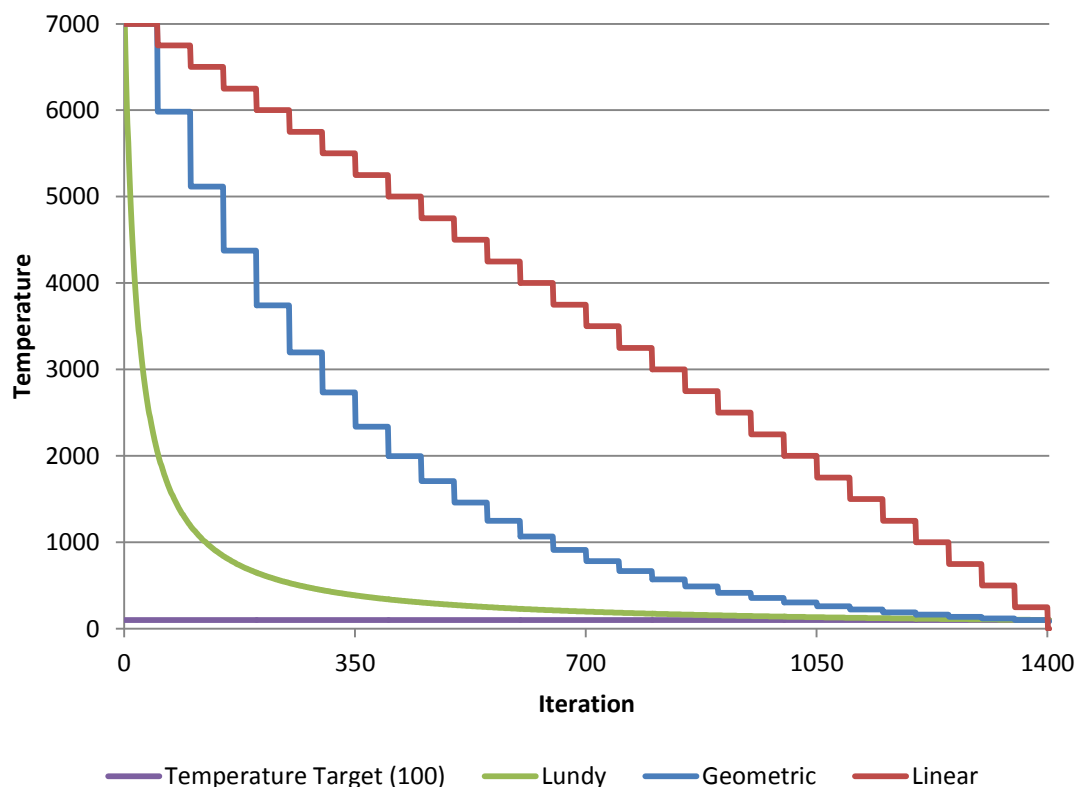


Figure 7.11: Lundy, Geometric and Linear Cooling Schedules (scaled for clarity)

Lundy and Mees cooling schedule

The Lundy and Mees Cooling Schedule was run with four different Lundy parameter values (β). These are reported in Table 7.15. The Lundy parameter values selected for tuning were solved for by setting four levels of maximum iterations (between approximately 500 and 2500 iterations) and estimating the corresponding β .

Table 7.15: Lundy parameter values

Lundy parameter (β)	Maximum number of algorithm iterations
4.0×10^{-6}	2500
6.5×10^{-6}	1517
1.4×10^{-5}	705
2.0×10^{-5}	493

With $t_0 = 7,000$ and final $t = 100$

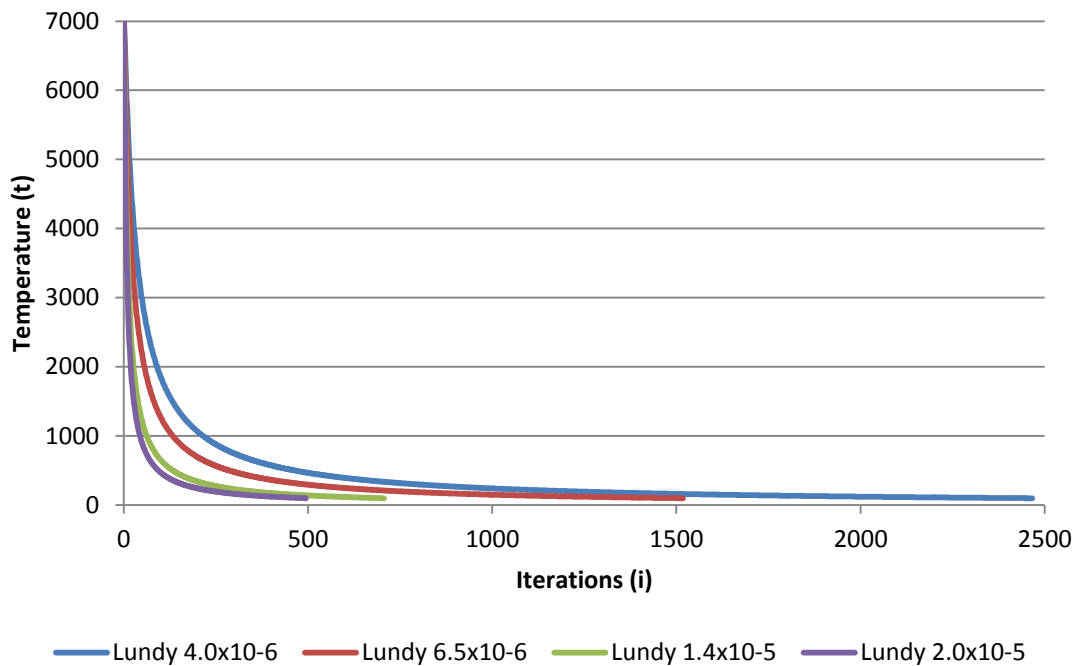


Figure 7.12: Lundy cooling schedules

The algorithm was run with each Lundy parameter value. The algorithm was restarted for each parameter five times, each time with a randomly generated starting solution. The results are provided in Table 7.16. With a smaller Lundy parameter (more iterations) the algorithm does not perform particularly well. When the algorithm does not find the best sequence {MTX, SSZ, HCQ, TICORA, BSC} it appears to have sorted the early part of the sequence reasonably well, with cDMARDs occupying early positions. However, many bDMARDs remain in the latter positions of the sequence. These sequences are more effective, but also much more costly,

resulting in lower NMB at £30k per QALY compared to the 'cDMARD only' sequence that is currently the best sequence identified.

However, with the largest Lundy parameter tested (2.0×10^{-5}), the algorithm performs well, finding the best solution (from all tuning experiments) in all five rounds. The results are surprising, and suggest that performance is not linear in the cooling schedules control parameter. Instead, it may be the case that this particular parameter value has found a particular 'sweet spot' in the decision space and neighbourhood function which allows the algorithm to extremely quickly find the best solution (between just 129 and 180 iterations).

Table 7.16: Lundy cooling schedule

Lundy Parameter 4.0×10^{-6} (Maximum possible runs per round = 2500)																	
Round	Sequence (line)													NMB*	Costs*	QALYs*	Iteration when best solution found
	1	2	3	4	5	6	7	8	9	10	11	12	13				
1	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	271
2	MTX	SSZ	HCQ	TICORA	GOL	RTX	CTZ	TCZ	ABT	ETN	ADA	IFX	BSC	£141,337	£59,717	6.702	508
3	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	361
4	MTX	SSZ	HCQ	TICORA	ADA	ETN	CTZ	RTX	ABTS	TCZ	BSC			£139,715	£59,874	6.653	438
5	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	277
Lundy Parameter 6.5×10^{-6} (Maximum possible runs per round = 1517)																	
1	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	248
2	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	107
3	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	148
4	SSZ	MTX	TICORA	HCQ	ADA	RTX	IFX	ETN	ABTS	GOL	TCZ	BSC		£136,518	£61,739	6.609	117
5	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	169
Lundy Parameter 1.4×10^{-5} (Maximum possible runs per round = 705)																	
1	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	141
2	MTX	HCQ	TICORA	SSZ	ADA	IFX	RTX	CTZ	BSC					£137,444	£61,153	6.620	73
3	MTX	HCQ	GOL	RTX	ADA	ETN	CTZ	TCZ	TICORA	SSZ	BSC			£128,255	£77,899	6.872	300
4	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	178
5	MTX	HCQ	TICORA	SSZ	CTZ	RTX	ETN	ADA	ABTS	BSC				£139,026	£59,325	6.612	260
Lundy Parameter 2.0×10^{-5} (Maximum possible runs per round = 493)																	
1	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	180
2	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	129
3	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	166
4	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	119
5	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	161

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7000, 1 repetition schedule. Random initial solution for every round. Stop when maximum runs completed (temperature < 100) or 50 consecutive solutions rejected. 2000 simulation model runs for tuning.

Linear cooling schedule

The algorithm was run with four alternative parameter settings for a linear cooling schedule. The two parameters for this cooling schedule (the number of repetitions at each cooling step, and the size of the change in temperature at each step) were estimated by setting an upper (2800) and lower (700) bound on the maximum number of iterations at each round. These values were selected based on the time required to run each algorithm, subject to the total time available for algorithm tuning.

A simple 2x2 experimental design was established, with number of repetitions either 50 or 100, and size of the decrement in temperature either 250 or 500. The corresponding cooling schedules are shown graphically in Figure 7.13.

The results are provided below in Table 7.17. The algorithm performed well across all four experiments, with 18 out of 20 rounds in total finding the current best sequence. The algorithm found the current best sequence in both experiments where the repetition parameter was set to 100, and was not sensitive in this case to the size of the temperature decrement. With 100 repetitions and a temperature change decrement of 500, the algorithm found the current best at all five independent rounds, and the maximum number of iterations of 1400 per restart is relatively efficient.

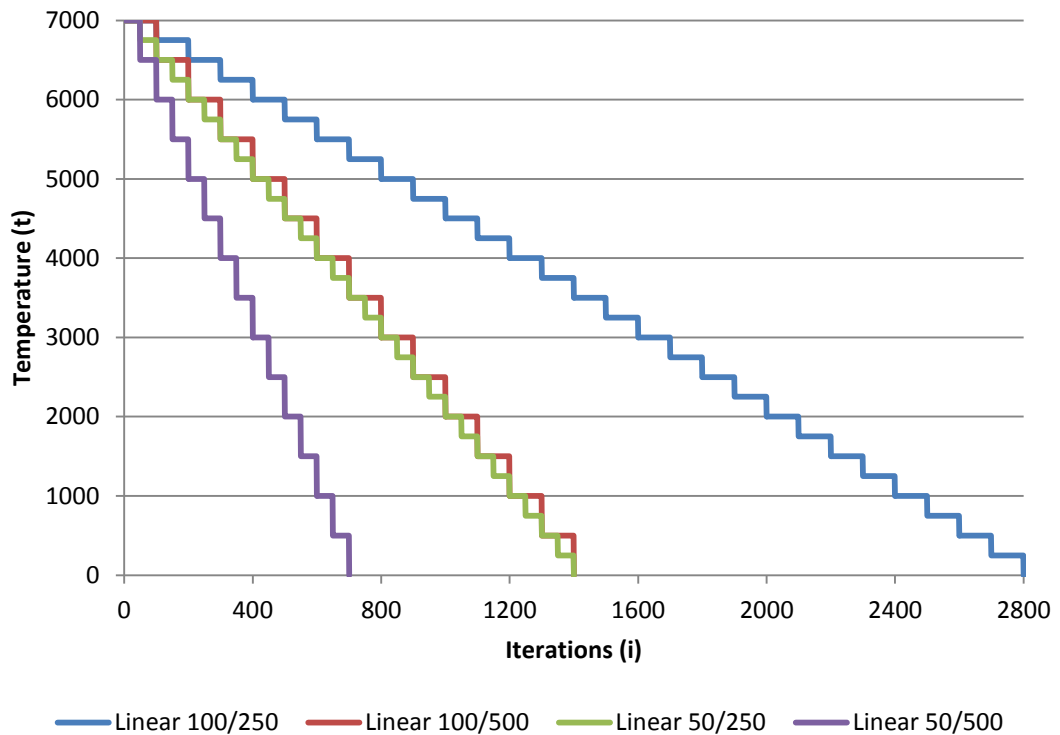


Figure 7.13: Linear cooling schedule

Table 7.17: Linear cooling schedule

Linear Cooling: 100 Repetitions per step. 250 Decrement per step. (Maximum possible iterations per round = 2800)																	
Round	Sequence (line)													NMB*	Costs*	QALYs*	Iteration when best solution found
	1	2	3	4	5	6	7	8	9	10	11	12	13				
1	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	634
2	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1237
3	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	326
4	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1344
5	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1337
Linear Cooling: 100 Repetitions per step. 500 Decrement per step. (Maximum possible iterations per round = 1400)																	
1	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	893
2	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1246
3	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	881
4	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1340
5	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1261
Linear Cooling: 50 Repetitions per step. 250 Decrement per step. (Maximum possible iterations per round = 1400)																	
1	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	893
2	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1246
3	MTX	SSZ	HCQ	TICORA	ADA	RTX	IFX	ETN	ABT	CTZ	GOL	BSC		£139,580	£60,277	6.662	881
4	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1340
5	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1261
Linear Cooling: 50 Repetitions per step. 500 Decrement per step. (Maximum possible iterations per round = 700)																	
1	HCQ	MTX	TICORA	SSZ	ADA	BSC								£138,612	£52,699	6.377	517
2	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	486
3	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	693
4	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	470
5	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	575

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7000. Random initial solution for every round. Stop when maximum runs completed (temperature < 100) or 50 consecutive solutions rejected. 2000 simulation model runs for tuning.

Geometric cooling schedule

The algorithm was run with four alternative parameter settings for a geometric cooling schedule. The first settings (50 repetitions and 0.9 cooling rate) were the initial settings for the tuning of the neighbourhood function. Three alternative experiments were established, with two levels for the repetition schedule (50 or 25), and a cooling rate of between 0.80 and 0.95. This resulted in a maximum number of algorithm iterations of between 1000 and 2075 (see Table 7.18).

Table 7.18: Geometric cooling parameters

Repetitions	Cooling rate	Maximum number of algorithm iterations
50	0.90	2050
25	0.95	2075
50	0.85	1350
50	0.8	1000

This range of cooling rate has been well established as a sensible range for a geometric cooling function in general.²⁴⁷ The resulting cooling schedules are shown graphically in Figure 7.14. The 50/0.9 schedules and 25/0.95 schedules are very similar in terms of resulting rate of change in the temperature parameter.

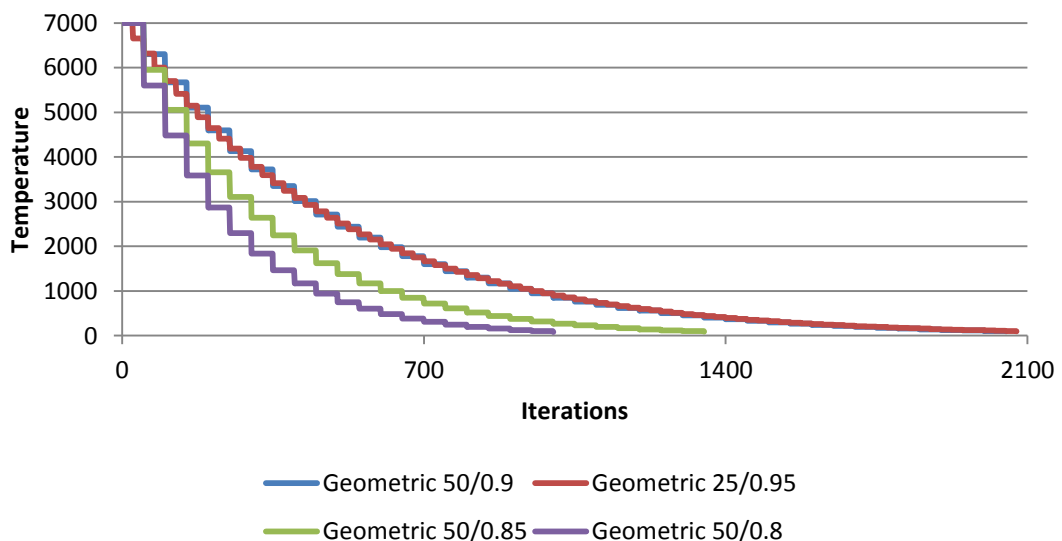


Figure 7.14: Geometric cooling schedule

The results are provided below in Table 7.19: Geometric cooling schedule. The algorithm performed well across all four experiments, with 17 out of 20 restarts finding the current best sequence. However, there are counterintuitive results which should be highlighted. Firstly, when comparing the same rate of cooling (0.85), the algorithm performed better with fewer

step repetitions (25 compared to 50). This result is similar to that found with the Lundy algorithm, where the more aggressive algorithm achieved greater success.

The algorithm also found that when comparing the two similar schedules (50/0.9 and 25/0.95), the algorithm failed to find the current best sequence in two of the five restarts with the latter parameterisation.

Table 7.19: Geometric cooling schedule

Geometric Cooling: 50 Repetitions per step. Cooling rate 0.9 (Maximum possible runs per round = 2050)																	
Round	Sequence (line)													NMB*	Costs*	QALYs*	Iteration when best solution found
	1	2	3	4	5	6	7	8	9	10	11	12	13				
1	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	821
2	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	893
3	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	640
4	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1073
5	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1052
Geometric Cooling: 25 Repetitions per step. Cooling rate 0.95 (Maximum possible runs per round = 2075)																	
1	HCQ	MTX	SSZ	TICORA	BSC									£143,321	£39,315	6.088	234
2	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1106
3	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	399
4	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	887
5	SSZ	MTX	HCQ	TICORA	BSC									£143,523	£39,114	6.088	487
Geometric Cooling: 50 Repetitions per step. Cooling rate 0.85 (Maximum possible runs per round = 1350)																	
1	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	473
2	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	562
3	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	227
4	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	505
5	MTX	SSZ	HCQ	TICORA	GOL	ABTS	CTZ	ETN	ADA	IFX	RTX	TCZ	BSC	£141,984	£59,813	6.727	899
Geometric Cooling: 50 Repetitions per step. Cooling rate 0.80 (Maximum possible runs per round = 1200)																	
1	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	543
2	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	366
3	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	767
4	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	383
5	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	287

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**

Algorithm settings: Initial temperature 7000. Random initial solution for every round. Stop when maximum runs completed (temperature < 100) or 50 consecutive solutions rejected. 2000 simulation model runs for tuning.

In Table 7.20, each chosen cooling schedule for each method is run with 10 random restarts. The Lundy cooling schedule finds the current best solution in six out of ten attempts, but is also relatively quick at finding these solutions (within 300 iterations). On the other hand, the Linear cooling schedule performs similarly, but is much slower (maximum number of iterations per restart is 1400). The Geometric cooling schedule is the best performing, in this particular experiment. It finds the current best solution in seven out of ten restarts, and is faster than the linear schedule, but slower than the Lundy schedule.

Similar results were found when using the current best solution for each restart. Each schedule performed well, and identified the current best well within ten restarts. These results are presented in Table 7.21.

Table 7.20: Final tuning - random starting solution for each round

Lundy Parameter 2.0x10 ⁻⁵ (Maximum possible runs per round = 493)																	
Round	Sequence (line)													NMB*	Costs*	QALYs*	Iteration when best solution found
	1	2	3	4	5	6	7	8	9	10	11	12	13				
1	MTX	SSZ	HCQ	ETN	RTX	TICORA	ABTS	TCZ	ADA	IFX	GOL	CTZ	BSC	£136,969	£66,605	6.786	131
2	SSZ	HCQ	TICORA	TCZ	RTX	GOL	CTZ	IFX	ETN	MTX	ADA	ABTS	BSC	£130,555	£72,455	6.767	181
3	MTX	SSZ	TICORA	ABT	CTZ	ADA	HCQ	IFX	BSC					£133,987	£66,842	6.694	174
4	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	228
5	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	170
6	MTX	SSZ	HCQ	TICORA	CTZ	IFX	GOL	BSC						£139,997	£58,243	6.608	101
7	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	119
8	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	87
9	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	237
10	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	109
Linear Cooling: 100 Repetitions per step. 500 Decrement per step. (Maximum possible runs per round = 1400)																	
1	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	339
2	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	908
3	MTX	SSZ	HCQ	BSC										£141,133	£34,748	5.862	27
4	SSZ	HCQ	TICORA	BSC										£141,304	£39,870	6.039	572
5	MTX	SSZ	HCQ	TICORA	CTZ	BSC								£142,545	£49,168	6.390	1236
6	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	908
7	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	996
8	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1049
9	HCQ	SSZ	MTX	TICORA	BSC									£143,317	£39,319	6.088	247
10	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	1226
Geometric Cooling: 50 Repetitions per step. Cooling rate 0.80 (Maximum possible runs per round = 1200)																	
1	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	748
2	SSZ	MTX	HCQ	TICORA	ETN	CTZ	GOL	IFX	ADA	RTX	TCZ	ABT	BSC	£139,686	£60,567	6.675	609
3	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	538
4	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	262
5	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	249
6	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	464
7	HCQ	MTX	TICORA	SSZ	CTZ	BSC								£141,849	£50,742	6.420	301
8	MTX	SSZ	HCQ	TICORA	CTZ	GOL	ETN	RTX	ADA	TCZ	ABTS	IFX	BSC	£141,212	£59,772	6.699	564
9	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	494
10	MTX	SSZ	HCQ	TICORA	BSC									£143,530	£39,106	6.088	334

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7000. Random initial solution for every round. Stop when maximum runs completed (temperature < 100) or 50 consecutive solutions rejected. 2000 simulation model runs for tuning.

Table 7.21: Final tuning - current best solution for each round

Lundy Parameter 2.0×10^{-5} (Maximum possible runs per round = 493)														
Round	Sequence (line)										NMB*	Costs*	QALYs*	Iteration when best solution found
	1	2	3	4	5	6	7	8	9	10				
1	MTX	SSZ	HCQ	TICORA	CTZ	GOL	RTX	ADA	IFX	BSC	£139,748	£59,705	6.648	153
2	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	110
3	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
4	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
5	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
6	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
7	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
8	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
9	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
10	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
Linear Cooling: 100 Repetitions per step. 500 Decrement per step. (Maximum possible runs per round = 1400)														
1	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	69
2	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
3	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
4	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
5	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
6	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
7	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
8	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
9	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
10	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
Geometric Cooling: 50 Repetitions per step. Cooling rate 0.80 (Maximum possible runs per round = 1200)														
1	MTX	SSZ	TICORA	BSC							£141,619	£39,555	6.039	430
2	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	208
3	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
4	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
5	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
6	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
7	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
8	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
9	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1
10	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7000. Previous best solution for every round. Stop when maximum runs completed (temperature < 100) or 50 consecutive solutions rejected. 2000 simulation model runs for tuning.

The Lundy cooling schedule continued to show good performance compared with the linear and geometric cooling schedules. All three converged to the same optima and the Lundy schedule was faster than the Geometric schedule at finding this optima. The Lundy cooling schedule was run one more time under the same conditions to confirm its performance. On this occasion it converged on the best-identified optima in the first restart, and after just 97 iterations (see Table 7.22).

Table 7.22: Final test of Lundy cooling schedule

Lundy Parameter 2.0×10^{-5} (Maximum possible runs per round = 493)									
Round	Sequence (line)					NMB*	Costs*	QALYs*	Iteration when best solution found
	1	2	3	4	5				
1	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	97
2	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1
3	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1
4	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1
5	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1
6	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1
7	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1
8	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1
9	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1
10	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7000. Previous best solution for every round. Stop when maximum runs completed (temperature < 100) or 50 consecutive solutions rejected. 2000 simulation model runs for tuning.

Efficiency of the algorithm

The Lundy (2.0×10^{-5}) algorithm was taken forward to conduct the basecase analysis. Tuning experiments were conducted to attempt to improve the efficiency of the algorithm.

The algorithm will continue to iterate indefinitely until a stopping criterion is met. The implemented algorithm contains three stopping rules, which can each be independently activated:

1. Maximum number of algorithm iterations
2. Minimum temperature value
3. Maximum consecutive rejected solutions

Stopping rule 1 acts as an overall control on the algorithm and is always active. Stopping rules 2 and 3 avoid inefficient running of the algorithm when stuck at a local optimum. When the temperature is zero or very near to zero, then the algorithm is performing a local search. When the algorithm is rejecting every consecutive neighbouring solution, then it is a signal that the algorithm has converged on a solution.

At the end of the optimisation process, including every restart, the algorithm has been designed to perform a short local search (temperature = 0, iterations = 100) to check that the best identified solution across all restarts is likely to be the true local optimum.

Tuning was conducted on stopping rule 3, the maximum consecutive number of rejected solutions. This was conducted to improve the overall efficiency of the algorithm while attempting to minimise any reduction in its performance.

Table 7.24 to Table 7.26 report the effect of reducing the number of consecutive failed attempts. From 50 to 15, there is no noticeable effect on the total number of iterations run per round. The results from the algorithm are consistent.

When reducing the number of consecutive failed attempts to 10 (Table 7.26), the algorithm continues to perform well, however it only requires approximately half the total number of iterations (see Table 7.23). Reducing this again to five (Table 7.26) has an impact on the performance of the algorithm (in two attempts, the algorithm required four and two rounds respectively, before finding the current best solution). The improved efficiency from 10 to 5 maximum consecutive failed attempts is at a cost of the reduced performance in the algorithm. For this reason, 10 consecutive failed attempts was selected as an appropriate stopping rule to avoid inefficient running of the algorithm. This decision was based on there not being any obvious increase in the rounds required to find the current best (3 rounds), but it also offered a significant decrease in the total number of iterations required (from 4448 to 2723).

Table 7.23: Tuning the maximum number of consecutive failed attempts

Maximum number of consecutive failed attempts	Round when current best solution identified	Iteration when best solution identified	Total number of iterations
50	2	673	4930
40	1	347	4930
30	3	791	4658
20	1	170	4930
15	1	363	4448
10	3	413	2723
5	4	313	646

Table 7.24: Maximum consecutive failed attempts (50 - 30)

Maximum consecutive failed attempts = 50																	
Round	Sequence (line)												NMB*	Costs*	QALYs*	Iteration when best solution found	Total Iterations
	1	2	3	4	5	6	7	8	9	10	11	12					
1	MTX	SSZ	HCQ	TICORA	GOL	ADA	RTX	ETN	TCZ	ABTS	CTZ	BSC	£139,969	£60,369	6.678	453	493
2	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	180	493
3	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
4	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
5	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
6	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
7	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
8	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
9	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
10	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
Maximum consecutive failed attempts = 40																	
1	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	347	493
2	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
3	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
4	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
5	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
6	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
7	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
8	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
9	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
10	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
Maximum consecutive failed attempts = 30																	
1	MTX	HCQ	TICORA	GOL	SSZ	RTX	ABT	BSC					£135,523	£63,120	6.621	84	221
2	MTX	SSZ	HCQ	TICORA	ADA	RTX	BSC						£140,912	£54,953	6.529	381	493
3	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	77	493
4	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
5	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
6	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
7	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
8	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
9	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493
10	MTX	SSZ	HCQ	TICORA	BSC								£143,530	£39,106	6.088	1	493

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7000. Previous best solution for every round. Stop when maximum runs completed (temperature < 100). 2000 simulation model runs for tuning. Lundy Parameter 2.0×10^{-5} (Maximum possible runs per restart = 493)

Table 7.25: Maximum consecutive failed attempts (20 and 15)

Maximum consecutive failed attempts = 20										
Round	Sequence (line)					NMB*	Costs*	QALYs*	Iteration when best solution found	Total Iterations
	1	2	3	4	5					
1	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	170	493
2	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
3	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
4	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
5	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
6	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
7	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
8	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
9	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
10	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
Maximum consecutive failed attempts = 15										
1	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	363	493
2	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
3	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	388
4	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
5	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
6	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
7	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
8	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	390
9	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	493
10	MTX	SSZ	HCQ	TICORA	BSC	£143,530	£39,106	6.088	1	219

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7000. Previous best solution for every round. Stop when maximum runs completed (temperature < 100). 2000 simulation model runs for tuning. Lundy Parameter 2.0×10^{-5} (Maximum possible runs per restart = 493)

Table 7.26: Maximum consecutive failed attempts (10 and 5)

Maximum consecutive failed attempts = 10															
Round	Sequence (line)										NMB*	Costs*	QALYs*	Iteration when best solution found	Total Iterations
	1	2	3	4	5	6	7	8	9	10					
1	MTX	TICORA	ADA	RTX	GOL	SSZ	IFX	HCQ	ABT	BSC	£127,682	£76,127	6.794	61	104
2	HCQ	MTX	SSZ	TICORA	GOL	RTX	ETN	BSC			£143,530	£39,106	6.088	101	113
3	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	196	272
4	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1	493
5	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1	164
6	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1	387
7	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1	274
8	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1	343
9	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1	329
10	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1	244
Maximum consecutive failed attempts = 5															
1	MTX	SSZ	HCQ	ABTS	ADA	GOL	BSC				£134,989	£66,785	6.726	39	56
2	MTX	SSZ	TICORA	HCQ	ABTS	GOL	ADA	RTX	BSC		£136,215	£61,828	6.601	5	59
3	MTX	SSZ	HCQ	TICORA	CTZ	BSC					£142,545	£49,168	6.390	41	66
4	MTX	SSZ	HCQ	TICORA	CTZ	BSC					£142,545	£49,168	6.390	1	74
5	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	58	79
6	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1	142
7	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1	9
8	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1	40
9	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1	22
10	MTX	SSZ	HCQ	TICORA	BSC						£143,530	£39,106	6.088	1	99

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7000. Previous best solution for every round. Stop when maximum runs completed (temperature < 100). 2000 simulation model runs for tuning. Lundy Parameter 2.0×10^{-5} (Maximum possible runs per restart = 493)

7.4 RESULTS

BASECASE ANALYSIS

The basecase analysis was run with the tuned parameter settings for the SA algorithm. These were the Lundy cooling schedule with a parameter of 2.0×10^{-5} , an initial temperature of 7,000 and the algorithm stopping after 10 consecutive rejected solutions or a temperature less than 100. The algorithm was run for 10 rounds with the best solution identified used at the initial solution for each restart. The results of the basecase analysis are provided in Table 7.27, along with the final parameter values. 40,000 patient simulations were run for each iteration of the algorithm, and the maximum possible number of iterations for each restart of the algorithm was 493. The algorithm converged on the current best solution during the first round after just 188 iterations. Nine subsequent rounds of the algorithm did not find any improvement.

The current best solution is exclusively a cDMARD sequence, given a willingness to pay of £30,000 per QALY gained. The best sequence found is {MTX, SSZ, HCQ, TICORA, BSC}. This sequence is consistent with the best sequence found in any of the tuning experiments. Therefore it is likely to be the global optima, although it is not possible to be certain.

Table 7.27: Basecase results

Basecase analysis										
Round	Sequence (line)					NMB*	Costs*	QALYs*	Iteration when best solution found	Total Iterations
	1	2	3	4	5					
1	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	188	291
2	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	183
3	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	60
4	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	362
5	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	493
6	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	261
7	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	166
8	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	264
9	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	173
9	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	180
END	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	100

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7,000. Previous best solution for every round. Stop when maximum runs completed (temperature < 100) or 10 consecutive solutions rejected. 40,000 simulation model runs for tuning. Lundy cooling schedule with parameter 2.0×10^{-5} (Maximum possible runs per restart = 493)

2,433 total iterations of the algorithm were performed until the algorithm terminated.* From rounds 2 to 10, a better solution was not identified, however the algorithm still iterated (between 60 and 493 times) due to accepting worse solutions at high temperatures. The final 'end run' local search was conducted (100 iterations) after 10 rounds, and confirmed the current best solution was a local optima.

For every potential solution evaluated by the simulation optimisation algorithm, the NMB was stored within a cache. Therefore the final cache can be ranked in order of performance to identify the potential solutions that were close to the current best solution found. This was undertaken and the results for the top 20 solutions reported. This ranked top 20 of the cache is displayed in Table 7.28. The top 18 are highlighted in green, because their own performance falls within 2 estimated standard errors of the current best solution, indicating that there may be indifference between these solutions due to the patient level variance. The top 20 solutions are all composed of purely cDMARD sequences. In fact, all bar two sequences are within the neighbourhood of the current best, because they can all be reached by one of the three neighbourhood operators.

* This required ~15 hours on a Windows 64-bit PC with i5 Quad Core @ 2.00GHZ and 8GB RAM

Table 7.28: Basecase results (top 20 ranked)

Solutions (ranked by NMB)	Sequence (line)					NMB*	Costs*	QALYs*
	1	2	3	4	5			
1	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006
2	SSZ	MTX	HCQ	TICORA	BSC	£141,816	£38,368	6.006
3	MTX	HCQ	SSZ	TICORA	BSC	£141,752	£38,433	6.006
4	SSZ	HCQ	MTX	TICORA	BSC	£141,740	£38,445	6.006
5	HCQ	MTX	SSZ	TICORA	BSC	£141,609	£38,575	6.006
6	HCQ	SSZ	MTX	TICORA	BSC	£141,605	£38,579	6.006
7	MTX	SSZ	TICORA	HCQ	BSC	£140,342	£39,604	5.998
8	SSZ	MTX	TICORA	HCQ	BSC	£140,334	£39,612	5.998
9	MTX	HCQ	TICORA	SSZ	BSC	£140,214	£39,732	5.998
10	SSZ	HCQ	TICORA	MTX	BSC	£140,199	£39,747	5.998
11	MTX	SSZ	HCQ	BSC		£140,152	£34,586	5.825
12	SSZ	MTX	HCQ	BSC		£140,145	£34,594	5.825
13	MTX	HCQ	SSZ	BSC		£140,080	£34,658	5.825
14	HCQ	MTX	TICORA	SSZ	BSC	£140,070	£39,876	5.998
15	SSZ	HCQ	MTX	BSC		£140,068	£34,670	5.825
16	HCQ	SSZ	TICORA	MTX	BSC	£140,062	£39,884	5.998
17	HCQ	MTX	SSZ	BSC		£139,937	£34,801	5.825
18	HCQ	SSZ	MTX	BSC		£139,933	£34,805	5.825
19	MTX	TICORA	SSZ	HCQ	BSC	£138,136	£43,095	6.041
20	SSZ	TICORA	MTX	HCQ	BSC	£138,123	£43,108	6.041

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Algorithm settings: Initial temperature 7000. Previous best solution for every round. Stop when maximum runs completed (temperature < 100) or 50 consecutive solutions rejected. 40000 simulation model runs for tuning. Lundy Parameter 2.0×10^{-5} (Maximum possible runs per restart = 493). Highlighted green if within 2 standard errors (£944) of best identified sequence

SCENARIO ANALYSIS

Alternative cost-effectiveness thresholds (λ)

By varying lambda (the cost-effectiveness threshold), the net monetary benefit of each sequence will change, *ceteris paribus*.

When lambda is zero, cost-minimisation is assumed, because any benefits are not valued at all. When lambda is very large, the decision problem tends towards benefit-maximisation, because the monetary value of any benefit dwarfs any costs associated with each sequence.

When using the net monetary benefit framework for an economic evaluation, it is important to test the results and explore how they are affected by varying lambda, especially when the ICER

between comparators may fall very close to lambda (and therefore the incremental NMB will be very close), because a decision-maker's belief about 'true lambda' may be very important when determining the true optimal solution. For NICE, their cost-effectiveness threshold is defined as £20,000 - £30,000 per QALY gained, and there are instances where technologies can be approved with an ICER above £30,000 per QALY gained. Therefore a single defined lambda is not available for the purposes of NICE. Therefore scenario analysis with alternative values of lambda is important. Scenario analysis was performed with lambda values of £0, £20,000, £50,000, £100,000 and £1,000,000.

Table 7.29 provides the scenario analysis results for lambda values of £0, £20,000 and £50,000 per QALY. The cost minimising (lambda of £0 per QALY) strategy was 'do nothing', with the optimal sequence purely best supportive care. At lambda of £20,000 per QALY, the results were consistent with the basecase analysis (lambda of £30,000 per QALY) with the best sequence found being {MTX, SSZ, HCQ, BSC}. When lambda is increased to £50,000 per QALY, bDMARDs contribute to the best identified sequence, but only after all cDMARD treatments have been used. The best sequence is {SSZ, HCQ, MTX, TICORA, CTZ, RTX, ETN, ADA, ABTS, GOL, BSC}.

Table 7.30 provides the scenario analysis results for lambda values of £100,000 and £1,000,000 per QALY. At £100,000 per QALY, first line use of bDMARDs contribute to the optimal treatment sequence. The best sequence identified is {ADA, RTX, MTX, SSZ, HCQ, TICORA, ABTS, ETN, CTZ, GOL, IFX, TCZ, BSC}. At £1,000,000 per QALY, the results are identical to £100,000. The best sequence identified is {ADA, RTX, MTX, SSZ, HCQ, TICORA, ABTS, ETN, CTZ, GOL, IFX, TCZ, BSC}

It should be noted that the algorithm was not tuned for these much larger estimates of the objective function (NMB). Therefore it is not surprising that the algorithm appears to perform less well when undertaking these scenario analyses with a high lambda. The same solution is identified at £100,000 per QALY and at £1,000,000 per QALY, but the algorithm appears to require more rounds before the best solution is identified, at least compared to the basecase analysis and the scenarios conducted with lower lambda values.

Table 7.29: Scenario analysis - Lambda values £0, £20,000, £50,000

Lambda £0 per QALY (Cost-Minimisation)																	
Round	Sequence (line)											NMB*	Costs*	QALYs*	Iteration when best solution found		
	1	2	3	4	5	6	7	8	9	10	11						
1	BSC											-£22,995	£22,995	4.502	205		
2	BSC											-£22,995	£22,995	4.502	1		
⋮																	
10	BSC											-£22,995	£22,995	4.502	1		
END	BSC											-£22,995	£22,995	4.502	1		
Lambda £20,000 per QALY																	
1	MTX	SSZ	HCQ	BSC								£81,906	£34,586	5.825	35		
2	MTX	SSZ	HCQ	BSC								£81,906	£34,586	5.825	1		
⋮																	
10	MTX	SSZ	HCQ	BSC								£81,906	£34,586	5.825	1		
END	MTX	SSZ	HCQ	BSC								£81,906	£34,586	5.825	1		
Lambda £50,000 per QALY																	
1	SSZ	MTX	HCQ	TICORA	ETN	ADA	RTX	IFX	GOL	BSC			£264,592	£61,950	6.531	87	
2	MTX	SSZ	HCQ	TICORA	CTZ	RTX	ETN	BSC					£266,383	£57,491	6.477	138	
3	MTX	SSZ	HCQ	TICORA	CTZ	RTX	ETN	BSC					£266,383	£57,491	6.477	1	
4	MTX	SSZ	HCQ	TICORA	ABTS	RTX	ADA	CTZ	ETN	BSC			£266,516	£60,074	6.532	162	
5	MTX	SSZ	HCQ	TICORA	ABTS	RTX	ADA	CTZ	ETN	BSC			£266,516	£60,074	6.532	1	
6	MTX	SSZ	HCQ	TICORA	ABTS	RTX	ADA	CTZ	ETN	BSC			£266,516	£60,074	6.532	1	
7	MTX	SSZ	HCQ	TICORA	ABTS	RTX	ADA	CTZ	ETN	BSC			£266,516	£60,074	6.532	1	
8	MTX	SSZ	HCQ	TICORA	ABTS	RTX	ADA	CTZ	ETN	BSC			£266,516	£60,074	6.532	1	
9	MTX	SSZ	HCQ	TICORA	ABTS	RTX	ADA	CTZ	ETN	BSC			£266,516	£60,074	6.532	1	
10	SSZ	HCQ	MTX	TICORA	CTZ	RTX	ETN	ADA	ABTS	GOL	BSC			£266,832	£59,746	6.532	276
END	SSZ	HCQ	MTX	TICORA	CTZ	RTX	ETN	ADA	ABTS	GOL	BSC			£266,832	£59,746	6.532	1

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7000. Previous best solution for every round. Stop when maximum runs completed (temperature < 100) or 50 consecutive solutions rejected. 40,000 simulation model.

Table 7.30: Scenario analysis - Lambda values £100,000, £1,000,000

Lambda £100,000 per QALY																	
Round	Sequence (line)													NMB*	Costs*	QALYs*	Iteration when best solution found
	1	2	3	4	5	6	7	8	9	10	11	12	13				
1	ADA	RTX	SSZ	MTX	HCQ	CTZ	TICORA	ABTS	ETN	GOL	IFX	BSC		£614,106	£112,587	7.267	426
2	ADA	RTX	SSZ	MTX	HCQ	TICORA	ABTS	ETN	CTZ	GOL	IFX	TCZ	BSC	£614,214	£111,885	7.261	227
3	ADA	RTX	SSZ	MTX	HCQ	TICORA	ABTS	ETN	CTZ	GOL	IFX	TCZ	BSC	£614,214	£111,885	7.261	1
4	ADA	RTX	SSZ	MTX	HCQ	TICORA	ABTS	ETN	CTZ	GOL	IFX	TCZ	BSC	£614,214	£111,885	7.261	1
5	ADA	RTX	SSZ	MTX	HCQ	TICORA	ABTS	ETN	CTZ	GOL	IFX	TCZ	BSC	£614,214	£111,885	7.261	1
6	ADA	RTX	SSZ	MTX	HCQ	TICORA	ABTS	ETN	CTZ	GOL	IFX	TCZ	BSC	£614,214	£111,885	7.261	1
7	ADA	RTX	SSZ	MTX	HCQ	TICORA	ABTS	ETN	CTZ	GOL	IFX	TCZ	BSC	£614,214	£111,885	7.261	1
8	ADA	RTX	MTX	SSZ	HCQ	TICORA	ABTS	ETN	CTZ	GOL	IFX	TCZ	BSC	£614,216	£111,883	7.261	3
9	ADA	RTX	MTX	SSZ	HCQ	TICORA	ABTS	ETN	CTZ	GOL	IFX	TCZ	BSC	£614,216	£111,883	7.261	1
10	ADA	RTX	MTX	SSZ	HCQ	TICORA	ABTS	ETN	CTZ	GOL	IFX	TCZ	BSC	£614,216	£111,883	7.261	1
END	ADA	RTX	MTX	SSZ	HCQ	TICORA	ABTS	ETN	CTZ	GOL	IFX	TCZ	BSC	£614,216	£111,883	7.261	1
Lambda £1,000,000 per QALY																	
1	ADA	ETN	TCZ	GOL	CTZ	RTX	MTX	ABTS	TICORA	IFX	SSZ	HCQ	BSC	£7,300,822	£148,251	7.449	321
2	ADA	RTX	ETN	GOL	CTZ	TCZ	IFX	TICORA	MTX	ABTS	SSZ	HCQ	BSC	£7,302,991	£133,829	7.437	185
3	ADA	RTX	ETN	GOL	CTZ	TCZ	IFX	TICORA	MTX	ABTS	SSZ	HCQ	BSC	£7,302,991	£133,829	7.437	1
⋮																	
10	ADA	RTX	MTX	SSZ	HCQ	TICORA	ABTS	ETN	CTZ	GOL	IFX	TCZ	BSC	£614,216	£111,883	7.261	1
END	ADA	RTX	MTX	SSZ	HCQ	TICORA	ABTS	ETN	CTZ	GOL	IFX	TCZ	BSC	£614,216	£111,883	7.261	1

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7000. Previous best solution for every round. Stop when maximum runs completed (temperature < 100) or 50 consecutive solutions rejected. 40,000 simulation model.

Alternative treatment decrement values

A treatment decrement parameter affects the effectiveness of the treatment within a sequence. It is applied to any cDMARD treatment where two prior cDMARDs are used, and also to any bDMARD treatment where two prior bDMARDs have been used. In the basecase analysis, the parameter is set at 10%. Once two prior (cDMARD or bDMARDs) have been used in the treatment sequence, the probability of treatment response is reduced by multiplying the probability of response by 1-10%. By varying the treatment decrement parameter, the downstream efficacy of treatments can be altered, to reflect diminished efficacy when similar treatments have already been attempted (classified by cDMARD or bDMARD status).

Scenario analyses were conducted with the treatment decrement parameter set at 0%, 20%, 30% and 40%, compared to the basecase value of 10%.

This analysis found that the best sequence found identical to that found in the basecase analysis when the treatment decrement parameter was varied in all instances. (Table 7.31).

Table 7.31: Scenario analysis - treatment decrement values

Treatment Decrement = 0%									
Round	Sequence (line)					NMB	Costs*	QALYs*	Iteration when best solution found
	1	2	3	4	5				
1	MTX	SSZ	HCQ	TICORA	BSC	£142,913	£40,314	6.108	120
2	MTX	SSZ	HCQ	TICORA	BSC	£142,913	£40,314	6.108	1
:									
10	MTX	SSZ	HCQ	TICORA	BSC	£142,913	£40,314	6.108	1
END	MTX	SSZ	HCQ	TICORA	BSC	£142,913	£40,314	6.108	1
Treatment Decrement = 20%									
1	MTX	SSZ	HCQ	TICORA	BSC	£139,471	£36,074	5.852	179
2	MTX	SSZ	HCQ	TICORA	BSC	£139,471	£36,074	5.852	1
:									
10	MTX	SSZ	HCQ	TICORA	BSC	£139,471	£36,074	5.852	1
END	MTX	SSZ	HCQ	TICORA	BSC	£139,471	£36,074	5.852	1
Treatment Decrement = 30%									
1	MTX	SSZ	HCQ	TICORA	BSC	£138,588	£34,552	5.771	311
2	MTX	SSZ	HCQ	TICORA	BSC	£138,588	£34,552	5.771	1
:									
10	MTX	SSZ	HCQ	TICORA	BSC	£138,588	£34,552	5.771	1
END	MTX	SSZ	HCQ	TICORA	BSC	£138,588	£34,552	5.771	1
Treatment Decrement = 40%									
1	MTX	SSZ	HCQ	TICORA	BSC	£138,217	£34,390	5.754	171
2	MTX	SSZ	HCQ	TICORA	BSC	£138,217	£34,390	5.754	1
:									
10	MTX	SSZ	HCQ	TICORA	BSC	£138,217	£34,390	5.754	1
END	MTX	SSZ	HCQ	TICORA	BSC	£138,217	£34,390	5.754	1

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7000. Previous best solution for every round. Stop when maximum runs completed (temperature < 100) or 50 consecutive solutions rejected. 40,000 simulation model.

Alternative cooling schedules

The final scenario analysis involved running the SA algorithm using the linear and geometric cooling schedules. These cooling schedules were run using the tuned parameters estimated in the tuning process (see Table 7.20 and Table 7.21). The results are reported in Table 7.32.

Table 7.32: Scenario analysis - alternative cooling schedules

Geometric Cooling: 50 Repetitions per step. Cooling rate 0.80 (Maximum possible runs per round = 1200)										
Round	Sequence (line)					NMB*	Costs*	QALYs*	Iteration when best solution found	Total Iterations
	1	2	3	4	5					
1	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	615	1001
2	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	693
3	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	874
4	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	671
5	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	961
6	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	828
7	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	350
8	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	746
9	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	540
10	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	549
END	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	100
Linear Cooling: 100 Repetitions per step. 500 Decrement per step. (Maximum possible runs per round = 1400)										
1	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	971	1400
2	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	1301
3	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	1400
4	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	1400
5	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	1400
6	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	1400
7	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	1400
8	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	1400
9	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	1400
10	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	1400
END	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006	1	100

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Green shading for 'current best' solution being identified by algorithm. bDMARDs in **bold**. Algorithm settings: Initial temperature 7,000. Previous best solution for every round. Stop when maximum runs completed (temperature < 100) or 10 consecutive solutions rejected. 40,000 simulation model runs

The best found sequence was consistent with that identified in the basecase analysis, irrespective of which cooling schedule was used. However, the basecase cooling schedule (Lundy) was much more efficient. The geometric and linear cooling schedules took 6,312 and 14,001 iterations, respectively, compared with the basecase analysis which took 2,433.

OPTIMAL cDMARD-ONLY TREATMENT SEQUENCE

The patient simulation model was run without the optimisation procedure, enabling a subset of user-specified cDMARD treatment sequences to be run. The total number of sequences which only include cDMARD treatments is 65 (see Table 7.33). All 65 sequences were run in

the simulation model to identify the optimal cDMARD-only treatment sequence, and the results are presented in Table 7.33. MTX, HCQ and SSZ have identical efficacy*, which is why many sequences generate identical total QALY estimates. Unsurprisingly, these results mirrored the basecase analysis, with an optimal sequence of {MTX, SSZ, HCQ, TICORA, BSC}. The top six solutions are within two standard errors of the apparent optimal solution {MTX, SSZ, HCQ, TICORA, BSC}, and are highlighted in green.

The most effective set of sequences {TICORA, cDMARD, BSC} (solutions 43-48 in Table 7.33) were also the most costly. From this analysis, some conclusions can be drawn. Firstly, at a lambda threshold of £30,000 per QALY, an intensive strategy (TICORA) of combination DMARDs may not be cost effective as first line treatment. Also, if treatments are assumed to have equal efficacy and licensed indication (as with MTX, SSZ and HCQ) then it is logical that treatments are prescribed in ascending order of treatment cost.

* They are identical in terms of initial treatment response, length of time spent on treatment, and their impact on HAQ progression over time.

Table 7.33: Optimisation of cDMARD treatment sequence

Sequence (ranked by NMB)	Line					NMB*	Costs*	QALYs*
	1	2	3	4	5			
1	MTX	SSZ	HCQ	TICORA	BSC	£141,824	£38,361	6.006
2	SSZ	MTX	HCQ	TICORA	BSC	£141,816	£38,368	6.006
3	MTX	HCQ	SSZ	TICORA	BSC	£141,752	£38,433	6.006
4	SSZ	HCQ	MTX	TICORA	BSC	£141,740	£38,445	6.006
5	HCQ	MTX	SSZ	TICORA	BSC	£141,609	£38,575	6.006
6	HCQ	SSZ	MTX	TICORA	BSC	£141,605	£38,579	6.006
7	MTX	SSZ	TICORA	HCQ	BSC	£140,342	£39,604	5.998
8	SSZ	MTX	TICORA	HCQ	BSC	£140,334	£39,612	5.998
9	MTX	HCQ	TICORA	SSZ	BSC	£140,214	£39,732	5.998
10	SSZ	HCQ	TICORA	MTX	BSC	£140,199	£39,747	5.998
11	MTX	SSZ	HCQ	BSC		£140,152	£34,586	5.825
12	SSZ	MTX	HCQ	BSC		£140,145	£34,594	5.825
13	MTX	HCQ	SSZ	BSC		£140,080	£34,658	5.825
14	HCQ	MTX	TICORA	SSZ	BSC	£140,070	£39,876	5.998
15	SSZ	HCQ	MTX	BSC		£140,068	£34,670	5.825
16	HCQ	SSZ	TICORA	MTX	BSC	£140,062	£39,884	5.998
17	HCQ	MTX	SSZ	BSC		£139,937	£34,801	5.825
18	HCQ	SSZ	MTX	BSC		£139,933	£34,805	5.825
19	MTX	TICORA	SSZ	HCQ	BSC	£138,136	£43,095	6.041
20	SSZ	TICORA	MTX	HCQ	BSC	£138,123	£43,108	6.041
21	MTX	TICORA	HCQ	SSZ	BSC	£138,095	£43,136	6.041
22	SSZ	TICORA	HCQ	MTX	BSC	£138,079	£43,152	6.041
23	HCQ	TICORA	MTX	SSZ	BSC	£137,855	£43,376	6.041
24	HCQ	TICORA	SSZ	MTX	BSC	£137,852	£43,379	6.041
25	MTX	SSZ	TICORA	BSC		£137,672	£38,370	5.868
26	SSZ	MTX	TICORA	BSC		£137,664	£38,377	5.868
27	MTX	HCQ	TICORA	BSC		£137,387	£38,655	5.868
28	SSZ	HCQ	TICORA	BSC		£137,360	£38,681	5.868
29	HCQ	MTX	TICORA	BSC		£137,243	£38,798	5.868
30	HCQ	SSZ	TICORA	BSC		£137,224	£38,817	5.868
31	MTX	SSZ	BSC			£136,339	£32,686	5.634
32	SSZ	MTX	BSC			£136,332	£32,694	5.634
33	MTX	HCQ	BSC			£136,053	£32,972	5.634
34	SSZ	HCQ	BSC			£136,027	£32,999	5.634
35	HCQ	MTX	BSC			£135,911	£33,114	5.634
36	HCQ	SSZ	BSC			£135,892	£33,133	5.634
37	MTX	TICORA	SSZ	BSC		£135,646	£41,894	5.918
38	SSZ	TICORA	MTX	BSC		£135,633	£41,907	5.918
39	MTX	TICORA	HCQ	BSC		£135,453	£42,086	5.918
40	SSZ	TICORA	HCQ	BSC		£135,427	£42,113	5.918
41	HCQ	TICORA	MTX	BSC		£135,213	£42,326	5.918
42	HCQ	TICORA	SSZ	BSC		£135,200	£42,340	5.918
43	TICORA	MTX	SSZ	HCQ	BSC	£134,294	£47,885	6.073
44	TICORA	SSZ	MTX	HCQ	BSC	£134,290	£47,889	6.073
45	TICORA	MTX	HCQ	SSZ	BSC	£134,252	£47,927	6.073
46	TICORA	SSZ	HCQ	MTX	BSC	£134,245	£47,934	6.073
47	TICORA	HCQ	MTX	SSZ	BSC	£134,177	£48,002	6.073
48	TICORA	HCQ	SSZ	MTX	BSC	£134,175	£48,005	6.073
49	MTX	TICORA	BSC			£132,292	£40,465	5.759
50	SSZ	TICORA	BSC			£132,266	£40,491	5.759
51	HCQ	TICORA	BSC			£131,845	£40,912	5.759
52	TICORA	MTX	SSZ	BSC		£131,437	£46,703	5.938
53	TICORA	SSZ	MTX	BSC		£131,432	£46,707	5.938
54	TICORA	MTX	HCQ	BSC		£131,232	£46,907	5.938
55	TICORA	SSZ	HCQ	BSC		£131,214	£46,926	5.938
56	TICORA	HCQ	MTX	BSC		£131,157	£46,982	5.938
57	TICORA	HCQ	SSZ	BSC		£131,143	£46,996	5.938
58	MTX	BSC				£127,929	£29,661	5.253
59	SSZ	BSC				£127,902	£29,687	5.253
60	TICORA	MTX	BSC			£127,678	£45,125	5.760
61	TICORA	SSZ	BSC			£127,660	£45,143	5.760
62	HCQ	BSC				£127,483	£30,107	5.253
63	TICORA	HCQ	BSC			£127,385	£45,418	5.760
64	TICORA	BSC				£121,357	£42,091	5.448
65	BSC					£112,062	£22,995	4.502

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY.

Highlighted green if within 2 standard errors (£484) of best identified sequence

COMPARISON TO NICE MTA APPRAISAL OF RA BIOLOGICS

The simulation model was run using the treatment sequences evaluated in the NICE MTA appraisal of RA bDMARDs (Severe RA patients, Population 1 and 2). 12 sequences were evaluated; one cDMARD comparator sequence {MTX, TICORA, BSC}; eight third-line bDMARD sequences {MTX, TICORA, bDMARD 1, RTX 2, bDMARD2*, BSC}, and three first line bDMARD sequences {bDMARD1, RTX, TCZ, MTX, TICORA, BSC}.

The full results are presented in Table 7.34. At a lambda threshold of £30,000 per QALY gained, the optimal sequence was the cDMARD comparator sequence. bDMARD sequences were not cost-effective, at both first line and third line use. The cDMARD comparator sequence was both the least costly and least effective. The bDMARD sequences provided additional benefits in terms of QALYs, but at too high a cost.

Table 7.34: Comparison to NICE MTA appraisal of RA biologics

Sequence	Line							NMB*	Costs*	QALYs*	ICER
	1	2	3	4	5	6	7				
<i>cDMARD comparator sequence</i>											
1	TICORA	MTX	BSC					£127,678	£45,125	5.760	-
<i>Third line bDMARD sequences</i>											
2	TICORA	ABT	RTX	TCZ	MTX	BSC		£105,055	£99,095	6.805	£51,646 ¹
3	TICORA	ABTS	RTX	TCZ	MTX	BSC		£105,658	£99,782	6.848	£50,236 ¹
4	TICORA	ADA	RTX	TCZ	MTX	BSC		£105,086	£99,874	6.832	£51,072 ¹
5	TICORA	CTZ	RTX	TCZ	MTX	BSC		£107,156	£96,994	6.805	£49,635 ¹
6	TICORA	ETN	RTX	TCZ	MTX	BSC		£104,527	£100,733	6.842	£51,394 ¹
7	TICORA	GOL	RTX	TCZ	MTX	BSC		£104,647	£99,773	6.814	£51,848 ¹
8	TICORA	IFX	RTX	TCZ	MTX	BSC		£99,160	£100,550	6.657	£61,789 ¹
9	TICORA	TCZ	RTX	MTX	BSC			£119,853	£95,277	7.171	£55,911 ¹
<i>Current NICE recommended treatment sequence</i>											
10	MTX	TICORA	ADA	RTX	TCZ	SSZ	BSC	£120,777	£77,043	6.594	-
<i>First line bDMARD sequence</i>											
11	ADA	RTX	TCZ	MTX	BSC			£93,074	£120,136	7.107	£84,002 ²

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY.

¹Compared to sequence 1

²Compared to sequence 10

The use of bDMARDs in third-line had an ICER of between £49,635 and £61,789 per QALY gained compared to cDMARDs only. The range was dependent on the particular bDMARD therapy used. These results are slightly lower than the original results estimated in the NICE RA MTA assessment group analysis. The basecase ICER using the original model was £56,000 per QALY gained for third line bDMARD therapy compared to cDMARDs only.

The use of bDMARDs in first line had an ICER of £84,002 per QALY gained compared a treatment sequence reflecting NICE guidance. These results are slightly lower than the original

* Where relevant

results estimated in the NICE RA MTA assessment group analysis. The range was dependent on the particular bDMARD therapy used. The basecase ICER using the original model was £98,000 per QALY gained for first line bDMARD therapy compared a treatment sequence reflecting NICE guidance.

Unfortunately the full NICE MTA results, in terms of absolute costs and QALYs, cannot be reported due to being Commercial in Confidence.

The deviation between the two analyses could come from several different sources:

1. The modifications undertaken (see page 155) within the Simul8 model may have significantly altered the results. Using metamodels and incorporating a treatment decrement parameter will undoubtedly alter the results slightly.
2. An error in either version of the model, which given the size and complexity of the models does remain a possibility
3. Slight different sequences of treatments being modelled – the SOSA analysis model was amended slightly to match the NICE MTA sequences evaluated
4. Slightly different patient populations – this analysis used the treatment naïve (Population 1) data for the patient population, and then models cDMARD use until bDMARDs (Population 2/3 in the MTA), and these patients may be different to those sampled in Population 2/3 used in the MTA.

Both models were fully validated, and the NICE MTA model underwent a full external peer-review. Therefore it is important to highlight that the software package used, and therefore the methods used to develop the model and simulate lifetime costs and QALYs may have an impact on the validity of the final results. It should also be noted that the objective of this analysis was not to replicate the NICE MTA model, but instead to provide a model which would enable the implementation of the simulation optimisation method, and therefore the deviation between the two sets of results does not detract from meeting this objective.

COMPARISON TO EXISTING NICE GUIDANCE

The existing NICE clinical pathway for patients with severe RA was detailed in Figure 5.1. After two cDMARDs, patients can receive a bDMARD. After this bDMARD, patients can receive RTX, and then TCZ. If a patient is contraindicated to RTX, then an alternative bDMARD can be used.

The model was run with the 21 sequences possible from this current NICE guidance pathway. It was assumed, for simplicity, that people are eligible for RTX treatment. The first line treatment was either TICORA, or two sequential cDMARD therapies before bDMARDs. If the first line treatment was TICORA, then Sequences 8-14 assume that cDMARDs are not used after

bDMARD treatment, and Sequences 15-21 assume that the remaining cDMARDs are used after bDMARD treatment.

The best sequence identified from the basecase simulation optimisation {MTX, SSZ, HCQ, TICORA, BSC} was run as a comparator (Sequence 22). This was undertaken to see if the new sequence identified was superior, and whether it was more efficacious than current NICE guidance.

Table 7.35 presents the results from this analysis. The best sequence in terms of maximising NMB remains the sequences found in the basecase simulation optimisation {MTX, SSZ, HCQ, TICORA, BSC}.

However, this sequence is both less costly and less efficacious when compared with the NICE guidance pathway. In these sequences, bDMARDs are used from third line position in the sequence. The addition of bDMARDs increases the lifetime QALYs provided by the treatment sequence. These sequences are also much more costly. Using {MTX, SSZ} as the first line treatment in the sequence is superior compared to using TICORA. This reinforces earlier findings (Table 7.28).

Table 7.35: NICE guidance comparison

Sequences	Line								NMB*	Costs*	QALYs*
	1	2	3	4	5	6	7	8			
<i>“MTX, SSZ” first line NICE sequence</i>											
1	MTX	SSZ	ADA	RTX	TCZ	TICORA	HCQ	BSC	£123,629	£75,467	6.637
2	MTX	SSZ	ETN	RTX	TCZ	TICORA	HCQ	BSC	£123,182	£76,521	6.657
3	MTX	SSZ	IFX	RTX	TCZ	TICORA	HCQ	BSC	£122,018	£76,652	6.622
4	MTX	SSZ	CTZ	RTX	TCZ	TICORA	HCQ	BSC	£125,364	£73,346	6.624
5	MTX	SSZ	GOL	RTX	TCZ	TICORA	HCQ	BSC	£123,799	£75,246	6.635
6	MTX	SSZ	ABT	RTX	TCZ	TICORA	HCQ	BSC	£123,105	£75,562	6.622
7	MTX	SSZ	ABTS	RTX	TCZ	TICORA	HCQ	BSC	£123,446	£75,648	6.636
<i>TICORA first line NICE sequence(no post-biologics cDMARDs)</i>											
8	TICORA	ADA	RTX	TCZ	BSC				£104,275	£99,266	6.786
9	TICORA	ETN	RTX	TCZ	BSC				£102,993	£100,690	6.789
10	TICORA	IFX	RTX	TCZ	BSC				£102,826	£100,085	6.764
11	TICORA	CTZ	RTX	TCZ	BSC				£105,439	£96,768	6.740
12	TICORA	GOL	RTX	TCZ	BSC				£104,999	£98,838	6.795
13	TICORA	ABT	RTX	TCZ	BSC				£103,996	£98,761	6.759
14	TICORA	ABTS	RTX	TCZ	BSC				£104,144	£99,210	6.778
<i>TICORA first line NICE sequence(including post-biologics cDMARDs)</i>											
15	TICORA	ADA	RTX	TCZ	MTX	SSZ	HCQ	BSC	£105,897	£100,240	6.871
16	TICORA	ETN	RTX	TCZ	MTX	SSZ	HCQ	BSC	£105,301	£101,415	6.891
17	TICORA	IFX	RTX	TCZ	MTX	SSZ	HCQ	BSC	£104,949	£101,039	6.866
18	TICORA	CTZ	RTX	TCZ	MTX	SSZ	HCQ	BSC	£107,269	£97,717	6.833
19	TICORA	GOL	RTX	TCZ	MTX	SSZ	HCQ	BSC	£106,004	£100,032	6.868
20	TICORA	ABT	RTX	TCZ	MTX	SSZ	HCQ	BSC	£105,661	£99,880	6.851
21	TICORA	ABTS	RTX	TCZ	MTX	SSZ	HCQ	BSC	£106,022	£100,297	6.877
<i>Best sequence from simulation optimisation</i>											
22	MTX	SSZ	HCQ	TICORA	BSC				£141,824	£38,361	6.006

*Costs (£) and QALYs discounted at 3.5%. NMB valued at £30,000 per QALY. Highlighted green if maximum, red if minimum.

7.5 DISCUSSION AND CONCLUSIONS

The aim of this chapter was to undertake a simulation optimisation of a discrete event simulation model, to enable the identification of an optimal or near-optimal DMARD treatment sequence for patients with RA.

SA with a memory function was selected as the simulation optimisation method to be implemented. Modifications to the DES model were required to enable the simulation optimisation algorithm to run. The SA algorithm required tuning of control parameters to enable it to function properly. The best solution found (across all simulations, as well as in the basecase analysis) was a cDMARD-exclusive sequence of {MTX, SSZ, HCQ, BSC}. For the basecase analysis, the algorithm was relatively quick to complete, requiring about 15hrs to run, and evaluating over 2,400 potential solutions.

The best solution found was robust when the treatment decrement parameter was varied. The best solution found was also robust when a cost effectiveness threshold of £20,000 per QALY was assumed. At a threshold of £50,000 per QALY, the best sequence found included bDMARDs after four previous cDMARD treatments. First line bDMARD use was found to be potentially optimal when the threshold was increased to £100,000 per QALY.

Compared to the NICE RA MTA model developed in Excel, the results using this model were found to be slightly different. This will be discussed further in the following paragraphs. Finally, the best sequence found was compared to the existing set of sequences recommended in the NICE guidance for RA. It was found that the best sequence identified {MTX, SSZ, HCQ, BSC} was optimal compared to NICE guidance, in that NMB would be increased. However, this increase in NMB was at the cost of a reduction in total QALYs.

There are several limitations which require further consideration.

Firstly, the Simul8 model is slightly different to the original model developed in Excel. Modifications were required to enable it to be developed in Simul8. The significant speedup offered by Simul8 came at a potential cost because the econometric models used to provide patient level parameters could not be properly incorporated. This represents a limitation of Simul8, but it also highlights the general trade-off between precision and speed, when developing a complex computer model.

The difference in the results is highlighted in Table 7.34. These differences may come from the modifications made to the Simul8 model, as discussed. However, it could be that an error remains within the analysis, either in the Simul8 model for this thesis or the Excel model for

the RA MTA appraisal. Both models were validated but there always remains a risk of an unidentified error within a model.

Secondly, the patient level variability within the model was such that a definitive statement about superiority between two very similar sequence was not possible. To account for this, the standard error of the best found sequence was used to illustrate which nearby solutions may be within a level of tolerance (twice an estimated standard error). In all cases, the results found to be within this margin of tolerance were very similar to that found to be the best. However, as reviewed in Chapter 6, there is a large body of methodological literature that looks to inform how to compare and rank the output of noisy simulation models. Due to limited time, it was not possible to investigate how these methods could be applied to the treatment sequencing problem, however this remains an area for further research.

The SA algorithm required tuning to enable good performance. Trial and error was used across some control parameters, and simple factorial design across others, to enable experimentation of different control parameters. Ideally, a full factorial set of experiments would have been conducted to obtain the optimal setting for the algorithm, however this was not feasible. In reality, designing and tuning an optimisation algorithm such as a metaheuristic itself represents an optimisation problem. Hyperheuristic methods (see Section 6.8) enable the automation of this tuning process; however they require a very efficient simulation model for their use in a simulation optimisation problem. Also there remains the possibility of algorithm parameters which would enable the algorithm to perform better. Repeated tuning experiments was undertaken to build confidence that it was a change in a parameter which affected performance, rather than chance. However, it does remain possible that the final set of tuned parameters were found by chance, and that superior parameter settings exist.

The algorithm appears to have performed well, given its simplicity compared to much more advanced SA implementations. Relatively simple static cooling schedules were evaluated and the SA algorithm uses a simple restart procedure to provide the algorithm with the best chance of escaping a local optima in the early rounds. The memory cache of previously evaluated solutions enabled a significant speed-up and should be considered in future. While the cooling schedules and SA algorithm performed well, more advanced schedules, such as dynamic schedules, and more advanced algorithms, including reheating procedures and adaptive neighbourhood functions could have improved performance even more.

While the algorithm performed well in the basecase, when the lambda threshold was increased to a point where there was a large change in the NMB objective values being evaluated, then the algorithm performed less well. This was likely to occur because the

algorithm was not tuned for objective values of this particular magnitude. This is an important implication for health economic evaluation. Using a NMB framework allows the problem to be characterised as a simple maximisation problem, but the NMB for any solution is dependent on the lambda value used to monetise benefits. Therefore further tuning may be required if a set of lambda values are to be fully evaluated.

The SO procedure utilises a link between Excel (the optimisation process) and Simul8 (the simulation model) to enable neighbouring sequences to be selected (implemented in Excel VBA) and then passed to Simul8 via Microsoft COM. This is potentially an area of inefficiency for the SO procedure. Also, the way the link was implemented meant that the number of simulation runs had to be predefined. That is, every run of the SO algorithm had a predefined number of patient simulations in the Simul8 model. Therefore it was not possible to attempt ordinal optimisation, or a statistical method which reduces the simulations required for solutions which are clearly inferior. This is a limitation of the approach taken, mainly due to current computing expertise, and also the fact that health economic models are normally developed in a package which enables a level of user interface. Executable models written in a true programming language may have offered significant speed up, as well as the possibility to test alternative SO methods, but it was not possible in this case.

Finally, it was not possible within these timescales to implement an alternative metaheuristic method. For example, it would have been valuable to implement GA and compare its performance to SA. In the review in Chapter 6, only one study compared GA and SA in a combinatorial simulation optimisation problem. It would have added to the body of evidence to implement another metaheuristic and compare their performance. GA's are harder to implement, and require more tuning due to having a greater number of user control parameters. The limited evidence from Lacksonen et al. (2001) is that GA's offer greater performance but at the expense of longer running time.²²⁴ Therefore it was decided to try to implement one method but with as full an experimentation and tuning as possible, rather than implement two methods but with less time for tuning and experimentation. The comparative evidence would have been less robust, compared to the evidence found in this chapter which broadly supports SA and its use for SO. SA has been found to be efficient at reaching what appears to be a global optima of the problem. The method was relatively straightforward to implement, although the modifications required to rebuild the model in Simul8 may have made the final results less robust. The SA algorithm also requires more tuning to ensure that the results found when lambda is varied.

(Blank page)

CHAPTER 8: DISCUSSIONS, RECOMMENDATIONS AND CONCLUSIONS

8.1 CHAPTER OVERVIEW

This chapter provides a discussion of the research presented within this thesis. It highlights the contribution of this work within the context of other related research. It also provides recommendations for further research, before drawing conclusions.

Section 8.2 presents the contribution of this work in the context of other research. Section 8.3 considers the strengths and limitations of this research. Section 8.4 presents recommendations for further research. Finally, Section 8.5 draws conclusions about the overall value and impact of this research.

8.2 CONTRIBUTION OF THIS WORK IN THE CONTEXT OF OTHER RESEARCH

Since beginning this PhD at the start of 2012, there have been several key publications and further advances in research related to this area.

Firstly, Tappenden *et al.* published a Whole Disease Modelling (WDM) methodological framework and an application of this framework.^{292,303} This WDM framework is a system-level approach to health economic modelling which captures the whole system of disease and treatment pathway within one consistent mathematical infrastructure. It allows the incorporation of multiple decision points, and the full quantification in terms of costs and QALYs of downstream consequences. An innovative aspect of the WDM framework is that it allows multiple decision points to be evaluated. Within the context of colorectal cancer, Tappenden *et al.* were able to evaluate screening, surgery, and metastatic treatment across the full pathway of cancer diagnosis and treatment.²⁹²

The DES model developed for implementing simulation optimisation for the RA treatment sequencing problem is consistent with the principles of the WDM framework. A whole disease model is required for a full treatment sequence to be evaluated. The full consequences of alternative treatment sequences have been captured, and alternative options at every decision point (e.g. first line, second line, third line therapy) have been formally evaluated in the RA simulation optimisation analysis.

While the framework was not formally applied when developing either the NICE RA model or the Simul8 model, our research adds value to the WDM research by highlighting the importance of a consistent model which allows the evaluation of all possible decision points in a patient's pathway. Piecemeal models and partial economic evaluations are likely to lead to sub-optimal decisions if downstream consequences are omitted.

The application of a simulation optimisation extends the WDM framework. Rather than a process of identifying key decision questions and using a whole disease model to evaluate each decision question, all possible questions can be considered as an optimisation problem – how do we optimise the complete treatment pathway for a particular population? Although our context has been treatment sequencing, it would not be unfeasible to incorporate alternative policy decisions about diagnosis and treatment, such as optimising screening intervals and population selection, and optimising treatment switching rules.

Secondly, there is an ongoing PhD on a similar topic by a post graduate student at SchARR. Their research is specific to hypertension treatment, and applies metaheuristic methods to a health economic model. There has not been dialogue between the two PhD candidates while research has been ongoing. The PhD topics were developed in very separate contexts and had separate supervisory teams. However, it will be of great interest to consider the final outputs of both PhDs to see where conclusions are similar and where differences are found.

Thirdly, a relevant paper was identified during the latter stages of this thesis. Brailsford *et al.* (2006) report the application an ant colony optimisation (ACO) model to identify optimal screening policies for diabetic retinopathy.³⁰⁴ It is not clear why this paper was not identified during the systematic review of simulation optimisation methods in Chapter 6. As such, this highlights a limitation of the search process used in the review.

ACO is a population-based stochastic optimisation method. When a combinatorial problem can be demonstrated through a graph, such as the travelling salesman problem, then ACO can be applied to find the optimal route or path through the graph.

Brailsford *et al.* used a previously published simulation based ACO method (S-ACO) for their analysis.^{305,306} The method used by Brailsford *et al.* is particularly interesting due to the dynamic process utilised. The development of a solution for evaluation (a particular screening programme) is guided by the simulation and ACO process in one model. This is different to an optimisation algorithm generating a solution, and then passing that solution to a simulation model for evaluation. The authors report that the method worked well and their results have

face validity. In particular, they maximise cost-effectiveness (identify the strategy with the lowest cost per year of sight saved), and also report the most effective screening strategy.

It is a weakness of this thesis that this research was not identified earlier and considered in the systematic review. ACO appears to be a promising method, and further research within a health economic evaluation context would be of significant value.

Finally, in 2013 NICE updated their Guide to the Methods of Technology Appraisal.⁴⁷ I was an invited expert to provide a briefing paper and attend a working party meeting during this process (before this PhD began).^{*} NICE had found that treatment sequences had caused issues for the development of guidance. My briefing paper regarding treatment sequences and downstream costs was considered by the working party. It highlighted that sequences should be considered, explores the key issues when modelling sequences, and identifies the key primary and sensitivity analyses which should be reported. This particular topic, as well as others, was influential in a modification to the Guide to the Methods of Technology Appraisal, which clarified that at the scoping stage, potentially relevant comparators should not be eliminated. Treatment sequences continue to represent a challenge for decision-makers and health economists. It is hoped that this thesis has provided a significant insight into the problem, and provides a potential method which can be applied.

8.3 STRENGTHS AND LIMITATIONS OF THIS RESEARCH

The rationale for undertaking this research is grounded by a systematic review (Chapter 3) which identifies how current methods have led to differing estimates of cost-effectiveness, and potentially sub-optimal decisions, with respect to RA sequences of treatments.

The systematic review reported in Chapter 4 is comprehensive. All major databases were searched with no date limit applied and across all possible disease modifying treatments. 57 studies and fully appraising them using the validated Drummond checklist.⁷⁸ Systematic reviews are regarded as the most robust of all study design types, within evidence based medicine.[†] Although this particular review concerned health economic evaluations, rather than randomised controlled trials, it is still a transparent and robust synthesis of the totality of health economic evidence pertaining to RA treatment.

^{*} [http://www.nicedsu.org.uk/NICE-Methods-Guide-updates\(1985333\).htm](http://www.nicedsu.org.uk/NICE-Methods-Guide-updates(1985333).htm) – Accessed June 2015

[†] <http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/> - Accessed June 2015

The systematic review was validated by cross-checking other reviews of RA health economic literature, including a systematic review also concerned with health economic evaluations for RA treatment sequences.¹²⁰ The conclusions found across these similar reviews were consistent, adding to its validity. The chapter was also peer-reviewed and published in an international RA journal.⁹⁴

The research in this thesis is further underpinned by another systematic review (Chapter 6). This pearl growing systematic review uses innovative searching methods to identify methods of potential applicability to the treatment sequencing problem. As before, a systematic review is robust due to its validated and transparent methodology. In this case, the search and review was conducted across a range of academic disciplines. The applicability of each identified method was judged using a bespoke framework. This was reported with full transparency (extraction tables provided in Appendix B.4), to avoid bias.

The breadth of search may have meant that some studies of relevance were missed. This is known to be true for the Brailsford *et al.* diabetic retinopathy study, and related research using Ant Colony Optimisation methods.³⁰⁴ It may have been beneficial to involve experts within the area of simulation optimisation to ensure that relevant methods and studies had been identified; however this was not feasible within the time available.

The review sought to identify relevant methods, irrespective of the academic field they were developed in or applied. This included fields such as engineering, computing and mathematics. As such, there may be bias and errors in the interpretation of particular aspects of the method or optimisation problem. To minimise the risk of this, intensive training was undertaken alongside this PhD research. Courses offered by the OR Society and the National Taught Course Centre in Operational Research (NATCOR) were attended. Courses focussed on stochastic modelling, metaheuristics, optimisation and simulation modelling. Skills were developed to ensure that alternative methods could be assessed and applied to the RA treatment sequencing problem.

The review of methods was specific to combinatorial problems and focussed on simulation optimisation methods. Therefore non-combinatorial problems were not included, although potentially they could work for a treatment sequencing problem. Also, statistical methods of sample approximation, and metamodeling/emulation were not included within this review. These could provide a solution to the treatment sequencing problem, but a decision was made to focus on metaheuristic methods due to the significant amount of current research that has been applied to combinatorial simulation optimisation problems.

Both systematic reviews would have benefitted from a second reviewer to validate and cross-check data extraction and data appraisal. This approach is standard practice in regular health research and systematic reviewing, but not feasible for a PhD project.

The timing of this PhD was ideal, because it enabled involvement with the NICE re-appraisal of biologics for RA. This enabled a model to be developed for the NICE appraisal and a separate model based on this to be developed for the PhD. The NICE MTA model was developed as a collaborative effort, involving health economists, statisticians, systematic reviews and modellers. This meant that the final model used the most appropriate evidence, and avoided simplifications which would have been inevitable if conducted by just one person. The model development process and team were carefully constructed so that I retained overall responsibility for the final model. However, much of the evidence synthesis used within the model is the result of other people's work. This reflects the reality of health economic model development, which is almost always a collaborative process. PhD's are normally the endeavour of one person and sometimes this can act as a constraint. However, good timing and an accommodating team within SCHARR enabled my PhD to benefit from a much more comprehensive model, and I was able to engage with the NICE appraisal process and further understand the NICE process (scoping, appraisal, guidance development).

The research would have benefitted from the experience of using the simulation optimisation methodology within a real world decision-making context. This was not possible during this PhD. However, with the SOSA method providing positive results, it is hoped to engage with upcoming NICE appraisals, in RA and in other chronic conditions, where flexible simulation modelling methods and optimisation methods can potentially be applied.

A weakness is that a comparative evaluation of simulated annealing and an alternative optimisation method was not possible. It would have been ideal to implement alternative methods (for example, a genetic algorithm) and explore whether one particular method is superior. However, it was not feasible to undertake this within the PhD. Programming, tuning and evaluating one simulation optimisation method was a substantial undertaking, even though simulated annealing is often seen as a relatively straight forward algorithm to implement (compared with population based and adaptive metaheuristics methods). Also, the value of a comparison of two or more methods is limited due to their performance often being problem-specific. Finding that one method is superior to another in one problem context may not be useful if the converse is true in an alternative problem context. Therefore further studies of alternative methods within alternative problem contexts are of value. The evidence found within the systematic review of methods (Chapter 5) did not report a clear finding as to

which method is potentially best for a combinatorial simulation optimisation problem. Therefore this unanswered question is wider than just health economic evaluation, but a question for simulation optimisation as a discipline.

The SOSA methodology would have been improved if it was possible to incorporate the precision of the estimated objective value into the algorithm's decision to run a set of patients. For example, some SO methods allow the decision about whether to run more simulations to be determined by the expected estimate of the objective function and its variance. Therefore solutions which are clearly inferior are only run for a small number of patient simulations, but solutions which are very close, or potentially an improving solution are run for a large number of patient simulations. This could have improved the accuracy of the results, due to simulation time being focussed on good solutions and improving the precision of the estimate of performance, however it was not feasible due to the software used. Simul8 has limits on the dimensions of its data and it is unable to easily and efficiently move large data in and out of Excel. Also, the number of patients run had to be determined *a priori* using the COM interface, and was therefore the same for every evaluation. Using an alternative software package, or a programming language, would have got around this limitation, but this was not feasible.

SOSA is limited by the software package or language it is implemented in. If using a bespoke simulation package such as Simul8 or Arena, then the complexity of the optimisation algorithm is limited to the bespoke programming language in each package (Visual Logic for Simul8, and a limited Visual Basic interface for Arena). Simul8 was chosen over Arena due to user familiarity, as well as being routinely used in our academic department; however it may be that it is feasible to program the simulation model and optimisation process within an Arena model. Further investigation and dialogue with OR simulation community may resolve this.

The alternative to using a bespoke simulation package is to develop the simulation model and optimisation process within a particular programming language (C++, Java, Fortran, R etc.), the benefit may be increased performance, but the downside is that often programmed models are less visual for decision making purposes, which are significant limitations when considering the NICE appraisal process and the fact that the stakeholders and independent assessment groups must be able to externally validate and critique the model.

Simul8 and Arena both have an optimisation software package associated with them (OptQuest), however this package is bespoke, black box and requires the full Professional version. It is not clear exactly what algorithm and process is being applied, although the documentation reports that composite search algorithm is used combining 'tabu search,

scatter search, integer programming and neural networks'.* Using a black box bespoke optimisation software package is less suitable for academic purposes, although it may well deliver good performance in practice.

A final strength of this research is that the generalisability and robustness of SA continues to be reinforced. The algorithm remains relatively simple and easy to implement, but has good performance across a range of simulation and non-simulation optimisation problems. Applying it to a treatment sequencing problem is yet another unique problem instance where it has performed well. There remain several other possibilities for optimisation within a health economic evaluation context and SA would appear to be a potential solution.

8.4 RECOMMENDATIONS FOR FURTHER RESEARCH

This thesis represents a single attempt to apply SOSA to the RA treatment sequencing problem. Therefore remains several unresolved methodological issues regarding SOSA which require further research.

FURTHER EVALUATION

Repeated evaluation would generate greater evidence regarding the appropriateness and feasibility of using simulation optimisation for DES models in a health economic evaluation context. Within simulated annealing, there are alternative neighbourhood functions, cooling schedules and algorithm methodologies which could provide better results and greater efficiency. The algorithm applied is relatively simplistic, and a significant improvement in using SA for the RA treatment sequencing problem could be made by applying more cutting edge SA methods.

Outside of SA, there remain other metaheuristics which may offer greater performance. GAs have been shown to perform well for combinatorial SO problems, but there remain questions regarding their efficiency. Brailsford *et al.* found that ACO performed well for their diabetic retinopathy screening optimisation problem. Further evaluation with these alternative methodologies is required.

FURTHER GENERALISABILITY

This research focussed on an application of SOSA for a treatment sequencing problem in RA. There are other clinical conditions where treatment sequencing issues remain, and where sub-optimal sequences of treatments may be used, both from a clinical and health economic viewpoint. These conditions include depression, hypertension, epilepsy and other chronic conditions. Also, wider than just clinical areas, the new advances in personalised medicine and

* <http://goo.gl/OB8wM9> - OptQuest documentation - Accessed June 2015

gene therapy open up possibilities for SOSA to be applied to individualised patient treatment pathways, including adaptive treatment decisions which involve patient characteristics and rapid simulation of future costs and benefits.

If SOSA is used in other clinical areas and problem contexts, then further evaluation should be undertaken. Qualitative research involving health economists could provide further evidence regarding the feasibility of SOSA within the particular decision context within which it is being applied. Also, more broadly, qualitative research involving decision-makers would help understand more fully the contexts in which SOSA may and may not be helpful.

PILOT STUDY WITH A DECISION-MAKING CONTEXT

The application of SOSA within a real-life decision-making context would be very valuable. The current window of time afforded to model development and evaluation within a NICE MTA is already very tight. It is unlikely that SOSA could be undertaken within this current window. The implementation, tuning and evaluation took the best part of 1 year to complete, and that was with an existing model structure to use, albeit in a separate platform. There may be ways to significantly reduce the time required to undertake SO. Computer code was written from scratch for this analysis; however components could be reused in the future (see Code in Appendix D). If further evaluations of alternative SO methods are undertaken then less tuning may be required as confidence in the performance of alternative metaheuristics grows. Efficiently coded DES models could theoretically be used with a hyperheuristic, which would remove a great deal of bespoke coding and algorithm development. Simulation methods continue to advance, along with computing power, which will enable more complex models to be evaluated more quickly.

The SO process could be brought forward in the NICE decision-making context, and instead of being used for optimisation, it could be used as a method to 'seek' decision problems and contexts which may or may not be most important for evaluation. For example, the results show from our evaluation that first line bDMARD use is not cost-effective. This could have been identified at the scoping stage and not taken forward for full evaluation. An existing model would be required, but it would enable NICE to avoid evaluating decision problems with a high likelihood of an intervention not being cost-effective. Where quite different sequences are identified to have very similar NMB, then a much more robust evaluation and comparison between those two (or more) sequences could be conducted.*

* for example, a non-bDMARD sequence has similar estimated NMB to a bDMARD sequence

Instead the results can be used to identify the iNMB between 'good sequences' and highlight where further evaluations (including many more patient simulations) should focus to gain more confidence in the optimality of a particular sequence.

Making a definitive decision to recommend one treatment sequence over another when the iNMB between them is very small is unlikely to happen. Monte Carlo error leads to a risk that a wrong decision is made in reality, and even if the right decision is made, the net gain is very small. There may be numerous reasons for why a decision maker may decide to avoid making a distinction between closely performing alternative sequences. These include: pharmaceutical competition; patient choice; clinical choice; benefits not captured within the QALY; and the continued generation of clinical evidence. Also, it is likely to suggest overconfidence in the DES model that underpins the analysis, and a metaheuristic cannot guarantee the optimal solution, just a near-optimal solution with varying degrees of confidence.

8.5 CONCLUSIONS

The aim of this research was to test the feasibility of simulation optimisation methods to find an optimal or near-optimal sequence of disease modifying treatments for RA in an economic evaluation framework. The thesis has looked to explore why treatment sequences are used and why they present a challenge for economic evaluation, and reviewed economic evaluations in RA to identify why sequences have not been fully considered. The thesis then framed the treatment sequencing problem as an optimisation problem, and reviewed simulation optimisation methods and assessed their appropriateness and feasibility for addressing the treatment sequencing problem. Finally, the thesis contains an implementation of SOSA, and considers its real world applicability within a health economic evaluation context.

The review of economic evaluations in Chapter 4 reaffirmed a prior belief, that existing analyses have not sought an optimal treatment sequence in RA. Coupled with NICE undertaking numerous partial evaluations and appraisals, there is a significant risk that sub-optimal guidance has been developed for RA. This concern could pass through to other chronic conditions where treatment sequences are common.

SOSA has been identified and applied, which represents a novel use of a methodology from another field, and its application to a health economic evaluation problem is unique. SOSA has shown promise as a potential method for resolving treatment sequencing health economic evaluation problems, where a simulation model is computationally expensive, and the number of alternative competing solutions is large.

The implementation of SOSA found that the best solution found by the analysis generates greater net monetary benefit than the current sequence recommended by NICE. However, the sequence identified is more cost-effective, but at reduced total QALYs. While plausible and economically valid, it is unlikely to see implementation by NICE. Identifiable patients would suffer reduced health benefits while unidentifiable patients would benefit.

When this research began, there was an absence of literature informing the development of health economic models for treatment sequencing problems. As such, there was variability in the approaches used to model treatment sequences.⁹⁴ The research undertaken has attempted to fill this void, and reduce this variability, by drawing across academic disciplines and applying and evaluating a methodology which has not previously been attempted.

Overall, the aims and objectives of this thesis have been achieved. The research represents a novel and innovative addition to the body of methodological research that underpins health economic evaluation. There is scope for further research; however this thesis represents a significant first step in ensuring that optimality remains a goal, even in complex chronic conditions

REFERENCES

1. Scott D, Steer S. The course of established rheumatoid arthritis. *Best Pract Res Clin Rheumatol*. 2007.
2. Allaire S, Wolfe F, Niu J. Current risk factors for work disability associated with rheumatoid arthritis: recent data from a US national cohort. *Arthritis Care* 2009.
3. Cooper NJ. Economic burden of rheumatoid arthritis: a systematic review. *Rheumatology*. 2000;39(1):28-33. doi:10.1093/rheumatology/39.1.28.
4. Naz S, Symmons D. Mortality in established rheumatoid arthritis. *Best Pract Res Clin Rheumatol*. 2007.
5. Dadoun S, Zeboulon-Ktorza N, Combescure C. Mortality in rheumatoid arthritis over the last fifty years: systematic review and meta-analysis. *Jt Bone Spine*. 2013.
6. Symmons D. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology*. 2002;41(7):793-800. doi:10.1093/rheumatology/41.7.793.
7. Symmons D, Barrett E. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. 1994.
8. Alamanos Y, Drosos A. Epidemiology of adult rheumatoid arthritis. *Autoimmun Rev*. 2005.
9. Arnett F, Edworthy S. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis* 1988.
10. Combe B, Landewé R, Lukas C. EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including. *Ann* 2007.
11. Felson D, Anderson J. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis* 1995.
12. Gestel A Van, Prevo M. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis: comparison with the preliminary American. *Arthritis* 1996.
13. Stiglitz J, Driffill J. *Economics*. New York and London: Norton; 1997.
14. Katz ML, Rosen HS. *Microeconomics*. New York. McGraw Hill Higher Education; 1998.
15. Morris S, Devlin N, Parkin D. *Economic Analysis in Health Care.*; 2007.
16. Bentham J. *An Introduction to the Principles of Morals and Legislation (Collected Works of Jeremy Bentham)*. Oxford: Clarendon Press; 1907.
17. Becker G. A Theory of the Allocation of Time. *Econ J*. 1965;75(299):493-517.
18. Grossman M. On the concept of health capital and the demand for health. *J Polit Econ*. 1972;80(2):223-255.

19. Frenk J. The global health system: strengthening national health systems as the next step for global progress. *PLoS Med.* 2010;7(1).
20. Kleczkowski B, Roemer M, Van Der Werff A. *National Health Systems and Their Reorientation towards Health for All: Guidance for Policy-Making.* Geneva; 1984.
21. Deutsch M. Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *J Soc Issues.* 1975.
22. Beauchamp T, Childress J. *Principles of Biomedical Ethics.*; 2001.
23. Culyer A., Wagstaff A. Equity and equality in health and health care. *J Health Econ.* 1993;12(4):431-457. doi:10.1016/0167-6296(93)90004-X.
24. Oliver A, Mossialos E. Equity of access to health care: outlining the foundations for action. *J Epidemiol Community Health.* 2004;58(8):655-658. doi:10.1136/jech.2003.017731.
25. Wagstaff A. QALYs and the equity-efficiency trade-off. *J Health Econ.* 1991;10(1):21-41.
26. Sen A. Welfare, preference and freedom. *J Econom.* 1991;50(1-2):15-29.
27. Ng Y-K. A case for happiness, cardinalism, and interpersonal comparability. *Econ J.* 1997;107(445):1848-1858. doi:10.1111/j.1468-0297.1997.tb00087.x.
28. Drummond MF, McGuire A. *Economic Evaluation in Health Care: Merging Theory with Practice.* Oxford: Oxford University Press; 2001.
29. Johansson P. *An Introduction to Modern Welfare Economics.* Cambridge University Press; 1991.
30. Sloman MJ, Wride PA. *Economics with MyEconLab.* Financial Times/ Prentice Hall
31. Pauly M. Taxation, health insurance, and market failure in the medical economy. *J Econ Lit.* 1986.
32. Donaldson C, Mitton C. *Priority Setting Toolkit: A Guide to the Use of Economics in Health Care Decision Making.* Wiley-Blackwell; 2004.
33. Fuchs V. *Who Shall Live?: Health, Economics, and Social Choice.* WSPC; 1974.
34. Brouwer WBF, Koopmanschap MA. On the economic foundations of CEA. Ladies and gentlemen, take your positions! *J Health Econ.* 2000;19(4):439-459.
35. Sen A. The possibility of social choice. *Am Econ Rev.* 1999;89(3):349-378.
36. Broome J. Qalys. *J Public Econ.* 1993;50(2):149-167.
37. Mason H, Jones-Lee M, Donaldson C. Modelling the monetary value of a QALY: a new approach based on UK data. *Health Econ.* 2009;18(8):933-950. doi:10.1002/hec.1416.
38. Drummond M, Torrance G, Mason J. Cost-effectiveness league tables: more harm than good? *Soc Sci Med.* 1993;37(1):33-40.
39. Towse A. Should NICE's threshold range for cost per QALY be raised? Yes. *BMJ.* 2009;338:b181.
40. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics.* 2008;26(9):733-744.

41. Claxton K, Martin S, Soares M, Rice N. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. 2015.
42. Brennan A, Akehurst R. Modelling in health economic evaluation. What is its place? What is its value? *Pharmacoeconomics*. 2000;17(5):445-459.
43. Briggs A, Sculpher M, Claxton K. *Decision Modelling for Health Economic Evaluation*. Oxford University Press; 2006.
44. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ*. 2006;15(12):1295-1310. doi:10.1002/hec.1148.
45. Rawlins M, Barnett D, Stevens A. Pharmacoeconomics: NICE's approach to decision-making. *Br J Clin Pharmacol*. 2010;70(3):346-349. doi:10.1111/j.1365-2125.2009.03589.x.
46. Secretary of State for Health. *Directions to Primary Care Trusts and NHS Trusts in England Concerning Arrangements for the Funding of Technology Appraisal Guidance from the National Institute for Clinical Excellence (NICE)*. London; 2003.
47. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*.; 2013.
48. Gafni A, Birch S. Incremental cost-effectiveness ratios (ICERs): the silence of the lambda. *Soc Sci Med*. 2006;62(9):2091-2100.
49. Birch S, Gafni A. Information created to evade reality (ICER): things we should not look to for answers. *Pharmacoeconomics*. 2006;24(11):1121-1131.
50. Birch S, Gafni A. On being NICE in the UK: guidelines for technology appraisal for the NHS in England and Wales. *Health Econ*. 2002;11(3):185-191.
51. Royston G. *Meeting Global Health Challenges through Operational Research and Management Science*.; 2011.
52. Rosenhead J. Reflections on fifty years of operational research. *J Oper Res Soc*. 2009;60:S5-S15.
53. Mingers J. Variety is the spice of life: combining soft and hard OR/MS methods. *Int Trans Oper Res*. 2000;7(6):673-691. doi:10.1111/j.1475-3995.2000.tb00224.x.
54. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M. Conceptualizing a Model: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. *Med Decis Making*. 2012;32(5):678-689. doi:10.1177/0272989X12454941.
55. Chong EKP, Zak SH. *An Introduction to Optimization*. John Wiley & Sons; 2013.
56. Stinnett AA, Paltiel AD. Mathematical programming for the efficient allocation of health care resources. *J Health Econ*. 1996.
57. Birch S, Gafni A. The biggest bang for the buck or bigger bucks for the bang: the fallacy of the cost-effectiveness threshold. *J Health Serv Res Policy*. 2006;11(1):46-51.
58. Torrance G, Thomas W, Sackett D. A utility maximization model for evaluation of health care programs. *Health Serv Res*. 1972;7(2):118-133.
59. Birch S, Gafni A. Cost effectiveness/utility analyses. Do current decision rules lead us to where we want to be? *J Health Econ*. 1992;11(3):279-296.

60. Dekker J, Van HL, Hendriksen M, et al. What Is the Best Sequential Treatment Strategy in the Treatment of Depression? Adding Pharmacotherapy to Psychotherapy or Vice Versa?. *Psychother Psychosom.* 2012;82(2):89-98. doi:10.1159/000341177.
61. Weintraub M, Waldman E, Koplewitz B, et al. A sequential treatment algorithm for infants with stage 4s neuroblastoma and massive hepatomegaly. *Pediatr Blood Cancer.* 2012;59(1):182-184. doi:10.1002/pbc.23186.
62. Trappe R, Oertel S, Leblond V, et al. Sequential treatment with rituximab followed by CHOP chemotherapy in adult B-cell post-transplant lymphoproliferative disorder (PTLD): the prospective international multicentre phase 2 PTLD-1 trial. *Lancet Oncol.* 2012;13(2):196-206. doi:10.1016/S1470-2045(11)70300-X.
63. Palmer S, Sculpher M. *Tocilizumab for the Treatment of Rheumatoid Arthritis. Report by the Decision Support Unit.*; 2010.
64. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-Effectiveness in Health and Medicine.* OUP USA; 1996.
65. Sarac SB, Rasmussen CH, Rasmussen MA, et al. A comprehensive approach to benefit-risk assessment in drug development. *Basic Clin Pharmacol Toxicol.* 2012;111(1):65-72. doi:10.1111/j.1742-7843.2012.00871.x.
66. Wagner EH, Austin BT, Korff M Von. Organizing Care for Patients with Chronic Illness. *Milbank Q.* 1996;74(4):511-544.
67. Goekoop-Ruiterman YPM, de Vries-Bouwstra JK, Allaart CF, et al. Comparison of treatment strategies in early rheumatoid arthritis: a randomized trial. *Ann Intern Med.* 2007;146(6):406-415.
68. Allaart CF, Goekoop-Ruiterman YPM, de Vries-Bouwstra JK, Breedveld FC, Dijkmans BAC. Aiming at low disease activity in rheumatoid arthritis with initial combination therapy or initial monotherapy strategies: the BeSt study. *Clin Exp Rheumatol.* 24(6 Suppl 43):S - 77-82.
69. Van der Kooij SM, de Vries-Bouwstra JK, Goekoop-Ruiterman YPM, et al. Limited efficacy of conventional DMARDs after initial methotrexate failure in patients with recent onset rheumatoid arthritis treated according to the disease activity score. *Ann Rheum Dis.* 2007;66(10):1356-1362.
70. Goekoop-Ruiterman YPM, de Vries-Bouwstra JK, Allaart CF, et al. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled. *Arthritis Rheum.* 2005;52(11):3381-3390.
71. Chan ESL., Cronstein BN. Methotrexate—how does it really work? *Nat Rev Rheumatol.* 2010;6:175-178.
72. Cutolo M. Anti-inflammatory mechanisms of methotrexate in rheumatoid arthritis. *Ann Rheum Dis.* 2001;60(8):729-735. doi:10.1136/ard.60.8.729.
73. Rachapalli SM, Williams R, Walsh DA, Young A, Kiely PDW, Choy EH. First-line DMARD choice in early rheumatoid arthritis—do prognostic factors play a role? *Rheumatology (Oxford).* 2010;49(7):1267-1271. doi:10.1093/rheumatology/kep389.
74. Alten R, Pohl C, Choy EH, et al. Developing a construct to evaluate flares in rheumatoid arthritis: a conceptual report of the OMERACT RA Flare Definition Working Group. *J Rheumatol.* 2011;38(8):1745-1750. doi:10.3899/jrheum.110400.

75. Lamb C, Price M, Jones D, Mansfield J. Patients with crohn's disease report high levels of fatigue during a symptomatic disease flare. *Gut*. 2011;60:A206.
76. Lublin F, Reingold S. Defining the clinical course of multiple sclerosis results of an international survey. *Neurology*. 1996.
77. Montori VM, Gafni A, Charles C. A shared treatment decision-making approach between patients with chronic conditions and their clinicians: the case of diabetes. *Health Expect*. 2006;9(1):25-36. doi:10.1111/j.1369-7625.2006.00359.x.
78. Drummond MF, Sculpher MJ, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford University Press; 2005.
79. D'Agostino RB. Estimating treatment effects using observational data. *JAMA J Am Med Assoc*. 2007;297(3).
80. Rubin D. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol*. 1974;66(5):688-701.
81. Stukel T, Fisher E, Wennberg D, Alter D, Gottlieb D, Vermeulen M. Analysis of observational studies in the presence of treatment selection bias. *JAMA J Am Med Assoc*. 2007;297(3):278-285.
82. Almbrand B, Johannesson M, Sjöstrand B, Malmberg K, Rydén L. Cost-effectiveness of intense insulin treatment after acute myocardial infarction in patients with diabetes mellitus; results from the DIGAMI study. *Eur Heart J*. 2000;21(9):733-739. doi:10.1053/euhj.1999.1859.
83. Björholt I, Andersson FL, Kahan T, Ostergren J. The cost-effectiveness of ramipril in the treatment of patients at high risk of cardiovascular events: a Swedish sub-study to the HOPE study. *J Intern Med*. 2002;251(6):508-517.
84. Johannesson M, Meltzer D, O'Conor RM. Incorporating future costs in medical cost-effectiveness analysis: implications for the cost-effectiveness of the treatment of hypertension. *Med Decis Making*. 17(4):382-389.
85. Scott DL, Cope A. New tumour necrosis factor inhibitors for rheumatoid arthritis: are there benefits from extending choice? *Ann Rheum Dis*. 2009;68(6):767-769.
86. Walker D, Hall S. NICE guidance on rheumatoid arthritis. *Rheumatology*. 2010;49(4):619-620.
87. Kessler D, Rose J, Temple R, Schapiro R, Griffin J. Therapeutic-class wars--drug promotion in a competitive marketplace. *N Engl J Med*. 1994;331:1350-1353.
88. Donaldson C, Shackley P. Does "process utility" exist? A case study of willingness to pay for laparoscopic cholecystectomy. *Soc Sci Med*. 1997;44(5):699-707.
89. Frank RG, Salkever DS. Generic Entry and the Pricing of Pharmaceuticals. *J Econ & Manag Strateg*. 1997;6(1):75-90. doi:10.1111/j.1430-9134.1997.00075.x.
90. Halimi S, Schweizer A, Minic B, Foley J, Dejager S. Combination treatment in the management of type 2 diabetes: focus on vildagliptin and metformin as a single tablet. *Vasc Health Risk Manag*. 2008;4(3):481-492.
91. Balzarini J. Suppression of resistance to drugs targeted to human immunodeficiency virus reverse transcriptase by combination therapy. *Biochem Pharmacol*. 1999;58(1):1-27.

92. Jacobi S, MacLeod R. Making sense of chronic illness--a therapeutic approach. *J Prim Health Care*. 2011;3(2):136-141.
93. Van Gestel A, Severens JL, Webers CAB, Beckers HJM, Jansonius NM, Schouten JSAG. Modeling complex treatment strategies: construction and validation of a discrete event simulation model for glaucoma. *Value Health*. 13(4):358-367. doi:10.1111/j.1524-4733.2009.00678.x.
94. Tosh J, Stevenson M, Akehurst R. Health economic modelling of treatment sequences for rheumatoid arthritis: a systematic review. *Curr Rheumatol Rep*. 2014;16(10):447. doi:10.1007/s11926-014-0447-2.
95. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339(jul21 1):b2535-b2535. doi:10.1136/bmj.b2535.
96. Arshad A, Sulaiman W. Optimizing the use of traditional DMARD in RA: getting the most out of what we can afford! *APLAR J Rheumatol*. 2007;10(1):3-8. doi:10.1111/j.1479-8077.2007.00247.x.
97. Bansback N, Regier D, Ara R. An overview of economic evaluations for drugs used in rheumatoid arthritis: focus on tumour necrosis factor-alpha antagonists. *Drugs*. 2005.
98. Bansback N, Ara R, Karnon J, Anis A. Economic evaluations in rheumatoid arthritis: a critical review of measures used to define health States. *Pharmacoeconomics*. 2008;26(5):395-408.
99. Bansback N, Brennan A, Anis AH. A pharmacoeconomic review of adalimumab in the treatment of rheumatoid arthritis. *Expert Rev Pharmacoecon Outcomes Res*. 2005;5(5):519-529. doi:10.1586/14737167.5.5.519.
100. Bansback N, Marra CA, Finckh A, Anis A. The economics of treatment in early rheumatoid arthritis. *Best Pract Res Clin Rheumatol*. 2009;23(1):83-92. doi:10.1016/j.berh.2008.11.005.
101. Benucci M, Saviola G, Manfredi M, Sarzi-Puttini P, Atzeni F. Cost effectiveness analysis of disease-modifying antirheumatic drugs in rheumatoid arthritis. A systematic review literature. *Int J Rheumatol*. 2011;2011:845496. doi:10.1155/2011/845496.
102. Doan Q V, Chiou C-F, Dubois RW. Review of eight pharmacoeconomic studies of the value of biologic DMARDs (adalimumab, etanercept, and infliximab) in the management of rheumatoid arthritis. *J Manag Care Pharm*. 2006;12(7):555-569.
103. Emery P. Review of health economics modelling in rheumatoid arthritis. *Pharmacoeconomics*. 2004;22(2 Suppl 1):55-69.
104. Ferraz MB, Maetzel A, Bombardier C. A summary of economic evaluations published in the field of rheumatology and related disciplines. *Arthritis Rheum*. 1997;40(9):1587-1593. doi:10.1002/1529-0131(199709)40:9<1587::AID-ART7>3.0.CO;2-C.
105. Fleurence R, Spackman E. Cost-effectiveness of biologic agents for treatment of autoimmune disorders: structured review of the literature. *J Rheumatol*. 2006;33(11):2124-2131.
106. Fries JF. Safety, cost and effectiveness issues with disease modifying anti-rheumatic drugs in rheumatoid arthritis. *Ann Rheum Dis*. 1999;58 Suppl 1:I86-I89.
107. Gabriel SE, Coyle D, Moreland LW. A clinical and economic review of disease-modifying antirheumatic drugs. *Pharmacoeconomics*. 2001;19(7):715-728.

108. Kavanaugh A. The pharmacoeconomics of newer therapeutics for rheumatic diseases. *Rheum Dis Clin North Am.* 2006;32(1):45-56.
109. Kobelt G, Jönsson B. The burden of rheumatoid arthritis and access to treatment: outcome and cost-utility of treatments. *Eur J Heal Econ.* 2008;8(2):95-106.
110. Lyseng-Williamson K, Foster R. Infliximab: a pharmacoeconomic review of its use in rheumatoid arthritis. *Pharmacoeconomics.* 2004;22(2):107-132.
111. Lyseng-Williamson KA, Foster R, Plosker GL. Etanercept: a pharmacoeconomic review of its use in rheumatoid arthritis. *Pharmacoeconomics.* 2004;22(16):1071-1095.
112. Maetzel A. Cost-effectiveness estimates reported for tumor necrosis factor blocking agents in rheumatoid arthritis refractory to methotrexate--a brief summary. *J Rheumatol Suppl.* 2005;72:51-53.
113. Maetzel A, Ferraz MB, Bombardier C. A review of cost-effectiveness analyses in rheumatology and related disciplines. *Curr Opin Rheumatol.* 1998;10(2):136-140.
114. Merkesdal S, Ruof J, Mittendorf T, Zeidler H. Cost-effectiveness of TNF- α -blocking agents in the treatment of rheumatoid arthritis. *Expert Opin Pharmacother.* 2004;5(9):1881-1886. doi:10.1517/14656566.5.9.1881.
115. Nurmohamed M, Dijkmans B. Efficacy, tolerability and cost effectiveness of disease-modifying antirheumatic drugs and biologic agents in rheumatoid arthritis. *Drugs.* 2005;65(5):661-694.
116. Ruchlin HS, Elkin EB, Paget SA. Assessing cost-effectiveness analyses in rheumatoid arthritis and osteoarthritis. *Arthritis Care Res.* 1997;10(6):413-421.
117. Schoels M, Wong J, Scott D, et al. Economic aspects of treatment options in rheumatoid arthritis: a systematic literature review informing the EULAR recommendations for the management of. *Ann Rheum Dis.* 2010;69(6):995-1003.
118. Tella MN, Feinglass J, Chang RW. Cost-effectiveness, cost-utility, and cost-benefit studies in rheumatology: a review of the literature, 2001-2002. *Curr Opin Rheumatol.* 2003;15(2):127-131.
119. Van der Velde G, Pham B, Machado M, et al. Cost-effectiveness of biologic response modifiers compared to disease-modifying antirheumatic drugs for rheumatoid arthritis: a systematic review. *Arthritis Care Res (Hoboken).* 2011;63(1):65-78. doi:10.1002/acr.20338.
120. Sullivan SD, Alfonso-Cristancho R, Carlson J, Mallya U, Ringold S. Economic consequences of sequencing biologics in rheumatoid arthritis: a systematic review. *J Med Econ.* 2013;16(3):391-396. doi:10.3111/13696998.2013.763812.
121. National Collaborating Centre for Chronic Conditions. *Rheumatoid Arthritis: National Clinical Guideline for Management and Treatment in Adults.* London; 2009.
122. Brazier J, Deverill M, Green C. A Review of the Use of Health Status Measures in Economic Evaluation. *J Health Serv Res Policy.* 1999;4(3):174-184. doi:10.1177/135581969900400310.
123. Räsänen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryyänen O-P, Roine R. Use of quality-adjusted life years for the estimation of effectiveness of health care: A systematic literature review. *Int J Technol Assess Health Care.* 2006;22(2):235-241. doi:10.1017/S0266462306051051.
124. Chen Y-F, Jobanputra P, Barton P, et al. *A Systematic Review of the Effectiveness of Adalimumab, Etanercept and Infliximab for the Treatment of Rheumatoid Arthritis in Adults and an Economic*

Evaluation of Their Cost-Effectiveness. England: West Midlands Health Technology Assessment Collaboration (WMHTAC), Department of Public Health and Epidemiology, University of Birmingham, UK.; 2006.

125. Davies A, Cifaldi MA, Segurado OG, Weisman MH. Cost-effectiveness of sequential therapy with tumor necrosis factor antagonists in early rheumatoid arthritis. *J Rheumatol*. 2009;36(1):16-26.
126. Finckh A, Bansback N, Marra CA, et al. Treatment of very early rheumatoid arthritis with symptomatic therapy, disease-modifying antirheumatic drugs, or biologic agents: a cost-effectiveness analysis. *Ann Intern Med*. 2009;151(9):612-621.
127. Hartman M, van Ede A, Severens JL, Laan RFJM, van de Putte L, van der Wilt GJ. Economic evaluation of folate supplementation during methotrexate treatment in rheumatoid arthritis. *J Rheumatol*. 2004;31(5):902-908.
128. Kobelt G, Lindgren P, Young A. Modelling the costs and effects of leflunomide in rheumatoid arthritis. *Eur J Health Econ*. 2002;3(3):180-187.
129. Kobelt G, Lekander I, Lang A, Raffener B, Botsios C, Geborek P. Cost-effectiveness of etanercept treatment in early active rheumatoid arthritis followed by dose adjustment. *Int J Technol Assess Health Care*. 2011;27(3):193-200.
130. Korthals-de Bos IBC, van Tulder MW, Boers M, et al. Indirect and total costs of early rheumatoid arthritis: A randomized comparison of combined step-down prednisolone, methotrexate and sulfasalazine with sulfasalazine alone. *Arthritis Rheum*. 2001;44(9):S313-S313.
131. Maetzel A, Strand V, Tugwell P, Wells G, Bombardier C. Cost effectiveness of adding leflunomide to a 5-year strategy of conventional disease-modifying antirheumatic drugs in patients with rheumatoid arthritis. *Arthritis Rheum*. 2002;47(6):655-661.
132. Schädlich PK, Zeidler H, Zink A, et al. Modelling cost effectiveness and cost utility of sequential DMARD therapy including leflunomide for rheumatoid arthritis in Germany: II. The contribution of leflunomide to efficiency. *Pharmacoeconomics*. 2005;23(4):395-420.
133. Schipper LG, Kievit W, den Broeder AA, et al. Treatment strategies aiming at remission in early rheumatoid arthritis patients: starting with methotrexate monotherapy is cost-effective. *Rheumatology (Oxford)*. 2011;50(7):1320-1330.
134. Spalding JR, Hay J. Cost effectiveness of tumour necrosis factor-alpha inhibitors as first-line agents in rheumatoid arthritis. *Pharmacoeconomics*. 2006;24(12):1221-1232.
135. Tosh JC, Wailoo AJ, Scott DL, Deighton CM. Cost-effectiveness of combination nonbiologic disease-modifying antirheumatic drug strategies in patients with early rheumatoid arthritis. *J Rheumatol*. 2011;38(8):1593-1600.
136. Van den Hout WB, Goekoop-Ruiterman YPM, Allaart CF, et al. Cost-utility analysis of treatment strategies in patients with recent-onset rheumatoid arthritis. *Arthritis Rheum Arthritis Care Res*. 2009;61(3):291-299.
137. Verhoeven AC, Bibo JC, Boers M, Engel GL, van der Linden S. Cost-effectiveness and cost-utility of combination therapy in early rheumatoid arthritis: randomized comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone. COBRA Trial Group. Combinatietherapie Bij Reumat. *Br J Rheumatol*. 1998;37(10):1102-1109.
138. Anis AH, Tugwell PX, Wells GA, Stewart DG. A cost effectiveness analysis of cyclosporine in rheumatoid arthritis. *J Rheumatol*. 1996;23(4):609-616.

139. Bansback NJ, Brennan A, Ghatnekar O. Cost effectiveness of adalimumab in the treatment of patients with moderate to severe rheumatoid arthritis in Sweden. *Ann Rheum Dis.* 2005;64(7):995-1002.
140. Barbieri M, Wong JB, Drummond M. The cost effectiveness of infliximab for severe treatment-resistant rheumatoid arthritis in the UK. *Pharmacoeconomics.* 2005;23(6):607-618.
141. Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. *The Use of Modelling to Evaluate New Drugs for Patients with a Chronic Condition: The Case of Antibodies against Tumour Necrosis Factor in Rheumatoid Arthritis.*; 2004.
142. Benucci M, Gobbi F, Sabadini L, Saviola G, Baiardi P, Manfredi M. The Economic Burden of Biological Therapy in Rheumatoid Arthritis in Clinical Practice: Cost-Effectiveness Analysis of Sub-Cutaneous Anti-Tnf Alpha Treatment in Italian Patients. *Int J Immunopathol Pharmacol.* 2009;22(4):1147-1152.
143. Benucci M, Saviola G, Baiardi P, Manfredi M. Cost-effectiveness treatment with Rituximab in patients with rheumatoid arthritis in real life. *Rheumatol Int.* 2011;31(11):1465-1469.
144. Beresniak A, Ariza-Ariza R, Garcia-Llorente JF, Ramirez-Arellano A, Dupont D. Modelling cost-effectiveness of biologic treatments based on disease activity scores for the management of rheumatoid arthritis in Spain. *Int J Inflamm.* 2011;2011.
145. Brennan A, Bansback N, Reynolds A, Conway P. Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK. *Rheumatology (Oxford).* 2004;43(1):62-72.
146. Brennan A, Bansback N, Nixon R, et al. Modelling the cost effectiveness of TNF-alpha antagonists in the management of rheumatoid arthritis: results from the British Society for Rheumatology Biologics Registry. *Rheumatology (Oxford).* 2007;46(8):1345-1354.
147. Chiou C-F, Choi J, Reyes CM. Cost-effectiveness analysis of biological treatments for rheumatoid arthritis. *Expert Rev Pharmacoeconomics Outcomes Res.* 2004;4(3):307-315.
148. Choi HK, Seeger JD, Kuntz KM. A cost-effectiveness analysis of treatment options for patients with methotrexate-resistant rheumatoid arthritis. *Arthritis Rheum.* 2000;43(10):2316-2327.
149. Choi HK, Seeger JD, Kuntz KM. A cost effectiveness analysis of treatment options for methotrexate-naive rheumatoid arthritis. *J Rheumatol.* 2002;29(6):1156-1165.
150. Cimmino MA, Leardini G, Salaffi F, et al. Assessing the cost-effectiveness of biologic agents for the management of moderate-to-severe rheumatoid arthritis in anti-TNF inadequate responders in Italy: a modelling approach. *Clin Exp Rheumatol.* 2011;29(4):633-641.
151. Clark W, Jobanputra P, Barton P, Burls A. The clinical and cost-effectiveness of anakinra for the treatment of rheumatoid arthritis in adults: a systematic review and economic analysis. *Health Technol Assess.* 2004;8(18):iii - 105.
152. Coyle D, Judd M, Blumenauer B, et al. *Infliximab and Etanercept in Patients with Rheumatoid Arthritis: A Systematic Review and Economic Evaluation (Technology Report No 64).* Ottawa: Ottawa: Canadian Coordinating Office for Health Technology Assessment; 2006.
153. Diamantopoulos A, Benucci M, Capri S, et al. Economic evaluation of tocilizumab combination in the treatment of moderate-to-severe rheumatoid arthritis in Italy. *J Med Econ.* 2012;15(3):576-585.

154. Hallinen TA, Soini EJO, Eklund K, Puolakka K. Cost-utility of different treatment strategies after the failure of tumour necrosis factor inhibitor in rheumatoid arthritis in the Finnish setting. *Rheumatology (Oxford)*. 2010;49(4):767-777.
155. Jobanputra P, Barton P, Bryan S, Burls A. *The Effectiveness of Infliximab and Etanercept for the Treatment of Rheumatoid Arthritis: A Systematic Review and Economic Evaluation*. England: Department of Public Health and Epidemiology, University of Birmingham, UK.; 2002.
156. Kavanaugh A, Heudebert G, Cush J, Jain R. Cost evaluation of novel therapeutics in rheumatoid arthritis (CENTRA): A decision analysis model. *Semin Arthritis Rheum*. 1996;25(5):297-307.
157. Kielhorn A, Porter D, Diamantopoulos A, Lewis G. UK cost-utility analysis of rituximab in patients with rheumatoid arthritis that failed to respond adequately to a biologic disease-modifying antirheumatic drug. *Curr Med Res Opin*. 2008;24(9):2639-2650.
158. Kievit W, Fransen J, Adang EM, et al. Evaluating guidelines on continuation of anti-tumour necrosis factor treatment after 3 months: clinical effectiveness and costs of observed care and different alternative strategies. *Ann Rheum Dis*. 2009;68(6):844-849.
159. Kobelt G, Jonsson L, Young A, Eberhardt K. The cost-effectiveness of infliximab (Remicade) in the treatment of rheumatoid arthritis in Sweden and the United Kingdom based on the ATTRACT study. *Rheumatology (Oxford)*. 2003;42(2):326-335.
160. Kobelt G, Eberhardt K, Geborek P. TNF inhibitors in the treatment of rheumatoid arthritis in clinical practice: costs and outcomes in a follow up study of patients with RA treated with etanercept or infliximab in southern Sweden. *Ann Rheum Dis*. 2004;63(1):4-10.
161. Kobelt G, Lindgren P, Singh A, Klareskog L. Cost effectiveness of etanercept (Enbrel) in combination with methotrexate in the treatment of active rheumatoid arthritis based on the TEMPO trial. *Ann Rheum Dis*. 2005;64(8):1174-1179.
162. Lekander I, Borgstrom F, Svarvar P, Ljung T, Carli C, van Vollenhoven RF. Cost-effectiveness of real-world infliximab use in patients with rheumatoid arthritis in Sweden. *Int J Technol Assess Health Care*. 2010;26(1):54-61.
163. Lindgren P, Geborek P, Kobelt G. Modeling the cost-effectiveness of treatment of rheumatoid arthritis with rituximab using registry data from Southern Sweden. *Int J Technol Assess Health Care*. 2009;25(2):181-189.
164. Maetzel A, Strand V, Tugwell P, Wells G, Bombardier C. Economic comparison of leflunomide and methotrexate in patients with rheumatoid arthritis: an evaluation based on a 1-year randomised controlled trial. *Pharmacoeconomics*. 2002;20(1):61-70.
165. Malottki K, Barton P, Tsourapas A, et al. Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor: a systematic review and economic evaluation. *Health Technol Assess*. 2011;15(14):1-278.
166. Marra CA, Marion SA, Guh DP, et al. Not all "Quality-adjusted life years" are equal. *J Clin Epidemiol*. 2007;60(6):616-624.
167. Merkesdal S, Kirchhoff T, Wolka D, Ladinek G, Kielhorn A, Rubbert-Roth A. Cost-effectiveness analysis of rituximab treatment in patients in Germany with rheumatoid arthritis after etanercept-failure. *Eur J Health Econ*. 2010;11(1):95-104.

168. Nuijten MJ, Engelfriet P, Duijn K, Bruijn G, Wierz D, Koopmanschap M. A cost-cost study comparing etanercept with infliximab in rheumatoid arthritis. *Pharmacoeconomics*. 2001;19(10):1051-1064.
169. Osiri M, Kamolratanakul P, Maetzel A, Tugwell P. Cost effectiveness analysis of disease modifying antirheumatic drugs in rheumatoid arthritis. *Rheumatol Int*. 2007;27(11):1063-1069.
170. Rubio-Terrés C, Domínguez-Gil A. Pharmacoeconomic analysis of the treatment of rheumatoid arthritis with leflunomide in comparison with the combination of infliximab and methotrexate. *J Drug Assess*. 2001;4(Part 1):39-54.
171. Russell A, Beresniak A, Bessette L, et al. Cost-effectiveness modeling of abatacept versus other biologic agents in DMARDs and anti-TNF inadequate responders for the management of moderate to severe rheumatoid arthritis. *Clin Rheumatol*. 2009;28(4):403-412.
172. Saraux A, Gossec L, Goupille P, et al. Cost-effectiveness modelling of biological treatment sequences in moderate to severe rheumatoid arthritis in France. *Rheumatology (Oxford)*. 2010;49(4):733-740.
173. Shini VK, Aboobacker S, Pahuja S, Revikumar KG, Bhasi R. Pharmacoeconomic study of DMARDs in the management of rheumatoid arthritis. *Int J Pharm Sci Rev Res*. 2010;5(3):148-154.
174. Soini EJ, Hallinen TA, Puolakka K, Vihervaara V, Kauppi MJ. Cost-effectiveness of adalimumab, etanercept, and tocilizumab as first-line treatments for moderate-to-severe rheumatoid arthritis. *J Med Econ*. 2012;15(2):340-351.
175. Tanno M, Nakamura I, Ito K, et al. Modeling and cost-effectiveness analysis of etanercept in adults with rheumatoid arthritis in Japan: a preliminary analysis. *Mod Rheumatol*. 2006;16(2):77-84.
176. Vera-Llonch M, Massarotti E, Wolfe F, et al. Cost-effectiveness of abatacept in patients with moderately to severely active rheumatoid arthritis and inadequate response to tumor necrosis factor-alpha antagonists. *J Rheumatol*. 2008;35(9):1745-1753.
177. Vera-Llonch M, Massarotti E, Wolfe F, et al. Cost-effectiveness of abatacept in patients with moderately to severely active rheumatoid arthritis and inadequate response to methotrexate. *Rheumatology (Oxford)*. 2008;47(4):535-541.
178. Wailoo AJ, Bansback N, Brennan A, Michaud K, Nixon RM, Wolfe F. Biologic drugs for rheumatoid arthritis in the Medicare program: a cost-effectiveness analysis. *Arthritis Rheum*. 2008;58(4):939-946.
179. Welsing PMJ, Severens JL, Hartman M, van Riel PLCM, Laan RFJM. Modeling the 5-year cost effectiveness of treatment strategies including tumor necrosis factor-blocking agents and leflunomide for treating rheumatoid arthritis in the Netherlands. *Arthritis Rheum*. 2004;51(6):964-973.
180. Wong JB, Singh G, Kavanaugh A. Estimating the cost-effectiveness of 54 weeks of infliximab for rheumatoid arthritis. *Am J Med*. 2002;113(5):400-408.
181. National Institute for Health and Clinical Excellence. *Adalimumab, Etanercept and Infliximab for the Treatment of Rheumatoid Arthritis (TA130)*.; 2007.
182. Deighton C, O'Mahony R, Tosh J, Turner C, Rudolf M. Management of rheumatoid arthritis: summary of NICE guidance. *BMJ*. 2009;338:b702.

183. Gartlehner G, Hansen RA, Jonas BL, Thieda P, Lohr KN. The comparative efficacy and safety of biologics for the treatment of rheumatoid arthritis: a systematic review and metaanalysis. *J Rheumatol*. 2006;33(12):2398-2408.
184. Singh J, Christensen R, Wells GA, et al. A network meta-analysis of randomized controlled trials of biologics for rheumatoid arthritis: a Cochrane overview. *Can Med Assoc J*. 2009;181(11):787-796.
185. Wiens A, Venson R, Correr CJ, Otuki MF, Pontarolo R. Meta-analysis of the efficacy and safety of adalimumab, etanercept, and infliximab for the treatment of rheumatoid arthritis. *Pharmacotherapy*. 2010;30(4):339-353. doi:10.1592/phco.30.4.339.
186. O'Dell JR, Mikuls TR, Taylor TH, et al. Therapies for active rheumatoid arthritis after methotrexate failure. *N Engl J Med*. 2013;369(4):307-318. doi:10.1056/NEJMoa1303006.
187. Tappenden P. Conceptual modelling for health economic model development. *HEDS Discuss Pap 12/05*. 2012.
188. Madan J, Ades AE, Welton NJ. An overview of models used in economic analyses of biologic therapies for arthritis--from current diversity to future consensus. *Rheumatology*. 2011;50(suppl 4):iv10-iv18. doi:10.1093/rheumatology/ker240.
189. Tosh J, Brennan A, Wailoo A, Bansback N. The Sheffield rheumatoid arthritis health economic model. *Rheumatology (Oxford)*. 2011;50 Suppl 4(suppl_4):iv26-iv31. doi:10.1093/rheumatology/ker243.
190. National Institute for Health and Clinical Excellence. *Guide to the Single Technology Appraisal Process.*; 2009.
191. National Institute for Health and Clinical Excellence. *Guide to the Multiple Technology Appraisal Process.*; 2009.
192. Emery P, Breedveld F, Hall S, et al. Comparison of methotrexate monotherapy with a combination of methotrexate and etanercept in active, early, moderate to severe rheumatoid arthritis (COMET): a randomised, double-blind, parallel treatment trial. *Lancet*. 2008;372(9636):375-382.
193. Plant MJ, O'Sullivan MM, Lewis PA, Camilleri JP, Coles EC, Jessop JD. What factors influence functional ability in patients with rheumatoid arthritis. Do they alter over time? *Rheumatology (Oxford)*. 2005;44(9):1181-1185. doi:10.1093/rheumatology/keh707.
194. Symmons D, Tricker K, Roberts C, Davies L, Dawes P, Scott DL. The British Rheumatoid Outcome Study Group (BROSG) randomised controlled trial to compare the effectiveness and cost-effectiveness of aggressive versus symptomatic therapy in established rheumatoid arthritis. *Health Technol Assess*. 2005;9(34):iii - iv, ix - x, 1-78.
195. Munro R, Hampson R, McEntegart A, Thomson EA, Madhok R, Capell H. Improved functional outcome in patients with early rheumatoid arthritis treated with intramuscular gold: results of a five year prospective study. *Ann Rheum Dis*. 1998;57(2):88-93.
196. Scott DL, Khoshaba B, Choy EH, Kingsley GH. Limited correlation between the Health Assessment Questionnaire (HAQ) and EuroQol in rheumatoid arthritis: questionable validity of deriving quality adjusted life. *Ann Rheum* 2007;66(11):1534-1537. doi:10.1136/ard.2007.073726.
197. Norton S, Fu B, Scott DL, et al. Health Assessment Questionnaire disability progression in early rheumatoid arthritis: Systematic review and analysis of two inception cohorts. *Semin Arthritis Rheum*. 2014;44(2):131-144. doi:10.1016/j.semarthrit.2014.05.003.

198. National Institute for Health and Clinical Excellence. *Tocilizumab for the Treatment of Rheumatoid Arthritis (TA247)*.; 2012.
199. Hernández Alava M, Wailoo A, Wolfe F, Michaud K. The relationship between EQ-5D, HAQ and pain in patients with rheumatoid arthritis. *Rheumatology (Oxford)*. 2013;52(5):944-950. doi:10.1093/rheumatology/kes400.
200. Hernández Alava M, Wailoo AJ, Ara R. Tails from the Peak District: Adjusted Limited Dependent Variable Mixture Models of EQ-5D Questionnaire Health State Utility Values. *Value Heal*. 2012;15(3):550-561.
201. Schlosser RW, Wendt O, Bhavnani S, Nail-Chiwetalu B. Use of information-seeking strategies for developing systematic reviews and engaging in evidence-based practice: the application of traditional and comprehensive Pearl Growing. A review. *Int J Lang Commun Disord*. 41(5):567-582. doi:10.1080/13682820600742190.
202. Black N, Brazier J, Fitzpatrick R. *Health services research methods: a guide to best practice*. 1999.
203. Higgins J, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. *Cochrane Collab*. 2011;Version 5.((Updated March 2011)).
204. Dissemination C for R and. *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*. York; 2009.
205. Chilcott J, Brennan A, Booth A, Karnon J, Tappenden P. The role of modelling in prioritising and planning clinical trials. *Health Technol Assess (Rockv)*. 2003;7(23).
206. Andradóttir S. *An Overview of Simulation Optimization via Random Search*. Elsevier; 2006. doi:10.1016/S0927-0507(06)13020-0.
207. Fu MC. Optimization via simulation: A review. *Ann Oper Res*. 1994;53(1):199-247. doi:10.1007/BF02136830.
208. Figueira G, Almada-Lobo B. Hybrid simulation–optimization methods: A taxonomy and discussion. *Simul Model Pract Theory*. 2014;46:118-134.
209. Swisher JR, Hyden PD, Jacobson SH, Schruben LW. A survey of simulation optimization techniques and procedures. In: *2000 Winter Simulation Conference Proceedings (Cat. No.00CH37165)*.Vol 1. IEEE; 2000:119-128. doi:10.1109/WSC.2000.899706.
210. Papadimitriou CH, Steiglitz K. *Combinatorial Optimization: Algorithms and Complexity*. Courier Corporation; 1998.
211. Hromkovič J. *Algorithmics for Hard Problems: Introduction to Combinatorial Optimization, Randomization, Approximation, and Heuristics*.; 2013.
212. Spall J. Stochastic optimization. *Handb Comput Stat*. 2012.
213. Boesel J, Nelson B, Kim S. Using ranking and selection to “clean up” after simulation optimization. *Oper Res*. 2003;51(2):814-825.
214. Jacobson SH, Sullivan KA, Johnson AW. Discrete manufacturing process design optimization using computer simulation and generalized hill climbing algorithms. *Eng Optim*. 1998;31(2):247-260. doi:10.1080/03052159808941372.

215. Kamrani F, Ayani R, Moradi F. A framework for simulation-based optimization of business process models. *Simulation*. 2011;88(7):852-869. doi:10.1177/0037549711417880.
216. Andradóttir S, Prudius AA. Balanced Explorative and Exploitative Search with Estimation for Simulation Optimization. *INFORMS J Comput*. 2009;21(2):193-208. doi:10.1287/ijoc.1080.0309.
217. Hong LJ, Nelson BL. Discrete Optimization via Simulation Using COMPASS. *Oper Res*. 2006;54(1):115-129. doi:10.1287/opre.1050.0237.
218. Huang C, Chang K, Lin J. Optimal vehicle allocation for an Automated Materials Handling System using simulation optimisation. *Int J Prod Res*. 2012;50(20):5734-5746. doi:10.1080/00207543.2011.622311.
219. Ahmed MA, Alkhamis TM, Hasan M. Optimizing discrete stochastic systems using simulated annealing and simulation. *Comput Ind Eng*. 1997;32(4):823-836. doi:10.1016/S0360-8352(97)00006-5.
220. Ahmed MA, Alkhamis TM. Simulation-based optimization using simulated annealing with ranking and selection. *Comput Oper Res*. 2002;29(4):387-402. doi:10.1016/S0305-0548(00)00073-3.
221. Alrefaei MH, Andradóttir S. A Simulated Annealing Algorithm with Constant Temperature for Discrete Stochastic Optimization. *Manage Sci*. 1999;45(5):748-764. doi:10.1287/mnsc.45.5.748.
222. Ghiani G, Legato P, Musmanno R, Vocaturo F. A combined procedure for discrete simulation-optimization problems based on the simulated annealing framework. *Comput Optim Appl*. 2007;38(1):133-145. doi:10.1007/s10589-007-9010-7.
223. Haddock J, Mittenthal J. Simulation optimization using simulated annealing. *Comput Ind Eng*. 1992;22(4):387-395. doi:10.1016/0360-8352(92)90014-B.
224. Lacksonen T. Empirical comparison of search algorithms for discrete event simulation. *Comput Ind Eng*. 2001;40(1-2):133-148. doi:10.1016/S0360-8352(01)00013-4.
225. Rosen SL, Harmonosky CM. An improved simulated annealing simulation optimization method for discrete parameter stochastic systems. *Comput Oper Res*. 2005;32(2):343-358. doi:10.1016/S0305-0548(03)00240-5.
226. Ding H, Benyoucef L, Xie X. A simulation optimization methodology for supplier selection problem. *Int J Comput Integr Manuf*. 2005;18:210-224.
227. Jun Z, Yu-An T, Xue-Lan Z, Jun L. An improved dynamic structure-based neural networks determination approaches to simulation optimization problems. *Neural Comput Appl*. 2010;19(6):883-901. doi:10.1007/s00521-010-0348-x.
228. Korytkowski P, Wiśniewski T, Rymaszewski S. An evolutionary simulation-based optimization approach for dispatching scheduling. *Simul Model Pract Theory*. 2013;35(Complete):69-85.
229. Yang T, Fu H-P, Yang K-Y. An evolutionary-simulation approach for the optimization of multi-constant work-in-process strategy—A case study. *Int J Prod Econ*. 2007;107(1):104-114. doi:10.1016/j.ijpe.2006.02.014.
230. Azadeh A, Maghsoudi A. Optimization of production systems through integration of computer simulation, design of experiment, and Tabu search: the case of a large steelmaking workshop. *Int J Adv Manuf Technol*. 2009;48(5-8):785-800. doi:10.1007/s00170-009-2305-3.

231. Yang * T, Kuo Y, Chang I. Tabu-search simulation optimization approach for flow-shop scheduling with multiple processors — a case study. *Int J Prod Res.* 2004;42(19):4015-4030. doi:10.1080/00207540410001699381.
232. Ho Y-C, Cassandras CG, Chen C-H, Dai L. Ordinal optimisation and simulation. *J Oper Res Soc.* 2000;51(4):490-500.
233. Shi L, Ólafsson S. Nested Partitions Method for Stochastic Optimization. *Methodol Comput Appl Probab.* 2000;2(3):271-291. doi:10.1023/A:1010081212560.
234. Kuo RJ, Yang CY. Simulation optimization using particle swarm optimization algorithm with application to assembly line design. *Appl Soft Comput.* 2011;11(1):605-613. doi:10.1016/j.asoc.2009.12.020.
235. Prudius AA, Andradóttir S. Averaging frameworks for simulation optimization with applications to simulated annealing. *Nav Res Logist.* 2012;59(6):411-429. doi:10.1002/nav.21496.
236. Xu WL, Nelson BL. Empirical stochastic branch-and-bound for optimization via simulation. *IIE Trans.* 2013;45(7):685-698. doi:10.1080/0740817X.2013.768783.
237. Rastrigin L. About Convergence of Random Search Method in Extremal Control of Multi-Parameter Systems. *Avtom Telemekhanika.* 1963.
238. Schumer M, Steiglitz K. Adaptive step size random search. *Autom Control IEEE Trans.* 1968;13(3):270-276.
239. Kregting J, White RC. *Adaptive Random Search.* University of Technology. Department of Electrical Engineering. Group Measurement and Control; 1971.
240. Gosavi A. *Simulation-Based Optimization.* (Springer, ed.). New York: Springer Science+Business Media New York; 2014.
241. Sörensen K, Glover F. Metaheuristics. In: *Encyclopedia of Operations Research and Management Science.*; 2013:960-970.
242. Bentley J. *Writing Efficient Programs.*; 1982.
243. Holland J. Genetic algorithms and the optimal allocation of trials. *SIAM J Comput.* 1973;2(2):88-105.
244. Holland J, Reitman J. Cognitive systems based on adaptive algorithms. *ACM SIGART Bull.* 1977;63:49-49.
245. Holland J. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence.*; 1975.
246. Goldberg D, Holland J. Genetic algorithms and machine learning. *Mach Learn.* 1988;3(2-3):95-99.
247. Henderson D, Jacobson S, Johnson A. The theory and practice of simulated annealing. In: *Handbook of Metaheuristics.*; 2003.
248. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of State Calculations by Fast Computing Machines. *J Chem Phys.* 1953;21(6):1087. doi:10.1063/1.1699114.
249. Černý V. Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *J Optim Theory Appl.* 1985;45(1):41-51.

250. Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. *Science*. 1983;220(4598):671-680.
251. Belisle CJP. Convergence Theorems for a Class of Simulated Annealing Algorithms on Rd. *J Appl Probab*. 1992;29(4):885. doi:10.2307/3214721.
252. Geman S, Geman D. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Trans Pattern Anal Mach Intell*. 1984;PAMI-6(6):721-741. doi:10.1109/TPAMI.1984.4767596.
253. *Simulated Annealing: Parallelization Techniques*. Wiley-Blackwell; 1992.
254. Szu H, Hartley R. Fast simulated annealing. *Phys Lett A*. 1987;122(3-4):157-162. doi:10.1016/0375-9601(87)90796-1.
255. Ingber L. Very fast simulated re-annealing. *Math Comput Model*. 1989;12(8):967-973. doi:10.1016/0895-7177(89)90202-1.
256. White S. Concepts of scale in simulated annealing. In: *The Physics of VLSI*; 1984.
257. Kirkpatrick S, Gelatt C, Vecchi M. Optimization by simulated annealing. *Science* (80-). 1983;220(4598):671-680.
258. Chick SE, Inoue K. New Two-Stage and Sequential Procedures for Selecting the Best Simulated System. *Oper Res*. 2001;49(5):732-743. doi:10.1287/opre.49.5.732.10615.
259. CHEN C-H. A lower for the correct subset-selection probability and its application to discrete-event system simulations. *IEEE Trans Automat Contr*. 41(8):1227-1231.
260. Inoue K, Chick SE, Chen C-H. An empirical evaluation of several methods to select the best system. *ACM Trans Model Comput Simul*. 1999;9(4):381-407. doi:10.1145/352222.352226.
261. Rinott Y. On two-stage selection procedures and related probability-inequalities. *Commun Stat methods*. 1978;7(8):799-811.
262. Almasi GS, Gottlieb A. *Highly Parallel Computing*. Benjamin-Cummings Publishing Co., Inc.; 1989.
263. Patterson D, Hennessy J. *Computer Organization and Design: The Hardware/software Interface*.; 2013.
264. Chu K-W, Deng Y, Reinitz J. Parallel Simulated Annealing by Mixing of States. *J Comput Phys*. 1999;148(2):646-662. doi:10.1006/jcph.1998.6134.
265. Greening DR. Parallel simulated annealing techniques. *Phys D Nonlinear Phenom*. 1990;42(1-3):293-306. doi:10.1016/0167-2789(90)90084-3.
266. Ingber L. Adaptive simulated annealing (ASA): Lessons learned. *Control Cybern*. 1992;16(11):87-100.
267. Roussel-Ragot P, Dreyfus G. A problem independent parallel implementation of simulated annealing: models and experiments. *IEEE Trans Comput Des Integr Circuits Syst*. 1990;9(8):827-835. doi:10.1109/43.57790.
268. Box G, Wilson K. On the Experimental Attainment of Optimum Conditions. *J R Stat Soc Ser B*. 1951;{XIII}(1):1-45.

269. Glover F. Tabu Search—Part I. *ORSA J Comput.* 1989;1(3):190-206. doi:10.1287/ijoc.1.3.190.
270. Glover F. Tabu Search—Part II. *ORSA J Comput.* 1990;2(1):4-32. doi:10.1287/ijoc.2.1.4.
271. Gendreau M, Potvin J. *Handbook of Metaheuristics.*; 2010.
272. Glover F, Laguna M. *Tabu Search.* Kluwer Academic Publishers; 1997.
273. Glover F, Kelly FJP, Laguna M. Genetic algorithms and tabu search: hybrids for optimization. *Comput Oper Res.* 1995;22(1):111-134.
274. Yang T, Rajasekharan M, Peters BA. Semiconductor fabrication facility design using a hybrid search methodology. *Comput Ind Eng.* 1999;36(3):565-583. doi:10.1016/S0360-8352(99)00151-5.
275. Shi L, Ólafsson S. Nested Partitions Method for Global Optimization. *Oper Res.* 2000;48(3):390-407.
276. Shi L, Ólafsson S, Sun N. New parallel randomized algorithms for the traveling salesman problem. *Comput Oper Res.* 1999;26(4):371-394. doi:10.1016/S0305-0548(98)00068-9.
277. Ólafsson S, Shi L. A method for scheduling in parallel manufacturing systems with flexible resources. *IIE Trans (Institute Ind Eng.* 2000;32(2):135-146.
278. Ho Y-C. An ordinal optimization approach to optimal control problems. *Automatica.* 1999;35(2):331-338.
279. Ho Y-C, Sreenivas RS, Vakili P. Ordinal optimization of DEDS. *Discret event Dyn Syst.* 1992;2(1):61-88.
280. Dai L. Convergence properties of ordinal comparison in the simulation of discrete event dynamic systems. *J Optim Theory Appl.* 1996;91(2):363-388.
281. Eberhart RC, Kennedy J. A new optimizer using particle swarm theory. In: *Proceedings of the Sixth International Symposium on Micro Machine and Human Science.* Vol 1. New York, NY; 1995:39-43.
282. Kennedy JF, Kennedy J, Eberhart RC. *Swarm Intelligence.* Morgan Kaufmann; 2001.
283. Cowling P, Soubeiga E. Neighborhood structures for personnel scheduling: a summit meeting scheduling problem. In: *Proceedings of the 3rd Interantional Conference on the Practice and Theory of Autmated Timetabling.*; 2000.
284. Burke E, Erben W, eds. *Practice and Theory of Automated Timetabling III.* Berlin, Heidelberg: Springer Berlin Heidelberg; 2001. doi:10.1007/3-540-44629-X.
285. Burke E, Hyde M, Kendall G, Ochoa G, Ozcan E, Woodward J. A classification of hyper-heuristic approaches. In: *Handbook of Metaheuristics.*; 2010:449-468.
286. Fisher H, Thompson G. Probabilistic learning combinations of local job-shop scheduling rules. In: *Industrial Scheduling.* Prentice-Hall, Englewood Cliffs; 1963:225-251.
287. Crowston W, Glover F, Trawick J. *Probabilistic and Parametric Learning Combinations of Local Job Shop Scheduling Rules.*; 1963.

288. Misir M, Verbeeck K, De Causmaecker P, Vanden Berghe G. *Learning and Intelligent Optimization*. (Hamadi Y, Schoenauer M, eds.). Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. doi:10.1007/978-3-642-34413-8.
289. Misir M, Verbeeck K. A new hyper-heuristic as a general problem solver: an implementation in HyFlex. *J Sched*. 2013.
290. Holmes MW, Goodacre S, Stevenson MD, Pandor A, Pickering A. The cost-effectiveness of diagnostic management strategies for adults with minor head injury. *Injury*. 2012;43(9):1423-1431. doi:10.1016/j.injury.2011.07.017.
291. Meng Y, Ward S, Cooper K, Harnan S, Wyld L. Cost-effectiveness of MRI and PET imaging for the evaluation of axillary lymph node metastases in early stage breast cancer. *Eur J Surg Oncol*. 2011;37(1):40-46. doi:10.1016/j.ejso.2010.10.001.
292. Tappenden P, Chilcott J, Brennan A, Squires H, Glynn-Jones R, Tappenden J. Using whole disease modeling to inform resource allocation decisions: economic evaluation of a clinical guideline for colorectal cancer using a single model. *Value Health*. 2013;16(4):542-553. doi:10.1016/j.jval.2013.02.012.
293. Tosh J, Kearns B, Brennan A, et al. Innovation in health economic modelling of service improvements for longer-term depression: demonstration in a local health community. *BMC Health Serv Res*. 2013;13:150. doi:10.1186/1472-6963-13-150.
294. Tappenden P, Chilcott JB. Avoiding and identifying errors and other threats to the credibility of health economic models. *Pharmacoeconomics*. 2014;32(10):967-979. doi:10.1007/s40273-014-0186-2.
295. Davis S, Stevenson M, Tappenden P, Wailoo A. *NICE DSU Technical Support Document 15: Cost-Effectiveness Modelling Using Patient-Level Simulation.*; 2014.
296. Mendivil F, Shonkwiler R, Spruill M. Restarting search algorithms with applications to simulated annealing. *Adv Appl Probab*. 2001;33(1):242-259.
297. Cheh K, Goldberg J, Askin R. A note on the effect of neighborhood structure in simulated annealing. *Comput Oper Res*. 1991;18(6):537-547.
298. Ogbu F, Smith D. The application of the simulated annealing algorithm to the solution of the n/m/C max flowshop problem. *Comput Oper Res*. 1990;17(3):243-253.
299. Goldstein L, Waterman M. Neighborhood size in the simulated annealing algorithm. *Am J Math Manag Sci*. 1988;8(3-4):409-423.
300. Laarhoven P Van, Aarts E. *Simulated Annealing.*; 1987.
301. Aarts E, Korst J. *Simulated Annealing and Boltzmann Machines*. New York: John Wiley & Sons, Ltd.; 1988.
302. Lundy M, Mees A. Convergence of an annealing algorithm. *Math Program*. 1986;34(1):111-124. doi:10.1007/BF01582166.
303. Tappenden P, Chilcott J, Brennan A, Squires H, Stevenson M. Whole Disease Modelling to inform resource allocation decisions in cancer: A methodological framework. *Value Heal*. 2012;15(8):1127-1136.

304. Brailsford SC, Gutjahr WJ, Rauner MS, Zeppelzauer W. Combined Discrete-event Simulation and Ant Colony Optimisation Approach for Selecting Optimal Screening Policies for Diabetic Retinopathy. *Comput Manag Sci.* 2006;4(1):59-83. doi:10.1007/s10287-006-0008-x.
305. Gutjahr W. A converging ACO algorithm for stochastic combinatorial optimization. *Stoch algorithms Found Appl.* 2003.
306. Gutjahr W. S-ACO: An ant-based approach to combinatorial optimization under uncertainty. *Ant colony Optim swarm Intell.* 2004.
307. Vazirani V. *Approximation Algorithms.*; 2001.
308. Cochran W, Cox G. *Experimental Designs .;* 1957.
309. Myers R, Anderson-Cook C. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments.;* 2009.
310. Law A, Kelton W. *Simulation Modeling and Analysis.;* 1991.
311. Fu M. Optimization for simulation: Theory vs. practice. *INFORMS J Comput.* 2002;14(3):192-215.
312. Nelder J, Mead R. A simplex method for function minimization. *Comput J.* 1965.
313. Kennedy J, Eberhart R. Particle swarm optimization. In: *IEE International Conference on Neural Networks.;* 1995.
314. A. G. Doig, Bya H. Land AGD. An automatic method for solving discrete programming problems. *Econometrica.* 1960;28(3):497-520.
315. Dantzig G. *Linear Programming and Extensions.;* 1965.
316. Ford L, Fulkerson D. *Maximal Flow through a Network.;* 1956.
317. Stuckman B. Comparison of global search methods for design optimization using simulation. In: *1991 Winter Simulation Conference Proceedings.* IEEE; 1991:937-944.
318. Bean JC. Genetic Algorithms and Random Keys for Sequencing and Optimization. *ORSA J Comput.* 1994;6(2):154-160. doi:10.1287/ijoc.6.2.154.
319. Starkweather T, Mcdaniel S, Whitley D, Mathias K. A Comparison of Genetic Sequencing Operators. In: *Proceedings of the Fourth International Conference on Genetic Algorithms.;* 1991.
320. Ficici S, Pollack J. Effects of Finite Populations on Evolutionary Stable Strategies. *GECCO.* 2000.
321. Baker JE. Adaptive Selection Methods for Genetic Algorithms. In: *Proceedings of the 1st International Conference on Genetic Algorithms.* L. Erlbaum Associates Inc.; 1985:101-111.
322. Goldberg DE, Deb K. *A Comparative Analysis of Selection Schemes Used in Genetic Algorithms.* Zurich; 1995.
323. Ochoa G, Harvey I, Buxton H. Optimal Mutation Rates and Selection Pressure in Genetic Algorithms. In: *Genetic and Evolutionary Computation Conference (GECCO).;* 2000.
324. Blickle T. Tournament selection. *Evol Comput.* 2000.

325. Whitley D. The GENITOR Algorithm and Selection Pressure: Why Rank-Based Allocation of Reproductive Trials is Best. In: *Proceedings of the 3rd International Conference on Genetic Algorithms.*; 1989.
326. Michalewicz Z, Fogel DB. *How to Solve It: Modern Heuristics*. Springer-Verlag New York, Inc.; 2004.
327. Rudolph G. Convergence of Evolutionary Algorithms in General Search Spaces. In: *IEEE Conference on Evolutionary Computation.*; 1996.
328. Homaifar A, Qi CX, Lai SH. Constrained Optimization Via Genetic Algorithms. *Simulation*. 1994;62(4):242-253. doi:10.1177/003754979406200405.
329. Hou Y, Chang Y. A New Efficient Encoding Mode of Genetic Algorithms for the Generalized Plant Allocation Problem. *J Inf Sci Eng*. 2004;20:1019-1034.
330. Balasubramanian Rangaswamy FG. Tabu Search Candidate List Strategies in Scheduling. In: *Advances in Computational and Stochastic Optimization, Logic Programming, and Heuristic Search Operations Research/Computer Science Interfaces Series*. New York; 1998:215-233.
331. Henderson SG, Nelson BL. *Handbooks in Operations Research and Management Science: Simulation: Simulation*. Elsevier; 2006.
332. Ólafsson S. Two-Stage Nested Partitions Method for Stochastic Optimization. *Methodol Comput Appl Probab*. 2004;6(1):5-27. doi:10.1023/B:MCAP.0000012413.54789.cc.
333. Lau TWE, Ho YC. Universal Alignment Probabilities and Subset Selection for Ordinal Optimization. *J Optim Theory Appl*. 93(3):455-489. doi:10.1023/A:1022614327007.
334. Ho Y-C, Zhao Q-C, Jia Q-S. *Ordinal Optimization: Soft Optimization for Hard Problems*. Springer; 2008.
335. Shi Y, Eberhart R. A modified particle swarm optimizer. In: *Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on.* IEEE; 1998:69-73.
336. Grigor C, Capell H, Stirling A, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet*. 2004;364(9430):263-269.

APPENDICES

APPENDIX A: RA ECONOMIC EVALUATION SYSTEMATIC REVIEW

APPENDIX A.1: PRISMA CHECKLIST

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	43
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	45
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	43
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	46
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	n/a
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	46

Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	46
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	256
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	46
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	47
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	47
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	47
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	47
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	n/a
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	n/a
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	n/a
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	50
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	52
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	58
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	n/a
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	n/a

Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	n/a
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	n/a
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	71
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	71
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	71
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	1

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

APPENDIX A.2: SEARCH STRATEGIES

Database	Search Strategy
BIOSIS (WoK) all databases	<ol style="list-style-type: none"> 1. Topic=(rheumatoid arthritis) 2. Topic=((disease modifying or disease-modifying or DMARD* or biologic* or therap* or treatment* or anti-rheumatic or anti rheumatic or TNF or tumor necrosis factor alpha or tumour necrosis factor alpha or TNF-alpha or TNF inhibitor* or TNF blocker* or interleukin 1 or IL-1 or monoclonal antibod* or costimulation blocker* or interleukin 6 or IL-6)) 3. Topic=((methotrexate or sulfasalazine or leflunomide or hydroxychloroquine or chloroquine or gold or minocycline or azathioprine or ciclosporin or cyclosporine or penicillamine or cyclophosphamide or etanercept or infliximab or adalimumab or certolizumab* or golimumab or anakinra or rituximab or abatacept or tocilizumab)) 4. Title=((economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or value for money or budget*)) 5. #2 OR #3 6. #1 AND #4 AND #5
Cochrane Database of Systematic Reviews (CDSR), The Cochrane database of methodology reviews, Cochrane Central Register of Controlled Trials (CCRCT), Database of Abstracts of Reviews of Effects (DARE)	<ol style="list-style-type: none"> 1. MeSH descriptor Arthritis, Rheumatoid explode all trees 2. (disease modifying or disease-modifying or DMARD* or biologic* or therap* or treatment* or anti-rheumatic or anti rheumatic or TNF or tumor necrosis factor alpha or tumour necrosis factor alpha or TNF-alpha or TNF inhibitor* or TNF blocker* or interleukin 1 or IL-1 or monoclonal antibod* or costimulation blocker* or interleukin 6 or IL-6):ti,ab,kw 3. (methotrexate or sulfasalazine or leflunomide or hydroxychloroquine or chloroquine or gold or minocycline or azathioprine or ciclosporin or cyclosporine or penicillamine or cyclophosphamide or etanercept or infliximab or adalimumab or certolizumab* or golimumab or anakinra or rituximab or abatacept or tocilizumab):ti,ab,kw 4. (economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or value for money or budget*):ti,ab,kw 5. (#2 OR #3) 6. (#1 AND #4 AND #5)
CINAHL	<ol style="list-style-type: none"> 1. (MH "Arthritis, Rheumatoid") 2. TI (disease modifying or disease-modifying or DMARD* or biologic* or therap* or treatment* or anti-rheumatic or anti rheumatic or TNF or tumor necrosis factor alpha or tumour necrosis factor alpha or TNF-alpha or TNF inhibitor* or TNF blocker* or interleukin 1 or IL-1 or monoclonal antibod* or costimulation blocker* or interleukin 6 or IL-6) OR AB (disease modifying or disease-modifying or DMARD* or biologic* or therap* or treatment* or anti-rheumatic or anti rheumatic or TNF or tumor necrosis factor alpha or tumour necrosis factor alpha or TNF-alpha or TNF inhibitor* or TNF blocker* or interleukin 1 or IL-1 or monoclonal antibod* or costimulation blocker* or interleukin 6 or IL-6) 3. TI (methotrexate or sulfasalazine or leflunomide or hydroxychloroquine or chloroquine or gold or minocycline or azathioprine or ciclosporin or cyclosporine or penicillamine or cyclophosphamide or etanercept or infliximab or adalimumab or certolizumab* or golimumab or anakinra or rituximab or abatacept or tocilizumab) OR AB (methotrexate or sulfasalazine or leflunomide or hydroxychloroquine or chloroquine or gold of minocycline or azathioprine or ciclosporin or cyclosporine or penicillamine or cyclophosphamide or etanercept or infliximab or adalimumab or certolizumab* or golimumab or anakinra or rituximab or abatacept or tocilizumab) 4. (MH "Economics") OR (MH "Economics, Dental") OR (MH "Economics, Pharmaceutical") OR (MH "Economic Value of Life") OR "Economic" OR (MH "Economic Aspects

EMBASE

of illness")

5. TI (economic* or cost or costs or costly or costing or price or prices or pricing or pharmaco-economic* or pharmaco-economic* or value for money or budget*) OR AB (economic* or cost or costs or costly or costing or price or prices or pricing or pharmaco-economic* or pharmaco-economic* or value for money or budget*)

6. S2 OR S3

7. S4 OR S5

8. S1 AND S6 AND S7

1. exp arthritis, rheumatoid/

2. (disease modifying or disease-modifying or DMARD\$ or biologic\$ or therap\$ or treatment\$ or anti-rheumatic or anti rheumatic or TNF or tumor necrosis factor alpha or tumour necrosis factor alpha or TNF-alpha or TNF inhibitor\$ or TNF blocker\$ or interleukin 1 or IL-1 or monoclonal antibod\$ or costimulation blocker\$ or interleukin 6 or IL-6).tw.

3. (methotrexate or sulfasalazine or leflunomide or hydroxychloroquine or chloroquine or gold or minocycline or azathioprine or ciclosporin or cyclosporine or penicillamine or cyclophosphamide or etanercept or infliximab or adalimumab or certolizumab\$ or golimumab or anakinra or rituximab or abatacept or tocilizumab).tw.

4. Economics/

5. exp "Costs and Cost Analysis"/

6. Economics, Dental/

7. exp Economics, Hospital/

8. economics, medical/

9. economics, nursing/

10. economics, pharmaceutical/

11. (economic\$ or cost or costs or costly or costing or price or prices or pricing or pharmaco-economic\$).ti,ab.

12. (expenditure\$ not energy).ti,ab.

13. value for money.ti,ab.

14. budget\$.ti,ab.

15. 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14

16. ((energy or oxygen) adj cost).ti,ab.

17. (metabolic adj cost).ti,ab.

18. ((energy or oxygen) adj expenditure).ti,ab.

19. 16 or 17 or 18

20. 15 not 19

21. letter.pt.

22. editorial.pt.

23. historical article.pt.

24. 21 or 22 or 23

25. 20 not 24
26. Animals/
27. Humans/
28. 26 not (26 and 27)
29. 25 not 28
30. 2 or 3
31. 1 and 29 and 30
- MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R)
1. exp arthritis, rheumatoid/
 2. (disease modifying or disease-modifying or DMARD\$ or biologic\$ or therap\$ or treatment\$ or anti-rheumatic or anti rheumatic or TNF or tumor necrosis factor alpha or tumour necrosis factor alpha or TNF-alpha or TNF inhibitor\$ or TNF blocker\$ or interleukin 1 or IL-1 or monoclonal antibod\$ or costimulation blocker\$ or interleukin 6 or IL-6).tw.
 3. (methotrexate or sulfasalazine or leflunomide or hydroxychloroquine or chloroquine or gold or minocycline or azathioprine or ciclosporin or cyclosporine or penicillamine or cyclophosphamide or etanercept or infliximab or adalimumab or certolizumab\$ or golimumab or anakinra or rituximab or abatacept or tocilizumab).tw.
 4. Economics/
 5. exp "Costs and Cost Analysis"/
 6. Economics, Dental/
 7. exp Economics, Hospital/
 8. economics, medical/
 9. economics, nursing/
 10. economics, pharmaceutical/
 11. (economic\$ or cost or costs or costly or costing or price or prices or pricing or pharmaco-economic\$).ti,ab.
 12. (expenditure\$ not energy).ti,ab.
 13. value for money.ti,ab.
 14. budget\$.ti,ab.
 15. 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14
 16. ((energy or oxygen) adj cost).ti,ab.
 17. (metabolic adj cost).ti,ab.
 18. ((energy or oxygen) adj expenditure).ti,ab.
 19. 16 or 17 or 18
 20. 15 not 19
 21. letter.pt.
 22. editorial.pt.
 23. historical article.pt.

- 24. 21 or 22 or 23
- 25. 20 not 24
- 26. Animals/
- 27. Humans/
- 28. 26 not (26 and 27)
- 29. 25 not 28
- 30. 2 or 3
- 31. 1 and 29 and 30

NHSEED/HTA
SCI WoK

MeSH DESCRIPTOR arthritis, rheumatoid EXPLODE ALL TREES IN NHSEED,HTA

- 1. Topic=(rheumatoid arthritis)
 - 2. Topic=((disease modifying or disease-modifying or DMARD* or biologic* or therap* or treatment* or anti-rheumatic or anti rheumatic or TNF or tumor necrosis factor alpha or tumour necrosis factor alpha or TNF-alpha or TNF inhibitor* or TNF blocker* or interleukin 1 or IL-1 or monoclonal antibod* or costimulation blocker* or interleukin 6 or IL-6))
 - 3. Topic=((methotrexate or sulfasalazine or leflunomide or hydroxychloroquine or chloroquine or gold or minocycline or azathioprine or ciclosporin or cyclosporine or penicillamine or cyclophosphamide or etanercept or infliximab or adalimumab or certolizumab* or golimumab or anakinra or rituximab or abatacept or tocilizumab))
 - 4. Topic=((economic* or cost or costs or costly or costing or price or prices or pricing or pharmaco-economic* or pharmaco-economic* or value for money or budget*))
 - 5. #2 OR #3
 - 6. #1 AND #4 AND #5
-

APPENDIX A.3: EXCLUDED ARTICLES

Reference	Reason for exclusion
Agarwal, Sukhpreet V. and Tiwari, Pramit Treatment and monitoring costs in rheumatoid arthritis: Preliminary results from an Indian setting Indian Journal of Pharmaceutical Sciences 2007, 69(2):226-231.	Not a comparative analysis
Agnew-Blais, J. C., Coblyn, J. S., Katz, J. N., Anderson, R. J., Mehta, J., and Solomon, D. H. Measuring quality of care for rheumatic diseases using an electronic medical record Annals of the Rheumatic Diseases 2009, 68(5):May	No long-term evaluation
Allaart, C. F., Breedveld, F. C., and Dijkmans, B. A. C. Treatment of recent-onset rheumatoid arthritis: lessons from the BeSt Study Journal of Rheumatology 2-11-2007, 34:25-33.	No long term costs
Anderson, R. B., Needleman, R. D., Gatter, R. A., Andrews, R. P., and Scarola, J. A. Patient outcome following inpatient vs outpatient treatment of rheumatoid arthritis Journal of Rheumatology 1988, 15(4):556-560.	No treatment specific long term evaluation
Andrews, G., Simonella, L., Lapsley, H., Sanderson, K., and March, L. Evidence-based medicine is affordable: The cost-effectiveness of current compared with optimal treatment in rheumatoid and osteoarthritis Journal of Rheumatology 2006, 33(4):671-680.	Not a comparative analysis
Anon. Patients with rheumatoid arthritis survive longer if treated with methotrexate Pharmaceutical Journal 2002, 268(7192):06	News article
Anon. Treatment of rheumatoid arthritis: unknown long-term effects. [Review] [71 refs] Prescrire International 2001, 10(52):55-61.	Article
Arce-Salinas, C. A., Vargas, J., Ivanova, E., and Villasenor-Ovies, P. Estimation of rheumatoid arthritis (RA) treatment costs in Mexico. Comparison between biologic agents and the ticora strategy using a markov's model Journal of Rheumatology 2006, 33(2):422-422.	Abstract
ArizaAriza, R., MestanzaPeralta, M., and Cardiel, M. H. Direct costs of medical attention to Mexican patients with rheumatoid arthritis in a tertiary care center Clinical and Experimental Rheumatology 1997, 15(1):75-78.	No long term costs
Arshad, A. and Sulaiman, W. Optimizing the use of traditional DMARD in RA: 2781Getting the most out of what we can afford! APLAR Journal of Rheumatology 2007, 10(1):April	Review
Bagust, A., Boland, A., Hockenhull, J., Fleeman, N., Greenhalgh, J., Dundar, Y., Proudlove, C., Kennedy, T., Moots, R., Williamson, P., and Dickson, R. Rituximab for the treatment of rheumatoid arthritis Health technology assessment (Winchester, England) 2009, 13(pp 23-29):Sep	NICE ERG report
Bansback, N. J., Regier, D. A., Ara, R., Brennan, A., Shojania, K., Esdaile, J. M., Anis, A. H., and Marra, C. A. An overview of economic evaluations for drugs used in rheumatoid arthritis - Focus on tumour necrosis factor-alpha antagonists Drugs 2005, 65(4):473-496.	Review
Bansback, N., Ara, R., Karnon, J., and Anis, A. Economic evaluations in rheumatoid arthritis Pharmacoeconomics 2008, 26(5):395-408.	Review
Bansback, N., Brennan, A., and Anis, A. H. A pharmaco-economic review of adalimumab in the treatment of rheumatoid arthritis Expert Review of Pharmacoeconomics and Outcomes Research 2005, 5(5):October	Review
Bansback, N., Marra, C. A., Finckh, A., and Anis, A. The economics of treatment in early rheumatoid arthritis Best Practice and Research: Clinical Rheumatology 2009, 23(1):February	Review
Bansback, N., and Marra, C. A. Now That We Know What's BeSt, What Is Good Value for the Money? Arthritis & Rheumatism-Arthritis Care & Research 2009, 61(3):289-290.	Editorial
Bansback, Nick, Marra, Carlo A., Finckh, Axel, and Anis, Aslam The economics of treatment in early rheumatoid arthritis Best practice & research.Clinical rheumatology 2009, 23(1):83-92.	No de novo analysis

Barra, L., Pope, J. E., and Payne, M. Real-world anti-tumor necrosis factor treatment in rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis: cost-effectiveness based on number needed to treat to improve health assessment questionnaire (Provisional abstract) <i>Journal of Rheumatology</i> 2009, 36:1421-1428.	Not a comparative analysis
Benucci, M., Saviola, G., Manfredi, M., Sarzi-Puttini, P., and Atzeni, F. Cost effectiveness analysis of disease-modifying antirheumatic drugs in rheumatoid arthritis. A systematic review literature <i>International journal of rheumatology</i> 2011, 2011	Review
Beresniak, Ariel, Gossec, Laure, Goupille, Philippe, Saraux, Alain, Bamberger, Marion, Bregman, Bruno, and Dupont, Danielle Direct Cost-Modeling of Rheumatoid Arthritis According to Disease Activity Categories in France <i>Journal of Rheumatology</i> 2011, 38(3):439-445.	No long term costs
Blomqvist, P., Feltelius, N., Ekblom, A., and Klareskog, L. Rheumatoid arthritis in Sweden. Drug prescriptions, costs, and adverse drug reactions <i>Journal of Rheumatology</i> 2000, 27(5):1171-1177.	No long term costs
Boonen, A. and Severens, J. L. Evaluating Treatment Strategies in Patients with Early Rheumatoid Arthritis: Merging Modeling Theory with Clinical Practice <i>Journal of Rheumatology</i> 2011, 38(8):1538-1540.	Editorial
Bullano, M. F., McNeeley, B. J., Yu, Y. F., Quimbo, R., Burawski, L. P., Yu, E. B., and Woolley, J. M. Health economics. Comparison of costs associated with the use of etanercept, infliximab, and adalimumab for the treatment of rheumatoid arthritis <i>Managed care interface</i> 2006, 19(9):47-53.	No long term costs
Colombo, G., Muzio, A., and Longhi, A. Economic evaluation of Infliximab (Remicade) vs Etanercept (Enbrel) in the treatment of rheumatoid arthritis <i>Farmeconomia e Percorsi Terapeutici</i> 2003, 4(2):77-86.	Not in English
Crespo, Carlos, Brosa, Max, Galvan, Jordi, Carbonell, Jordi, Maymo, Jordi, Marengo, Jose Luis, Del Pino-Montes, Javier, Alonso, Alberto, and Rodriguez, Carlos [Pharmacoeconomic analysis of Metoject() in the treatment of rheumatoid arthritis in Spain] <i>Reumatologia clinica</i> 2010, 6(4):203-211.	Not in English
Doan, Q. V., Chiou, C. F., and Dubois, R. W. Review of eight pharmacoeconomic studies of the value of biologic DMARDs (adalimumab, etanercept, and infliximab) in the management of rheumatoid arthritis <i>Journal of managed care pharmacy : JMCP</i> 2006, 12(7):Sep	Review
Drummond, M. Pharmacoeconomics: friend or foe? <i>Annals of the Rheumatic Diseases</i> 2006, 65 Suppl 3:iii44-iii47.	No de novo analysis
Emery, P. Review of health economics modelling in rheumatoid arthritis <i>Pharmacoeconomics</i> 2004, 22(2):55-69.	Review
Ferraz, M. B., Maetzel, A., and Bombardier, C. A summary of economic evaluations published in the field of rheumatology and related disciplines <i>Arthritis and Rheumatism</i> 1997, 40(9):1587-1593.	Review
Finckh, A., Bansback, N., and Liang, M. H. Cost-effectiveness of biologics in early rheumatoid arthritis <i>Annals of Internal Medicine</i> 2-3-2010, 152(5):333-334.	Letter
Fleurence, Rachael and Spackman, Eldon Cost-effectiveness of biologic agents for treatment of autoimmune disorders: Structured review of the literature <i>Journal of Rheumatology</i> 2006, 33(11):2124-2131.	Review
Fries, J. F. Safety, cost and effectiveness issues with disease modifying anti-rheumatic drugs in rheumatoid arthritis <i>Annals of the Rheumatic Diseases</i> 1999, 58:86-89.	Review
Gabriel, S. E., Coyle, D., and Moreland, L. W. A clinical and economic review of disease-modifying antirheumatic drugs <i>Pharmacoeconomics</i> 2001, 19(7):715-728.	Review
Gabriel, S. E., Crowson, C. S., Luthra, H. S., Wagner, J. L., and O'Fallon, W. M. Modeling the lifetime costs of rheumatoid arthritis <i>Journal of Rheumatology</i> 1999, 26(6):1269-1274.	Not a comparative analysis
Gabrielle, Pham, Ba', Machado, M. +, Ieraci, Luciano, Witteman, William, Bombardier, Claire, and Krahn, Murray Cost-effectiveness of biologic response modifiers compared to disease-modifying antirheumatic drugs for rheumatoid arthritis: a systematic review <i>Arthritis Care & Research (2151464X)</i> 2011, 63(1):65-78.	Review
Hoving, J., Bartelds, G., Sluiter, J., Sadiraj, K., Groot, I., Lems, W., Dijkmans, B. A. C., Wijbrandts, C. A., Tak, P. P., Nurmohamed, M. T., Voskuyl, A. E., and Frings-Dresen, M. H. W. Perceived work ability, quality of life, and fatigue in patients with rheumatoid arthritis after a 6-month course of TNF inhibitors: prospective intervention study and	Not a comparative analysis

partial economic evaluation *Scandinavian Journal of Rheumatology* 2009, 38(4):246-250.

Hulsemann, J. L., Ruof, J., Zeidler, H., and Mittendorf, T. Costs in rheumatology: results and lessons learned from the 'Hannover Costing Study' *Rheumatology International* 2006, 26(8):704-711. No long term costs

Kavanaugh, A. The pharmacoeconomics of newer therapeutics for rheumatic diseases *Rheumatic Disease Clinics of North America* 2006, 32(1):45-+. Review

Kobelt, G. and Joensson, B. The burden of rheumatoid arthritis and access to treatment: outcome and cost-utility of treatments *European Journal of Health Economics* 2008, 8:S95-S106. Review

Kobelt, G., Jonsson, L., Lindgren, P., Young, A., and Eberhardt, K. Modeling the progression of rheumatoid arthritis - A two-country model to estimate costs and consequences of rheumatoid arthritis *Arthritis and Rheumatism* 2002, 46(9):2310-2319. Not a comparative analysis

Kobelt, G., Lindgren, P., and Geborek, P. Costs and outcomes for patients with rheumatoid arthritis treated with biological drugs in Sweden: a model based on registry data *Scandinavian Journal of Rheumatology* 2009, 38(6):409-418. Not a comparative analysis

Kobelt, G., Lindgren, P., Lindroth, Y., Jacobson, L., and Eberhardt, K. Modelling the effect of function and disease activity on costs and quality of life in rheumatoid arthritis *Rheumatology* 2005, 44(9):1169-1175. Not a comparative analysis

Laas, K., Peltomaa, R., Kautiainen, H., Puolakka, K., and Leirisalo-Repo, M. Pharmacoeconomic study of patients with chronic inflammatory joint disease before and during infliximab treatment *Annals of the Rheumatic Diseases* 2006, 65(7):924-928. Not a comparative analysis

Lyseng-Williams, K. A. and Foster, R. H. Infliximab - A pharmacoeconomic review of its use in rheumatoid arthritis *Pharmacoeconomics* 2004, 22(2):107-132. Review

Lyseng-Williamson, K. A. and Plosker, G. L. Etanercept - A pharmacoeconomic review of its use in rheumatoid arthritis *Pharmacoeconomics* 2004, 22(16):1071-1095. Review

Maetzel, A. Cost-effectiveness estimates reported for tumor necrosis factor blocking agents in rheumatoid arthritis refractory to methotrexate--a brief summary *The Journal of rheumatology.Supplement* 2005, 72(pp 51-53):Jan Review

Maetzel, A., Ferraz, M. B., and Bombardier, C. A review of cost effectiveness analyses in rheumatology and related disciplines *Current Opinion in Rheumatology* 1998, 10(2):1998 Review

McLeod, C., Bagust, A., Boland, A., Dagenais, P., Dickson, R., Dundar, Y., Hill, R., Jones, A., Mota, R. Mujica, and Walley, T. Adalimumab, etanercept and infliximab for the treatment of ankylosing spondylitis: A systematic review and economic evaluation *Health Technology Assessment* 2007, 11(28):IX-+. Not Rheumatoid Arthritis

Merkesdal, S. and Ruof, J. Current aspects of cost effectiveness of TNF-alpha blocking agents in patients with rheumatoid arthritis *Zeitschrift fur Rheumatologie* 2002, 61:29-32. Not in English

Merkesdal, S., Ruof, J., Mittendorf, T., and Zeidler, H. Cost-effectiveness of TNF-alpha-blocking agents in the treatment of rheumatoid arthritis *Expert Opinion on Pharmacotherapy* 2004, 5(9):1881-1886. Review

Nurmohamed, M. T. and Dijkmans, B. A. C. Efficacy, tolerability and cost effectiveness of disease-modifying antirheumatic drugs and biologic agents in rheumatoid arthritis *Drugs* 2005, 65(5):661-694. Review

Ostrov, B., Robbins, L., Ferriss, J., Newman, E., Maclary, S., Ayoub, W., Harrington, T., and Perruquet, J. Improved tolerance and cost-effectiveness of subcutaneous vs oral methotrexate in rheumatoid arthritis (RA) and juvenile rheumatoid arthritis (JRA) *Arthritis and Rheumatism* 1998, 41(9 SUPPL.):S75 Abstract

Prashker, M. J. and Meenan, R. F. The Total Costs of Drug-Therapy for Rheumatoid-Arthritis - A Model-Based on Costs of Drug, Monitoring, and Toxicity *Arthritis and Rheumatism* 1995, 38(3):318-325. Cost analysis

Ruchlin, H. S., Elkin, E. B., and Paget, S. A. Assessing cost-effectiveness analyses in rheumatoid arthritis and osteoarthritis *Arthritis Care and Research* 1997, 10(6):413-421. Review

Sany, J., Cohen, J. D., Combescore, C., Bozonnat, M. C., Roch-Bras, F., Lafon, G., and Dures, J. P. Medico-economic evaluation of infliximab in rheumatoid arthritis--prospective French study of a cohort of 635 patients monitored for two years <i>Rheumatology (Oxford, England)</i> 2009, 48(10):Oct	Not a comparative analysis
Schadlich, P. K., Zeidler, H., Zink, A., Gromnica-Ihle, E., Schneider, M., Straub, C., Brecht, J. G., and Huppertz, E. Contribution of leflunomide to the cost effectiveness of sequential DMARD therapy of rheumatoid arthritis in Germany <i>Zeitschrift fur Rheumatologie</i> 2004, 63(1):61-77.	Not in English
Schoels, Monika, Wong, John, Scott, David L., Zink, Angela, Richards, Pamela, Landewe, Robert, Smolen, Josef S., and Aletaha, Daniel Economic aspects of treatment options in rheumatoid arthritis: a systematic literature review informing the EULAR recommendations for the management of rheumatoid arthritis <i>Annals of the Rheumatic Diseases</i> 2010, 69(6):995-1003.	Review
Schulze-Koops, H., Deeg, M., Runge, C., Volmer, T., and Brecht, J. Health-economic assessment of combination therapy for rheumatoid arthritis with methotrexat and etanercept based on the TEMPO Study <i>Zeitschrift fur Rheumatologie</i> 2009, 68(10):836-841.	Not in English
Suarez-Almazor, M. E., Ortiz, Z., Lopez-Olivo, M., Moffett, M., Pak, C., Skidmore, B., Kimmel, B., Kallen, M., and Cox, V. Long-term clinical and cost-effectiveness of infliximab and etanercept for rheumatoid arthritis 2007	Summary report
Suka, M. and Yoshida, K. Cost effectiveness of leflunomide in the treatment of rheumatoid arthritis in Japan <i>Expert Review of Pharmacoeconomics and Outcomes Research</i> 2004, 4(6):December	Not a comparative analysis
Tella, M. N., Feinglass, J., and Chang, R. W. Cost-effectiveness, cost-utility, and cost-benefit studies in rheumatology: a review of the literature, 2001-2002 <i>Current Opinion in Rheumatology</i> 2003, 15(2):127-131.	Review
van den Hout, W. B., Goekoop-Ruiterman, Y. P., Allaart, C. F., Vries-Bouwstra, J. K., Hazes, J. M., Kerstens, P. J., Van, Zeven D., Hulsmans, H. M., De Jonge-Bok, J. M., De Sonnaville, P. B., Dijkmans, B. A., and Breedveld, F. C. Cost-utility analysis of treatment strategies in patients with recent-onset rheumatoid arthritis <i>Arthritis and Rheumatism (Arthritis Care and Research)</i> 2009, 61(3):291-299.	Duplicate citation
van der Velde, G., Pham, B., Machado, M., Ieraci, L., Witterman, W., Bombardier, C., and Krahn, M. Cost-Effectiveness of Biologic Response Modifiers Compared to Disease-Modifying Antirheumatic Drugs for Rheumatoid Arthritis: A Systematic Review <i>Arthritis Care & Research</i> 2011, 63(1):65-78.	Duplicate citation
van, der, V, Pham, B., Machado, M., Ieraci, L., Witterman, W., Bombardier, C., and Krahn, M. Cost-effectiveness of biologic response modifiers compared to disease-modifying antirheumatic drugs for rheumatoid arthritis: a systematic review. [Review] <i>Arthritis Care & Research</i> 2011, 63(1):65-78.	Review
Verhoeven, A., Bibo, J., Boers, M., Engel, G., Schouten, H., and van der Linden, S. Combination of step-down steroids, methotrexate, and sulphasalazine is cost-effective compared to sulphasalazine alone in early rheumatoid arthritis <i>British Journal of Rheumatology</i> 1997, 36(SUPPL. 1):181	Abstract
Virkki, L., Konttinen, Y., Peltomaa, R., Suontama, K., Saario, R., Immonen, K., Jantti, J., Tuomiranta, T., Nykanen, P., Hameenkorpi, R., Heikkila, S., Isomaki, P., and Nordstrom, D. Cost-effectiveness of infliximab in the treatment of rheumatoid arthritis in clinical practice <i>Clinical and Experimental Rheumatology</i> 2008, 26(6):1059-1066.	Not a comparative analysis
Walsh, C., Minnock, P., Slatery, C., Kennedy, N., Pang, F., Veale, D., Bresnihan, B., and FitzGerald, O. Quality of life and economic impact of switching from established infliximab therapy to adalimumab in patients with rheumatoid arthritis <i>Rheumatology</i> 2007, 46(7):1148-1152.	Not a comparative analysis
Welsing, P. M. J., Severens, J. L., and Laan, R. F. J. M. Optimistic assumptions in modelling studies have a substantial influence on the cost-effectiveness result <i>Rheumatology</i> 2003, 42(12):1574-1575.	Letter
Welsing, Paco M., Severens, Johan L., Hartman, Margriet, van Gestel, Anke M., van Riel, Piet L., and Laan, Roland F. The initial validation of a Markov model for the economic evaluation of (New) treatments for rheumatoid arthritis <i>Pharmacoeconomics</i> 2006, 24(10):1011-1020.	Not a comparative analysis
Wong, J. B. Cost-effectiveness of anti-tumor necrosis factor agents <i>Clinical and Experimental Rheumatology</i> 2004, 22(5):S65-S70.	Article

APPENDIX A.4: DATA EXTRACTION

Name	Sequencing information				Modelling information				
	If a sequence was evaluated, how many lines?	Were all relevant treatments included in each sequence?	Was the methodology used amenable to vary the sequence and comparing a full range of alternatives	Was an attempt to find a truly 'optimal' sequence undertaken?	If the analysis used modelling, what method was used?	How was initial treatment response modelled?	What determined a switch to an alternative therapy?	How were the costs and effectiveness of subsequent treatments in a sequence modelled?	Where data were not available for a treatment in a sequence, how was this accounted for?
Anis 1996 ¹³⁸	1	Unclear	N	No	Decision tree	Functional index improvement	n/a	n/a	n/a
Bansback 2005 ¹³⁹	5 (comparator bDMARD, DMARD 1,2,3, Pall Care)	Potentially, however no sequential use of bDMARDs	Yes	No	Individual level Markov model	ACR response to HAQ change	Short term withdrawal for non-response. Long term withdrawal for loss of efficacy or an AE	For sequential DMARDs, efficacy is modified by a OR for disease duration.	Adjustments based on registry data.
Barbieri 2005 ¹⁴⁰	1	No	Markov model. Unlikely	No	Markov model	HAQ score	An undescribed transition probability	n/a	n/a
Barton 2004 ¹⁴¹	Up to 11	Potentially	Potentially	No	Individual Sampling Model	HAQ score	Probability of early withdrawal and time to later withdrawal	Modelled explicitly.	Not clear
Benucci 2009 ¹⁴²	1	No. Infliximab not included	No	No	Observational analysis	HAQ score	n/a	observational data	n/a
Benucci 2011 ¹⁴³	1	No	No	No	Observational analysis	n/a	n/a	n/a	n/a

Beresniak 2011 ¹⁴⁴	2	No. IFX not included in second line.	Potentially	No. No incremental analysis performed and full range of possible sequences not modelled	Unclear	Remission or LDAS	Failure to achieve the constant probability of effectiveness while on maintenance therapy	Modelled explicitly within model time horizon and comparators. No future projections of costs and benefits	Assumed same as other tx
Brennan 2004 ¹⁴⁵	4	No	Potentially	No	Individual Sampling Model	ACR20 to HAQ	Short term withdrawal for non-response. Long term withdrawal for loss of efficacy or an AE	Explicitly	Data were available. Assumed class effect on HAQ progression
Brennan 2007 ¹⁴⁶	6 lines	Potentially, however sequential use of TNFa's only considered in scenario analysis	Potentially yes	No	Individual Sampling Model	EULAR response categories	withdrawal rates or non-response	Observational data and maintenance therapy after 6 treatments	Generalised DMARDs used based on position in sequence.
Chen 2006 ¹²⁴	Approximately 10	Almost. Sequential considered	Potentially	No	Individual Sampling Model	HAQ score	Probability of early withdrawal and time to later withdrawal	Modelled explicitly.	Not clear
Chiou 2004 ¹⁴⁷	1	No	No	No	Decision tree	ACR response	n/a	n/a	n/a
Choi 2000 ¹⁴⁸	n/a	n/a	n/a	n/a	Decision tree	ACR response	n/a	n/a	n/a
Choi 2002 ¹⁴⁹	1	No	No	No	Decision tree	ACR response	n/a	n/a	n/a

Cimmino 2011 ¹⁵⁰	2	No. IFX not included in second line.	Potentially	No. No incremental analysis performed and full range of possible sequences not modelled	Unclear	Remission or LDAS	Failure to achieve the constant probability of effectiveness while on maintenance therapy	Modelled explicitly within model time horizon and comparators. No future projections of costs and benefits	Assumed same as other tx. Expert opinion
Clark 2004 ¹⁵¹	Up to 11	Potentially	Potentially	No	Individual Sampling Model	HAQ score	Probability of early withdrawal and time to later withdrawal	Modelled explicitly.	Not clear
Coyle 2006 ¹⁵²	6	No	No	No	Markov model	ACR Response	Stopping for AE or stopping for lack of benefit	Trial data or simulated	Assumed the same as other treatments
Davies 2009 ¹²⁵	5 (unique)	No. No combination DMARDs for early RA. No full comparators	Potentially yes	No	Individual Sampling Model	ACR50	Withdrawal for AE or inefficacy	Not reported	Not reported
Diamantopoulos 2012 ¹⁵³	5 (with 6 for one sensitivity analysis)	Yes	Potentially	No	Individual Sampling Model	ACR response to HAQ change	withdrawal rates or non-response	Trial data	Assumed equal to other treatments. Registry data
Finckh 2009 ¹²⁶	Approximately 3	Alternative tapered DMARD treatments not considered	Yes	Within comparator set yes. Wider no	Individual Sampling Model	Multivariate distribution for response level (excellent, good, moderate, no)	Database withdrawal rates, and response level	Using trial evidence, although downstream past 3 TNFs not stated	
Hallinen 2010 ¹⁵⁴	Up to 4	Potentially Yes	Potentially	No	Individual level Markov model	ACR response to HAQ change	withdrawal rates or non-response	Trial data	Assumptions

Hartman 2004 ¹²⁷	1	No	No	No	Trial analysis	n/a	n/a	n/a	n/a
Jobanputra 2002 ¹⁵⁵	Up to 11	Potentially	Potentially	No	Individual Sampling Model	HAQ score	Probability of early withdrawal and time to later withdrawal	Modelled explicitly.	Not clear
Kavanaugh 1996 ¹⁵⁶	1	Yes potentially	N	No	Decision tree	Unclear	n/a	n/a	n/a
Kielhorn 2008 ¹⁵⁷	6 base case, 7 SA	Maybe	Y	N	Individual level Markov model	ACR response to HAQ change	withdrawal rates or non-response	Data where possible, however downstream cDMARDs assumed equal effectiveness	See before, equal efficacy used
Kievit 2009 ¹⁵⁸	1	Yes. Guidelines for Netherlands followed and alternative options considered.	No	No	Trial analysis	DAS	Guideline and trial protocols. DAS28 improvement in general	DAS28 response and costs of drugs	n/a
Kobelt 2002 ¹²⁸	1	No	No	No	Markov model	Observational data to develop HAQ transition matrices	Observational data	Observational data	n/a
Kobelt 2003 ¹⁵⁹	1	No	No	No	Markov model	EQ-5D	n/a	n/a	n/a
Kobelt 2004 ¹⁶⁰	1	No	No	No	Trial analysis	EQ-5D	n/a	n/a	n/a
Kobelt 2005 ¹⁶¹	1	No. No comparison to other aTNFs	No	No	Markov model	Change in functional capacity and disease activity	n/a	n/a	n/a
Kobelt 2011 ¹²⁹	3 (maybe more) not clear	No	Potentially	No	Markov model	Remission	Trial discontinuation	Trial and registry data	n/a

Korthals-de Bos 2004 ¹³⁰	n/a	n/a	n/a	n/a	n/a	n/a	n/a	rates	n/a	n/a
Lekander 2010 ¹⁶²	1	No	No	No	Markov model	HAQ and DAS28	n/a	n/a	n/a	n/a
Lindgren 2009 ¹⁶³	Open ended	No	Potentially	No	Discrete Event Simulation	HAQ score	Treatment withdrawal rate taken from trial data	Assumed same as previous line for RTX sequence (i.e. RTX has no impact on treatment effectiveness). For TNF data provided on 2nd and 3rd line from registry but not for subsequent lines	Not clearly reported	
Maetzel 2002 ¹³¹	6	Yes	Yes	No	Decision tree	ACR20 response	LoE or Toxicity	All from trial evidence	n/a	
Maetzel 2002a ¹⁶⁴	1	No	No	No	n/a	n/a	n/a	n/a	n/a	
Malottki 2011 ¹⁶⁵	6	No. No sequential biologic use	Potentially	No	Individual Sampling Model	HAQ score	Probability of early withdrawal and time to later withdrawal	Modelled explicitly.	Not clear	
Marra 2007 ¹⁶⁶	n/a	n/a	n/a	n/a	Markov model	HAQ score	n/a	n/a	n/a	
Merkesdal 2010 ¹⁶⁷	6	Potentially	Yes	No	Individual level Markov model	ACR response to HAQ change	LoE or AE	Trial data adjusted crudely	Not clear. Trial data not appropriately references	
Nuijten 2001 ¹⁶⁸	1	Potentially	No	No	Unclear	n/a	n/a	n/a	n/a	
Osiri 2007 ¹⁶⁹	1	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
Rubio-Terrés 2001 ¹⁷⁰	n/a	n/a	n/a	n/a	Unclear	n/a	n/a	n/a	n/a	

Russell 2009 ¹⁷¹	3	No. Abatacept not considered as 3rd line treatment	Yes, however data limited	No	Decision tree	Remission or LDAS	Not achieving remission	From trial data	n/a
Saraux 2010 ¹⁷²	2	No	Yes	No	Unclear	DAS	Not achieving remission	From trial data	n/a
Schadlich 2005 ¹³²	6	No	No	No	Markov model	DAS	Not achieving remission	Registry data	2nd line aTNF set equal to 1st line.
Schipper 2011 ¹³³	1	Maybe. Wide choice of DMARDs	No	No	n/a	n/a	n/a	n/a	n/a
Shini 2010 ¹⁷³	Only 1 line evaluated. Sequence included 6 therapies (bDMARD, RTX, IFX, LEF, CSA, MTX)	Alternative aTNFs and bDMARDs not included	Yes	No	Individual Sampling Model	ACR response	Lack of ACR20 response or treatment discontinuation	Trials and indirect comparisons. Some assumptions	All data accounted for, however data not position specific
Soini ¹⁷⁴ 2012	2	No	No	No	Markov model	HAQ score	Transition probabilities for toxicity and loss of efficacy	Pooled estimate of cost and effect	Not clearly reported
Spalding 2006 ¹³⁴	4	No	Probably not	No	Markov model	ACR20 response	Lack of ACR20 response or treatment discontinuation	Trial data	Assumed equal to other treatments (MTX+SSZ assumed equal to just SSZ)
Tanno 2006 ¹⁷⁵	3 (biologics merged)	No	Potentially	No	Individual Sampling Model	ACR response	Short term withdrawal for non-response.	Bolton-on costs and QALYs	as before, from another study

							Long term withdrawal for loss of efficacy or an AE		
Tosh 2011 ¹³⁵	8 (in one strategy)	Potentially yes	Trial based evaluation	No	Trial analysis	DAS	DAS > 2.4	Trial	Trial
van den Hout 2009 ¹³⁶	1	No	Unclear	No	Individual Sampling Model	ACR response or HAQ change	n/a	n/a	n/a
Vera-Llonch 2008 ¹⁷⁶	1	No	Unclear	No	Microsimulation	ACR response or HAQ change	n/a	n/a	n/a
Vera-Llonch 2008a ¹⁷⁷	1	Y	N	N	n/a	n/a	n/a	n/a	n/a
Verhoeven 1998 ¹³⁷	1+ (unclear tx regimen)	Potentially. Not clear	Potentially yes	No	Individual Sampling Model	ACR response to HAQ change	LoE or AE	Not clear. HAQ progression?	Not clear
Wailoo 2008 ¹⁷⁸	3 at most	No	No	No	Markov model	Not clear	Not clear	Trial data	n/a
Welsing 2004 ¹⁷⁹	1	No	No	No	Markov model	Trial data	n/a	n/a	n/a

APPENDIX A.5: DRUMMOND CHECKLIST

	Drummond Checklist (Y=Yes, N=No, C=Can't tell, n/a=Not applicable)									
	1	2	3	4	5	6	7	8	9	10
	Was a well-defined question posted in an answerable form?	Was a comprehensive description of the competing alternatives given?	Was there evidence that the programme's effectiveness had been established?	Were all important and relevant costs and consequences for each alternative identified?	Were costs and consequences measured accurately in appropriate physical units?	Were costs and consequences valued credibly?	Were costs and consequences adjusted for differential timing?	Was an incremental analysis of costs and consequences of alternatives performed	Was allowance made for the uncertainty in the estimates of costs and consequences?	Did the presentation and discussion of study results include all issues of concern to users?
Anis 1996 ¹³⁸	Y	Y	C	Y	Y	Y	N	N	Y	N
Bansback 2005 ¹³⁹	Y	Y	Y	Y	Y	Y	Y	N	Y	N
Barbieri 2005 ¹⁴⁰	Y	Y	Y	Y	Y	Y	Y	Y	N	N
Barton 2004 ¹⁴¹	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Benucci 2009 ¹⁴²	Y	Y	Y	N	N	N	N	N	N	N
Benucci 2011 ¹⁴³	Y	N	N	C	C	C	N	Y	N	N
Beresniak 2011 ¹⁴⁴	Y	Y	Y	N	N	Y	C	N	Y	N
Brennan 2004 ¹⁴⁵	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Brennan 2007 ¹⁴⁶	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Chen 2006 ¹²⁴	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Chiou 2004 ¹⁴⁷	Y	Y	Y	Y	Y	Y	n/a	Y	N	Y
Choi 2000 ¹⁴⁸	Y	Y	N	N	Y	Y	n/a	Y	N	N
Choi 2002 ¹⁴⁹	Y	Y	Y	Y	Y	Y	n/a	Y	N	Y

	1	2	3	4	5	6	7	8	9	10
Cimmino 2011 ¹⁵⁰	Y	Y	Y	N	N	Y	C	N	Y	N
Clark 2004 ¹⁵¹	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Coyle 2006 ¹⁵²	Y	Y	Y	Y	Y	Y	Y	N	Y	N
Davies 2009 ¹²⁵	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Diamantopoulos 2012 ¹⁵³	Y	Y	Y	Y	Y	Y	Y	N	Y	N
Finckh 2009 ¹²⁶	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Hallinen 2010 ¹⁵⁴	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Hartman 2004 ¹²⁷	Y	Y	Y	N	Y	Y	N	N	Y	Y
Jobanputra 2002 ¹⁵⁵	Y	Y	N	Y	Y	Y	Y	Y	Y	N
Kavanaugh 1996 ¹⁵⁶	Y	N	Y	Y	Y	Y	N	N	Y	N
Kielhorn 2008 ¹⁵⁷	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Kievit 2009 ¹⁵⁸	Y	Y	Y	N	Y	Y	N	N	Y	N
Kobelt 2002 ¹²⁸	Y	Y	Y	N	Y	Y	Y	Y	Y	Y
Kobelt 2003 ¹⁵⁹	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Kobelt 2004 ¹⁶⁰	Y	N	C	Y	Y	Y	Y	Y	Y	N
Kobelt 2005 ¹⁶¹	Y	Y	Y	Y	Y	Y	Y	N	Y	N
Kobelt 2011 ¹²⁹	Y	Y	N	Y	Y	Y	Y	Y	Y	Y
Korthals-de Bos 2004 ¹³⁰	Y	Y	N	N	Y	Y	n/a	Y	Y	N
Lekander 2010 ¹⁶²	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lindgren 2009 ¹⁶³	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Maetzel 2002 ¹³¹	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Maetzel 2002a ¹⁶⁴	Y	Y	Y	Y	Y	Y	Y	N	Y	N

	1	2	3	4	5	6	7	8	9	10
Malottki 2011 ¹⁶⁵	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Marra 2007 ¹⁶⁶	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Merkesdal 2010 ¹⁶⁷	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Nuijten 2001 ¹⁶⁸	Y	Y	N	Y	Y	Y	Y	N	Y	N
Osiri 2007 ¹⁶⁹	Y	Y	N	N	Y	Y	n/a	N	N	N
Rubio-Terrés 2001 ¹⁷⁰	Y	Y	N	N	Y	Y	n/a	N	N	N
Russell 2009 ¹⁷¹	Y	Y	C	Y	Y	Y	C	N	Y	N
Saraux 2010 ¹⁷²	Y	Y	Y	Y	Y	Y	C	N	N	N
Schadlich 2005 ¹³²	Y	Y	Y	N	Y	Y	N	N	Y	Y
Schipper 2011 ¹³³	Y	Y	N	Y	N	Y	Y	N	Y	N
Shini 2010 ¹⁷³	Y	Y	N	N	N	Y	N	N	N	N
Soini 2012 ¹⁷⁴	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
Spalding 2006 ¹³⁴	Y	Y	N	Y	Y	Y	Y	N	N	N
Tanno 2006 ¹⁷⁵	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Tosh 2011 ¹³⁵	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
van den Hout 2009 ¹³⁶	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Vera-Llonch 2008 ¹⁷⁶	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Vera-Llonch 2008a ¹⁷⁷	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Verhoeven 1998 ¹³⁷	Y	Y	Y	N	Y	Y	n/a	N	Y	Y
Wailoo 2008 ¹⁷⁸	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Welsing 2004 ¹⁷⁹	Y	N	Y	N	Y	Y	Y	Y	Y	Y
Wong 2002 ¹⁸⁰	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

APPENDIX B: METHODOLOGICAL REVIEW

APPENDIX B.1: ARTICLES EXCLUDED AT INITIAL SIFT STAGE

Reference	Reason for exclusion
1 Ali M, Vukovic V, Sahir MH, Fontanella G. Energy analysis of chilled water system configurations using simulation-based optimization. <i>Energy and Buildings</i> . 2013;59:111-22.	Unrelated problem type
2 Almeder C, Preusser M, Hartl RF. Simulation and optimization of supply chains: alternative or complementary approaches? <i>Or Spectrum</i> . 2009;31(1):95-119.	Unrelated problem type (mixed integer)
3 Arena C, Mazzola MR, Scordo G. A simulation/optimization model for selecting infrastructure alternatives in complex water resource systems. <i>Water Science and Technology</i> . 2010;61(12):3050-60.	Not a combinatorial problem
4 Azadeh A, Moghaddam M, Asadzadeh SM, Negahban A. An integrated fuzzy simulation-fuzzy data envelopment analysis algorithm for job-shop layout optimization: The case of injection process with ambiguous data. <i>European Journal of Operational Research</i> . 2011;214(3):768-79.	Unrelated problem type – production efficiency
5 Bachelet B, Yon L. Model enhancement: Improving theoretical optimization with simulation. <i>Simulation Modelling Practice and Theory</i> . 2007;15(6):703-15.	Not optimisation of a simulation model
6 Bhatnagar S, Mishra VK, Hemachandra N. Stochastic Algorithms for Discrete Parameter Simulation Optimization. <i>Ieee Transactions on Automation Science and Engineering</i> . 2011;8(4):780-93.	Looking to optimise the performance of a simulation
7 Chen Y, Mockus L, Orcun S, Reklaitis GV. Simulation-optimization approach to clinical trial supply chain management with demand scenario forecast. <i>Computers & Chemical Engineering</i> . 2012;40:82-96.	Not a combinatorial problem
8 Fu MC. Optimization for simulation: Theory vs. practice. <i>Inform Journal on Computing</i> . 2002;14(3):192-215.	Not a combinatorial problem
9 Fu Z, Mo J, Chen L, Chen W. Using genetic algorithm-back propagation neural network prediction and finite-element model simulation to optimize the process of multiple-step incremental air-bending forming of sheet metal. <i>Materials & Design</i> . 2010;31(1):267-77.	Not a combinatorial problem
10 Geyik F, Dosdogru AT. Process plan and part routing optimization in a dynamic flexible job shop scheduling environment: an optimization via simulation approach. <i>Neural Computing & Applications</i> . 2013;23(6):1631-41.	Not a combinatorial problem
11 Gosavi A, Ozkaya E, Kahraman AF. Simulation optimization for revenue management of airlines with cancellations and overbooking. <i>Or Spectrum</i> . 2007;29(1):21-38.	Not a combinatorial problem
12 Gulyanitskii LF, Koshlai LB, Sergienko IV. CONVERGENCE OF A SIMULATION METHOD FOR SOLUTION OF COMBINATORIAL OPTIMIZATION PROBLEMS. <i>Cybernetics and Systems Analysis</i> . 1993;29(3):445-9.	Not concerned with stochastic/simulation based problem
13 Gupta A, Evans GW, Heragu SS. Simulation and optimization modeling for drive-through mass vaccination - A generalized approach. <i>Simulation Modelling Practice and Theory</i> . 2013;37:99-106.	Only an application of a ‘off the shelf’ solution
14 Hong JH, Seo K-M, Kim TG. Simulation-based optimization for design parameter exploration in hybrid system: a defense system example. <i>Simulation-Transactions of the Society for Modeling and Simulation International</i> . 2013;89(3):362-80.	Mixed (integer and continuous) problem

- 15 Horng S-C, Lin S-Y, Lee H, Chen C-H. Memetic Algorithm for Real-Time Combinatorial Stochastic Simulation Optimization Problems With Performance Analysis. *Ieee Transactions on Cybernetics*. 2013;43(5):1495-509. Off-line and on-line problem
- 16 Huang Y, Li YP, Chen X, Bao AM, Zhou M. Simulation-based optimization method for water resources management in Tarim River Basin, China. In: Yang Z, Chen B, editors. *International Conference on Ecological Informatics and Ecosystem Conservation*. *Procedia Environmental Sciences*. 2010. p. 1451-60. Not a combinatorial problem
- 17 Iassinovski S, Artiba A, Bachelet V, Riane F. Integration of simulation and optimization for solving complex decision making problems. *International Journal of Production Economics*. 2003;85(1):3-10. Not specific for combinatorial problems
- 18 Jeong KY. Conceptual frame for development of optimized simulation-based scheduling systems. *Expert Systems with Applications*. 2000;18(4):299-306. Not a combinatorial problem
- 19 Keskin BB, Melouk SH, Meyer IL. A simulation-optimization approach for integrated sourcing and inventory decisions. *Computers & Operations Research*. 2010;37(9):1648-61. Not a combinatorial problem
- 20 Kleijnen JPC, Rubinstein RY. Optimization and sensitivity analysis of computer simulation models by the score function method. *European Journal of Operational Research*. 1996;88(3):413-27. Not a combinatorial problem
- 21 Klemmt A, Horn S, Weigert G, Wolter K-J. Simulation-based optimization vs. mathematical programming: A hybrid approach for optimizing scheduling problems. *Robotics and Computer-Integrated Manufacturing*. 2009;25(6):917-25. Not a combinatorial problem
- 22 Latorre J-I, Jimenez E. Simulation-based optimization of discrete event systems with alternative structural configurations using distributed computation and the Petri net paradigm. *Simulation-Transactions of the Society for Modeling and Simulation International*. 2013;89(11):1310-34. A method for parallel programming for DES optimisation
- 23 Li J, Sava A, Xie X. Simulation-Based Discrete Optimization of Stochastic Discrete Event Systems Subject to Non Closed-Form Constraints. *Ieee Transactions on Automatic Control*. 2009;54(12):2900-4. Not a combinatorial problem
- 24 Lu M, Lam H-C, Dai F. Resource-constrained critical path analysis based on discrete event simulation and particle swarm optimization. *Automation in Construction*. 2008;17(6):670-81. Not a combinatorial problem
- 25 Lu M, Wu D-p, Zhang J-p. A particle swarm optimization-based approach to tackling simulation optimization of stochastic, large-scale and complex systems. In: Yeung DS, Liu ZQ, Wang XZ, Yan H, editors. *Advances in Machine Learning and Cybernetics. Lecture Notes in Artificial Intelligence*. 39302006. p. 528-37. Not a combinatorial problem
- 26 Mahdavi I, Shirazi B, Sahebjamnia N. Development of a simulation-based optimisation for controlling operation allocation and material handling equipment selection in FMS. *International Journal of Production Research*. 2011;49(23):6981-7005. Multi-objective optimisation
- 27 Mardan N, Klahr R. Combining optimisation and simulation in an energy systems analysis of a Swedish iron foundry. *Energy*. 2012;44(1):410-9. Not relevant
- 28 Marseguerra M, Zio E, Podofilini L. Condition-based maintenance optimization by means of genetic algorithms and Monte Carlo simulation. *Reliability Engineering & System Safety*. 2002;77(2):151-65. Not a combinatorial problem
- 29 Matott LS, Tolson BA, Asadzadeh M. A benchmarking framework for simulation-based optimization of environmental models. *Environmental Modelling & Software*. 2012;35:19-30. Not a combinatorial problem
- 30 Medaglia AL, Fang SC, Nuttle HLW. Fuzzy controlled simulation optimization. *Fuzzy Sets and Systems*. 2002;127(1):65-84. Not identifying optimal or near-optimal solutions
- 31 Mele FD, Guillen G, Espuna A, Puigjaner L. A simulation-based optimization framework for parameter optimization of supply-chain networks. *Industrial & Engineering Chemistry Research*. 2006;45(9):3133-48. Not a combinatorial problem

- 32 Napalkova L, Merkurjeva G. MULTI-OBJECTIVE STOCHASTIC SIMULATION-BASED OPTIMISATION APPLIED TO SUPPLY CHAIN PLANNING. *Technological and Economic Development of Economy*. 2012;18(1):132-48. Multi-objective optimisation
- 33 Narciso M, Piera MA, Guasch A. A Methodology for Solving Logistic Optimization Problems through Simulation. *Simulation-Transactions of the Society for Modeling and Simulation International*. 2010;86(5-6):369-89. Not relevant
- 34 Nielsen AL, Hilwig H, Kisoos N, Teelucksingh S. Discrete event simulation as a tool in optimization of a professional complex adaptive system. In: Andersen SK, Klein GO, Schulz S, Aarts J, Mazzoleni MC, editors. *Ehealth Beyond the Horizon - Get It There. Studies in Health Technology and Informatics*. 1362008. p. 247-52. Simple optimisation via enumeration
- 35 Park C, Telci IT, Kim S-H, Aral MM. Designing an optimal water quality monitoring network for river systems using constrained discrete optimization via simulation. *Engineering Optimization*. 2014;46(1):107-29. Not a combinatorial problem
- 36 Ramwadhoebe S, Buskens E, Sakkers RJB, Stahl JE. A tutorial on discrete-event simulation for health policy design and decision making: Optimizing pediatric ultrasound screening for hip dysplasia as an illustration. *Health Policy*. 2009;93(2-3):143-50. Not a combinatorial problem
- 37 Rosen SL, Harmonosky CM, Trabant MT. A simulation optimization method that considers uncertainty and multiple performance measures. *European Journal of Operational Research*. 2007;181(1):315-30. Multi-objective optimisation
- 38 Roux O, Jamali MA, Kadi DA, Chatelet E. Development of simulation and optimization platform to analyse maintenance policies performances for manufacturing systems. *International Journal of Computer Integrated Manufacturing*. 2008;21(4):407-14. Not relevant
- 39 Rubinstein RY. Optimization of computer simulation models with rare events. *European Journal of Operational Research*. 1997;99(1):89-112. Not a combinatorial problem
- 40 Salama A, Nehring M, Greberg J. Operating value optimisation using simulation and mixed integer programming. *International Journal of Mining Reclamation and Environment*. 2014;28(1):25-46. Not a combinatorial problem
- 41 Sanchez D, Amodeo L, Prins C. Meta-heuristic Approaches for Multi-objective Simulation-based Optimization in Supply Chain Inventory Management. In: Benyoucef L, Grabot B, editors. *Artificial Intelligence Techniques for Networked Manufacturing Enterprises Management. Springer Series in Advanced Manufacturing*2010. p. 249-69. Multi-objective optimisation
- 42 Waanders BGvB, Carnes BR. Optimization under adaptive error control for finite element based simulations. *Computational Mechanics*. 2011;47(1):49-63. Not relevant
- 43 Wu F, Dantan J-Y, Etienne A, Siadat A, Martin P. Improved algorithm for tolerance allocation based on Monte Carlo simulation and discrete optimization. *Computers & Industrial Engineering*. 2009;56(4):1402-13. Not a combinatorial problem
- 44 Yang T. An evolutionary simulation-optimization approach in solving parallel-machine scheduling problems - A case study. *Computers & Industrial Engineering*. 2009;56(3):1126-36. Not a combinatorial problem
- 45 Yang Y, Blum RS. Sensor Placement in Gaussian Random Field Via Discrete Simulation Optimization. *Ieee Signal Processing Letters*. 2008;15:729-32. Not a combinatorial problem
- 46 Yeo IY, Guldmann JM. Global spatial optimization with hydrological systems simulation: application to land-use allocation and peak runoff minimization. *Hydrology and Earth System Sciences*. 2010;14(2):325-38. Not a combinatorial problem
- 47 Yoo T, Cho H, Yuecesan E. Web Services-Based Parallel Replicated Discrete Event Simulation for Large-Scale Simulation Optimization. *Simulation-Transactions of the Society for Modeling and Simulation International*. 2009;85(7):461-75. Not relevant
- 48 Zamora-Cristales R, Boston K, Sessions J, Murphy G. Stochastic simulation and optimization of mobile chipping and transport of forest biomass from harvest residues. *Silva Fennica*. 2013;47(5). Not a combinatorial problem

- 49 Zhang R, Chiang W-C, Wu C. Investigating the impact of operational variables on manufacturing cost by simulation optimization. International Journal of Not a combinatorial problem Production Economics. 2014;147:634-46.
-

APPENDIX B.2: ARTICLES EXCLUDED AT FULL PAPER STAGE

Reference	Reason for exclusion
1 Abspoel SJ, Etman LFP, Vervoort J, van Rooij RA, Schoofs AJG, Rooda JE. Simulation based optimization of stochastic systems with integer design variables by sequential multipoint linear approximation. <i>Structural and Multidisciplinary Optimization</i> . 2001;22(2):125-38.	Statistical method of approximation
2 Alkhamis TM, Ahmed MA. Sequential stochastic comparison algorithm for simulation optimization. <i>Engineering Optimization</i> . 2004;36(5):513-24.	Statistical method for comparison of results
3 Andradottir S. An Overview of Simulation Optimization via Random Search. <i>Simulation</i> . 2006;13:617-31.	Review paper
4 Azimi P. SIMSUM1: A GENERAL OPTIMISATION VIA SIMULATION APPROACH FOR 0-1 PROGRAMMING MODELS. <i>International Journal of Simulation Modelling</i> . 2012;11(3):150-64.	Not optimisation of a simulation model
5 Barton RR, Ivey JS. Nelder-Mead simplex modifications for simulation optimization. <i>Management Science</i> . 1996;42(7):954-73.	Direct search method, and not specific for combinatorial problems
6 Barton RR, Meckesheimer M. Metamodel-Based Simulation Optimization. <i>Simulation</i> . 2006;13:535-74.	Meta-model methods for continuous, not combinatorial problems
7 Chang X, Dong M, Yang D. Multi-objective real-time dispatching for integrated delivery in a Fab using GA based simulation optimization. <i>Journal of Manufacturing Systems</i> . 2013;32(4):741-51.	Multi-objective optimisation
8 Fu MC, Healy KJ. Techniques for optimization via simulation: An experimental study on an (s,S) inventory system. <i>IIE Transactions</i> . 1997;29(3):191-9.	Not a combinatorial problem
9 Ho, Yu-Chi, R_S Sreenivas, and P. Vakili. "Ordinal optimization of DEDS." <i>Discrete event dynamic systems 2.1</i> (1992): 61-88.	Not optimisation of a simulation model
10 Hurrion RD. An example of simulation optimisation using a neural network metamodel: finding the optimum number of kanbans in a manufacturing system. <i>Journal of the Operational Research Society</i> . 1997;48(11):1105-12.	Metamodel method
11 Jia Q-S. An Adaptive Sampling Algorithm for Simulation-Based Optimization With Descriptive Complexity Preference. <i>Ieee Transactions on Automation Science and Engineering</i> . 2011;8(4):720-31.	Sampling strategy for search
12 Koulouriotis DE, Xanthopoulos AS, Tourassis VD. Simulation optimisation of pull control policies for serial manufacturing lines and assembly manufacturing systems using genetic algorithms. <i>International Journal of Production Research</i> . 2010;48(10):2887-912.	Not a combinatorial problem
13 Martins MSR, Fuchs SC, Pando LU, Lueders R, Delgado MR. PSO with path relinking for resource allocation using simulation optimization. <i>Computers & Industrial Engineering</i> . 2013;65(2):322-30.	Not a combinatorial problem
14 Miranda AK, Del Castillo E. Robust parameter design optimization of simulation experiments using stochastic perturbation methods. <i>Journal of the Operational Research Society</i> . 2011;62(1):198-205.	Robust parameter design methods. Not optimisation of a simulation
15 Pan Y, Zhou M, Chen Z. SIMULATION-BASED OPTIMIZATION FOR RESOURCE ALLOCATION AT THIRD-PARTY LOGISTICS SYSTEMS. <i>International Journal of Industrial Engineering-Theory Applications and Practice</i> . 2012;19(2):101-15.	Not accessible
16 Pinho AFd, Montevechi JAB, Marins FAS, Costa RFdS, Miranda RdC, Friend JD. Evaluation of a proposed optimization method for discrete-event simulation models. <i>Pesquisa Operacional</i> . 2012;32(3):543-60.	Only reporting a black box solver
17 Pichitlamken J, Nelson BL, Hong LJ. A sequential procedure for neighborhood selection-of-the-best in optimization via simulation. <i>European Journal</i>	Sampling strategy for search

of Operational Research. 2006;173(1):283-98.

18 Smew W, Young P, Geraghty J. SUPPLY CHAIN ANALYSIS USING SIMULATION, GAUSSIAN PROCESS MODELLING AND OPTIMISATION. International Journal of Simulation Modelling. 2013;12(3):178-89. Not a combinatorial problem

19 Subramanian D, Pekny JF, Reklaitis GV, Blau GE. Simulation-optimization framework for stochastic optimization of R&D pipeline management. Aiche Journal. 2003;49(1):96-112. Not a combinatorial problem

20 Swisher JR, Hyden PD, Jacobson SH, Schruben LW. A survey of recent advances in discrete input parameter discrete-event simulation optimization. IIE Transactions. 2004;36(6):591-600. Review paper

21 Yoo T, Cho H, Yuecesan E. Hybrid algorithm for discrete event simulation based supply chain optimization. Expert Systems with Applications. 2010;37(3):2354-61. Not a combinatorial problem

APPENDIX B.3: ARTICLES EXCLUDED AT CITATION/REFERENCE REVIEW STAGE

Reference	Reason for exclusion
1 Andradottir S. A method for discrete stochastic optimization. Management Science 1995;41(12):1946-61.	Not a combinatorial problem
2 D. Costa, E.A. Silver, Tabu search when noise is present: An illustration in the context of cause and effect analysis, Journal of Heuristics 4 (1998) 5–23.	Not of a simulation model
3 Guo Yet al (2006) SimOpt: a new simulation optimization system based virtual simulation for manufacturing system. Simul Model Pract Theory 14:577–585	Computer engineering of an algorithm / program
4 Kleijnen, Jack PC, Wim van Beers, and Inneke van Nieuwenhuysse. "Constrained optimization in expensive simulation: Novel approach." European Journal of Operational Research 202.1 (2010): 164-174.	Use of black box optimiser
5 Nemhauser, G.L.; Wolsey, L.A. 1988: Integer and combinatorial optimization. New York: John Wiley & Sons	Not of a simulation model
6 Spinellis, D.D. and Papadopoulos, C.T., 2000. A simulated annealing approach for buffer allocation in reliable production lines. Annals of Operations Research, 93, 373–384.	Not of a simulation model
7 Sriver, Todd A. Pattern search ranking and selection algorithms for mixed-variable optimization of stochastic systems. AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF ENGINEERING AND MANAGEMENT, 2004.	Not of a simulation model
8 Zeng, Q.,&Yang,Z.(2009).Integrating simulation and optimization to schedule loading operations in container terminals. Computers & Operations Research,36, 1935–1944	Not a combinatorial problem
9 Zhang H, Li H (2004) Simulation-based optimization for dynamic resource allocation. Autom Constr 13:409–420	Dynamic modelling

APPENDIX B.4: REFERENCES FOR INCLUDED STUDIES

-
- | Reference |
|--|
| 1 Ahmed MA, Alkhamis TM, Hasan M. Optimizing discrete stochastic systems using simulated annealing and simulation. <i>Computers & Industrial Engineering</i> . 1997;32(4):823-36. |
| 2 Ahmed MA, Alkhamis TM. Simulation-based optimization using simulated annealing with ranking and selection. <i>Computers & Operations Research</i> . 2002;29(4):387-402. |
| 3 M.H. Alrefaei, S. Andradoittir, A simulated annealing algorithm with constant temperature for discrete stochastic optimization, <i>Management Science</i> 45 (1999) 748–764. |
| 4 Andradottir S, Prudius AA. Balanced Explorative and Exploitative Search with Estimation for Simulation Optimization. <i>Inform Journal on Computing</i> . 2009;21(2):193-208. |
| 5 Azadeh A, Maghsoudi A. Optimization of production systems through integration of computer simulation, design of experiment, and Tabu search: the case of a large steelmaking workshop. <i>International Journal of Advanced Manufacturing Technology</i> . 2010;48(5-8):785-800. |
| 6 Ding HW, Benyoucef L, Xie XL. A simulation optimization methodology for supplier selection problem. <i>International Journal of Computer Integrated Manufacturing</i> . 2005;18(2-3):210-24. |
| 7 Ghiani G, Legato P, Musmanno R, Vocaturo F. A combined procedure for discrete simulation-optimization problems based on the simulated annealing framework. <i>Computational Optimization and Applications</i> . 2007;38(1):133-45. |
| 8 Haddock J, Mittenthal J. Simulation optimization using simulated annealing. <i>Computers and Industrial Engineering</i> 1992;20(4):87-395. |
| 9 Ho YC, Cassandras CG, Chen CH, Dai L. Ordinal optimisation and simulation. <i>Journal of the Operational Research Society</i> . 2000;51(4):490-500. |
| 10 Hong, L. Jeff, and Barry L. Nelson. "Discrete optimization via simulation using COMPASS." <i>Operations Research</i> 54.1 (2006): 115-129. |
| 11 Huang C-J, Chang K-H, Lin JT. Optimal vehicle allocation for an Automated Materials Handling System using simulation optimisation. <i>International Journal of Production Research</i> . 2012;50(20):5734-46. |
| 12 Jacobson SH, Sullivan KA, Johnson AW. Discrete manufacturing process design optimization using computer simulation and generalized hill climbing algorithms. <i>Engineering Optimization</i> . 1998;31(2):247-60. |
| 13 Jun Z, Yu-An T, Xue-Lan Z, Jun L. An improved dynamic structure-based neural networks determination approaches to simulation optimization problems. <i>Neural Computing & Applications</i> . 2010;19(6):883-901. |
| 14 Kamrani F, Ayani R, Moradi F. A framework for simulation-based optimization of business process models. <i>Simulation-Transactions of the Society for Modeling and Simulation International</i> . 2012;88(7):852-69. |
| 15 Korytkowski P, Wisniewski T, Rymaszewski S. An evolutionary simulation-based optimization approach for dispatching scheduling. <i>Simulation Modelling Practice and Theory</i> . 2013;35:69-85. |
| 16 Kuo, R. J., and C. Y. Yang. "Simulation optimization using particle swarm optimization algorithm with application to assembly line design." <i>Applied Soft Computing</i> 11.1 (2011): 605-613. |
| 17 Lacksonen, T., Empirical comparison of search algorithms for discrete event simulation. <i>Computers & Industrial Engineering</i> , 2001, 40, 133–148. |
| 18 Prudius AA, Andradottir S. Averaging frameworks for simulation optimization with applications to simulated annealing. <i>Naval Research Logistics</i> . 2012;59(6):411-29. |
| 19 Rosen SL, Harmonosky CM. An improved simulated annealing simulation optimization method for discrete parameter stochastic systems. <i>Computers & Operations Research</i> . 2005;32(2):343-58. |
| 20 L. Shi, S. Olafsson, Nested partitions method for stochastic optimization, <i>Methodology and Computing in Applied Probability</i> 2(3) (2000) 271–291. |

- 21 Xu WL, Nelson BL. Empirical stochastic branch-and-bound for optimization via simulation. *lie Transactions*. 2013;45(7):685-98.
- 22 Yang T, Kuo Y, Chang I. Tabu-search simulation optimization approach for flow-shop scheduling with multiple processors - a case study. *International Journal of Production Research*. 2004;42(19):4015-30.
- 23 Yang T, Fu H-P, Yang K-Y. An evolutionary-simulation approach for the optimization of multi-constant work-in-process strategy - A case study. *International Journal of Production Economics*. 2007;107(1):104-14.

APPENDIX B.5: TAXONOMIES OF SIMULATION OPTIMISATION

Numerous simulation optimisation methods exist for the identification of local optima, including approximation,³⁰⁷ response surface methods (RSM),^{308,309} and ranking and selection.³¹⁰ A taxonomy of local and global simulation optimisation methods is provided in Figure AB.1. A comprehensive review of local simulation optimisation methods is provided by Fu *et al.* in 1994 and 2002.^{207,311}

Identified in this review was the contribution that LS methods had made to the development of global search methods. In particular, many global methods use LS algorithms to identify local optima before moving to another area within the search space. Therefore these local search methods will be considered where relevant within the results of this review.

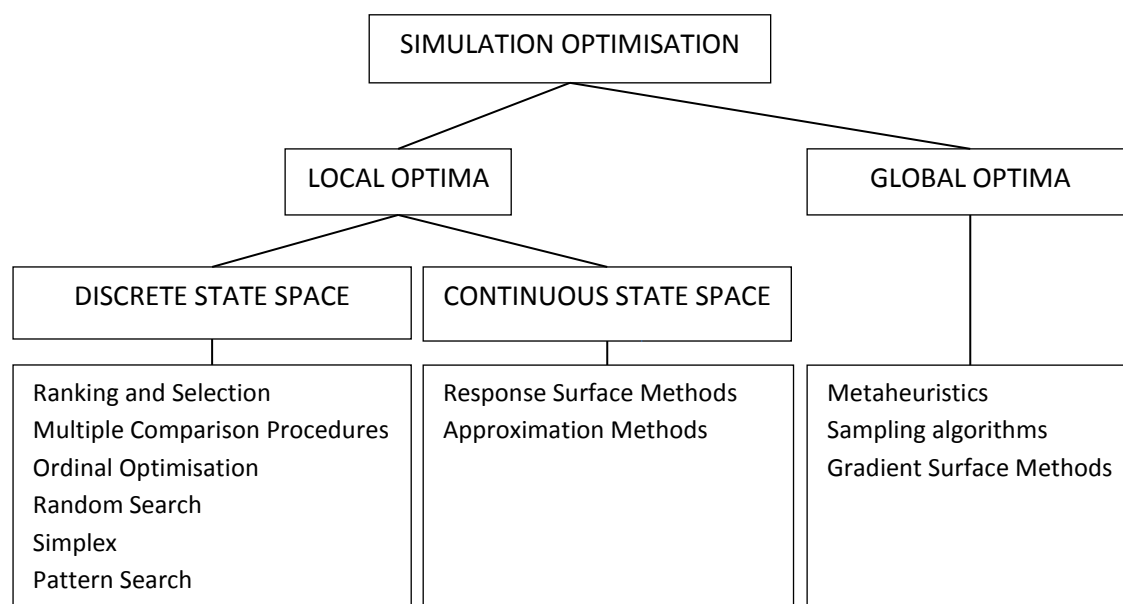


Figure AB.1: Local/Global taxonomy of simulation optimisation methods

Optimisation problems can also be classified by function argument variables that the problem contains. In fact, this step is crucial in determining an appropriate method for an optimisation problem, since many optimisation algorithms and solutions are specifically tailored for a type of

problem. Even methods which promote the ability to work across a range of problem types may still perform better or worse in particular instances.

Optimisation problems with variables that take values from a discrete set are discrete problems, with a discrete input state space. Optimisation problems with variables that can take any real value, either finite or infinite, are continuous optimisation problems. Often, although not always, continuous problems are easier to solve, because the smoothness of the objective function means the function can be used to determine information about solutions within a neighbourhood. Often methods estimate derivatives to determine movements towards an optima. Some problems may have a mixture of both discrete and continuous variables – called mixed state optimisation problems.

This classification of the treatment sequencing problem allows several methods applicable only to continuous state spaces to be ruled out. These include derivative based approaches including gradient based and hessian based methods via standard calculus, and derivative free approaches for continuous problems, including the Nelder-Mead simplex method,³¹² and standard applications of Particle Swarm Optimisation.³¹³

Many continuous optimisation methods rely on the ability to continue moving in a favourable direction. In a discrete search space, the concept of 'direction' may not have any significant meaning. Therefore the theory that underpins many methods may be distinct to either discrete or continuous optimisation, however some methods have proven successful across both state space types. The exact methods contained in this taxonomy are global methods for small (or simple) optimisation problems. These are methods which will guarantee to find a global optima. Therefore they are only relevant for relatively small problems. Examples of these methods include exhaustive and implicit enumeration, Branch and Bound,³¹⁴ Dantzig's simplex method,³¹⁵ and column generation.³¹⁶ For large or more complex problems, particular features of the problem may guarantee to find a global optimum, however due to the size of the problem (or the computational complexity of the problem) this may require a long time to execute.

A taxonomy of simulation optimisation methods grouped by the state space they apply to is provided in Figure AB.2.

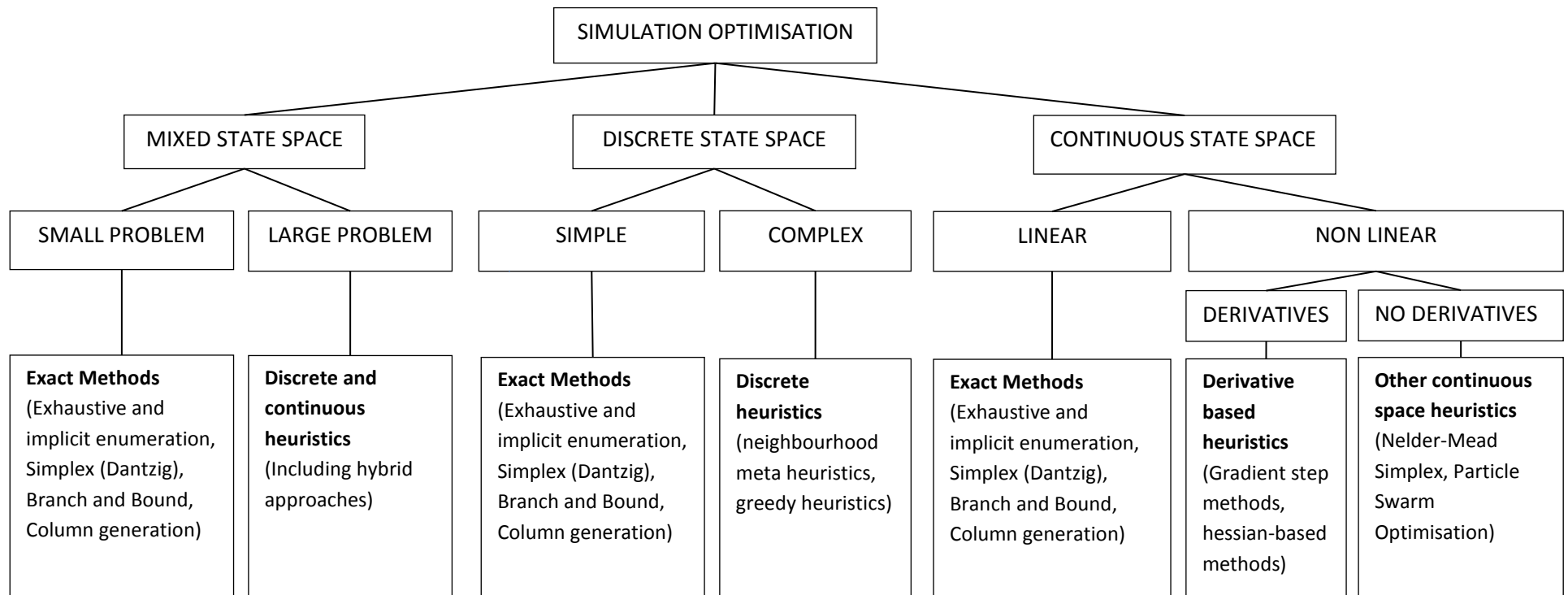


Figure AB.2: State space taxonomy of simulation optimisation methods

APPENDIX B.6: HOW THE METHODS WORK

Traditional Random Search

Random search (RS) is a stochastic optimisation procedure, with a probabilistic (or random) procedure determining which neighbouring solutions are selected during the iteration. The general structure of RS for simulation optimisation involves a defined sampling strategy for the selection of neighbouring solutions. These solutions are evaluated, the best solution selected, and the algorithm iterates to a new point in the search space. The sampling strategy is often a key component of the algorithm. A generic RS algorithm is provided in Table AB.1.²⁰⁶

The number of sample solutions at iteration n is given by M_n , and is a parameter of the sampling strategy. There are no other user-defined (tuning) parameters required for this algorithm. The generic RS algorithm does not include a stopping criterion, because RS literature has shown that the sequence will converge asymptotically on the global optima ($x_n \rightarrow x^*$ for $n \rightarrow \infty$). However, in practice stopping rules are applied to ensure the RS algorithm terminates in a finite period of time.²⁰⁶

Table AB.1: Traditional random search algorithm for simulation optimisation (Adapted from Andradottir (2006))²⁰⁶

Step	Process	Details
0	Initialise	Choose initial sampling strategy S_1 and set iteration count $n = 1$
1	Sample	Select $x_n^{(1)}, \dots, x_n^{(M_n)} \in X$ according to the sampling strategy S_n
2	Simulate	Estimate $g(x_n^{(i)})$, for $i = 1, \dots, M_n$, using the simulation model
3	Update	Select optimal solution x_n^* , update S_{n+1} , $n = n+1$, Go to Step 1

M_n is the number of sample solutions at iteration n

Adaptive Random Search – Balanced Explorative and Exploitative Search (BEES)

The concept underpinning the adaptive random search method is that there is an optimal point where each random search method should switch from the exploration of the global search space, to an exploitation of the local search space. The authors, Andradottir & Prudius,²¹⁶ present two variations of their methodological framework. Deterministic optimisation using randomised search (R-BEES) uses a probability sampled from a uniform distribution to switch the algorithm between either a local or global search. Stochastic optimisation using randomised search (R-BEES with estimation = R-BEES-E) is applicable for stochastic problems, and also uses a probabilistic switch between local and global search.

The authors also present adaptive variations of the two methods. These require tracking the change in the optimal solution and the distance between the two optimal solutions. The

algorithm switches from local to global if the change in optimal solution is small (suggesting a local optimum nearby). A switch from global to local can happen when a promising region is found (a small improvement) or if a large improvement is found in a short distance (to focus on that area).

Adaptive Random Search – Convergent Optimisation via Most Promising Area Stochastic Search (COMPASS)

Initially, the most promising area identified is the whole feasible search space. At every iteration of the algorithm, a small number of simulation runs are undertaken of a sample of solutions. All visited solutions are collected into a set, and full simulations are undertaken as assigned by the simulation-allocation rule (SAR) for each solution within the set. As the algorithm proceeds, the set enlarges and more runs are required for each solution. The sample average of each solution therefore updates iteratively. As the sample becomes sufficiently large, the algorithm can select the best performing solution.

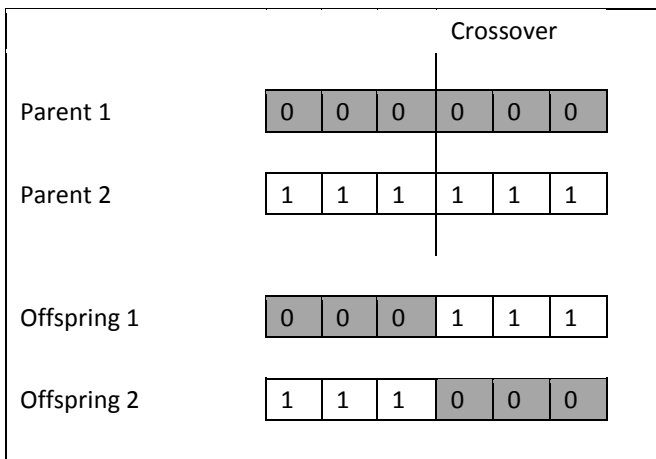
For each iteration, the algorithm selects the best current solution, and the most-promising area is defined as the set of feasible solutions that are at least as close to the current best as they are to other visited solutions. Therefore as more solutions are sampled, the most-promising area shrinks in size.

Genetic algorithms

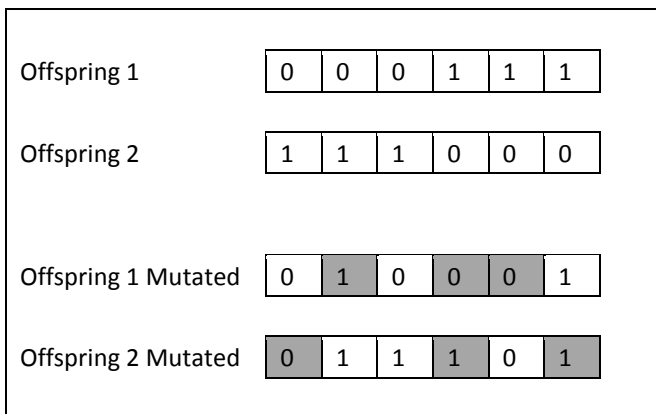
Genetic algorithms are a population-based metaheuristic, with a pool of potential solutions maintained in the algorithm. They mimic the process of natural evaluation and use concepts of natural selection and genetic inheritance to navigate the search space.³¹⁷ Each potential solution is represented as a chromosome with each decision variable (often called an element) a gene. This representation type naturally leads GAs to be a popular method for combinatorial problems.

There are two key operations within the GA which evolve the pool of potential solutions – crossover and mutation. Crossover is the operation of taking parent solutions from the population, and generating offspring with a ‘crossed over’ set of chromosomes (see Box AB.1). Mutation is the altering of each gene independently via a specified probability (the mutation rate - see Box AB.2).

Evolutionary algorithms are related to genetic algorithms, but they only include offspring mutation, and not crossover.



Box AB.1: One point crossover



Box AB.2: Mutation

The SGA has a simple iterative process to replace the whole population with the offspring generated by crossover and mutation:

1. Select parents for the mating pool (size of mating pool = population size)
2. Shuffle the mating pool
3. Select a set of parents from the mating pool. For each consecutive pair apply crossover with a defined probability, otherwise copy parents. This requires two parents and results in two offspring.
4. For each offspring apply mutation (for a binary representation - bit-flip with a defined probability independently for each bit)
5. Replace the set of parents with the resulting offspring and return to mating pool.

Even more so than other metaheuristic methods, the representation of the solution is crucial for genetic algorithms. This is because the newer and more complex crossover and mutation operators are directly informed by the representation type. The SGA was represented by a bit array (binary representation). However, as already discussed, binary representation is only suitable for a subset of combinatorial optimisation problems.

Permutation representation as already discussed is more appropriate for ordering and sequencing problems (e.g. TSP, sequencing problem), and the RA treatment sequencing problem. However, permutation representation introduces challenges for genetic algorithms and standard crossover and mutation operations. This is because the order and the adjacency of elements within the solution sequence are crucial.³¹⁸ Normal crossover and mutation operators lead to inadmissible solutions, and therefore at least two elements/genes within the chromosome must be changed to ensure a valid solution. This increases the neighbourhood size for the GA.

Four common mutation operators for permutation representation and provided below (Box AB.3, with details on their ability to retain order and adjacency information within the chromosome.

Like mutation, normal crossover operations will often result in inadmissible solutions. Therefore a large number of specialised operators have been devised which focus on combining order and adjacency information from the two parents. These include order one crossover, partially mapped crossover, cycle crossover, edge recombination and multi-parent recombination.³¹⁹

Within a genetic algorithm, there are two points where selection of solutions occurs. Firstly, when selecting from the current population which solutions are going to take part in mating (parent selection). Secondly, when selecting which parents and offspring go into the next generation (survivor selection).

<p>1. Insert mutation</p> <table border="1"> <tr> <td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td> </tr> </table> <table border="1"> <tr> <td>1</td><td>2</td><td>5</td><td>3</td><td>4</td><td>6</td> </tr> </table> <p>Select two values, move second to follow first and shift rest to accommodate. This method preserves most order and adjacency information.</p>	1	2	3	4	5	6	1	2	5	3	4	6
1	2	3	4	5	6							
1	2	5	3	4	6							
<p>2. Swap mutation</p> <table border="1"> <tr> <td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td> </tr> </table> <table border="1"> <tr> <td>1</td><td>5</td><td>3</td><td>4</td><td>2</td><td>6</td> </tr> </table> <p>Pick two values and swap. Preserves most adjacency information but disruptive to order.</p>	1	2	3	4	5	6	1	5	3	4	2	6
1	2	3	4	5	6							
1	5	3	4	2	6							
<p>3. Inversion mutation</p> <table border="1"> <tr> <td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td> </tr> </table> <table border="1"> <tr> <td>1</td><td>5</td><td>4</td><td>3</td><td>2</td><td>6</td> </tr> </table> <p>Pick two values and invert substring between them. Preserved most adjacency information but disruptive to order.</p>	1	2	3	4	5	6	1	5	4	3	2	6
1	2	3	4	5	6							
1	5	4	3	2	6							
<p>4. Scramble Mutation</p> <table border="1"> <tr> <td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td> </tr> </table> <table border="1"> <tr> <td>1</td><td>4</td><td>3</td><td>5</td><td>2</td><td>6</td> </tr> </table> <p>Select a subset and randomly rearrange. Disruptive to order and adjacency.</p>	1	2	3	4	5	6	1	4	3	5	2	6
1	2	3	4	5	6							
1	4	3	5	2	6							

Box AB.3: Mutation operators

Traditionally, fitness-proportionate selection has been a commonly applied method, including roulette-wheel algorithms and Baker's stochastic universal sampling (SUS) algorithm.³²⁰ However, these can be problematic if a highly fit member dominates the population and can lead to premature convergence. Also, when a population is very similar in terms of fitness, the selection pressure of the algorithm can drop and convergence does not occur.^{321,322} Selection pressure is the informal term within the evolutionary algorithm literature to reflect the balance between exploration and exploitation. Some authors quantify selection pressure as the ratio of maximum to average fitness within the population.³²³

Rank based selection is a potential solution to the limitations of fitness-proportionate selection, and is based on relative rather than absolute fitness. This is a common way to maintain selection pressure within the algorithm. However, a sorting overhead is imposed on the algorithm when this is introduced. Tournament selection is similar, with a subset of solutions randomly selected from the pool and the best solution then selected as a parent. The size of the subset (tournament) is a user-defined parameter and the selection pressure is highly sensitive to the tournament size.³²⁴

For survivor selection, there are two general methods: age based selection where the oldest is deleted, and fitness based selection. However, two special cases have been introduced and widely used - Elitism (always keep the best solution), and GENITOR (GENetic ImplemaTOR - always delete the worst solution).³²⁵ Elitism means that the fitness of the best solution in the population never deteriorates as the algorithm iterates. The ability of a GA to asymptotically converge has been found to rely on elitism, and several studies have shown that algorithms with elitism converge faster than those without.^{326,327}

If there are constraints to the problem, then genetic algorithms can incorporate penalties to the objective function.³²⁸ However, it is much preferable to encode out any infeasible solutions, rather than evaluate and then apply a subsequent penalty.³²⁹

Tabu search

Tabu search (TS) is a local search method with a flexible memory structure. Unlike branch and bound (rigid memory structure) and simulated annealing (no memory structure), TS has a short-term memory structure to avoid moving back to a recently visited solution (tabu moves). TS has the same foundations as an ordinary local or neighbourhood search. Each vector of decision variables from the feasible space $x \in \mathbf{X}$ has an associated neighbourhood $\mathbf{N}(x) \subset \mathbf{X}$ and each neighbouring solution $x' \in \mathbf{N}(x)$ is reached from x via a move operation.

Unlike genetic algorithms and simulated annealing, which are probabilistic/stochastic search methods, TS accepts non-improving solutions deterministically by guiding a steepest-descent/ascent hill-climbing heuristic based on the best neighbouring non-tabu solution.

However, for problem where $\mathbf{N}(x)$ is large or each element of $\mathbf{N}(x)$ is computationally costly to retrieve, the steepest descent method may be impractical. Therefore any improving move can be selected as a relaxation to the algorithm.

The decision regarding the evaluation (and subsequent selection) of a solution is called the 'candidate list strategy' and is an important component of the TS method. In particular, the relevance of choosing a good solution is magnified when looking to move out of a local optimum (where a descent method would normally terminate). At this point, the method requires the selection of the best non-improving solution.²⁷²

If the best neighbouring solution (descent) heuristic is relaxed to avoid evaluating every possible move in a current neighbourhood, then it needs to be the case that neighbours are identified which are meaningful for the particular problem. Candidate lists are methods to isolate good candidate moves from the current neighbourhood. Candidate moves use an intelligent process, rather than random or naïve processes.³³⁰ It allows a reduction in the number of evaluations conducted and allows the problem structure to be exploited (called 'context related rules' in Glover & Laguna (1997)).^{272,330}

Examples of candidate list strategies include the 'elite candidate list strategies', which records the best moves encountered and then implement this list of moves for each solution. 'Aspiration plus strategy' is an adaptive method of only implementing moves which currently have proved to improve a solution by a given threshold level. More detail regarding these strategies is provided in Rangaswamy et al (1998).³³⁰ It should be noted that candidate list strategies are memory based methods, in keeping with the general paradigm of TS. The memory that is used to record recently visited (tabu) solutions is called the Tabu list. This list is typically short term, and therefore updated with each move. Diversification within the algorithm can be encouraged by increasing the length of the tabu list (tenure), or using a dynamic tabu list structure.

Nested partitions

The underlying concept of the algorithm is to systematically partition the feasible region into subregions, evaluate the potential of each subregion, and then focus on the most promising region. The process iterates with each partition nested within the previous most promising partition.

More formally, a promising region at iteration k , $\sigma(k)$, is defined, and the initial state of the algorithm contains no knowledge regarding the most promising region, $\sigma(0) = \mathbf{X}$. At the k -th iteration of the algorithm, a region $\sigma(k) \subseteq \mathbf{X}$ is considered the most promising. The most promising region is partitioned into M subsets which cover the entire best region, and the remaining surrounding region $\mathbf{X}/\sigma(k)$ is aggregated into one partition. At each iteration, $M + 1$ subsets are generated and each region is randomly sampled to estimate a set of solutions. The next most promising region is the subset with the best sampled value. If the surrounding region is the most promising, then the algorithm backtracks. The new most promising region is partitioned and sampled which generates a sequence of set partitions, with each partition nested within the last.

A stochastic version of the method is developed by Shi et al. (2000) where the only modification from the original algorithm is the way the optimum is estimated. This modification is required to ensure convergence with stochastic problems. Specifically, the best solution is the most frequently visited region (the region most often in the most-promising region).

The NP method is particularly robust for simulation optimisation, because nested partition is a set-based method with a stochastic move operation. Like genetic algorithms, this stochastic move appears to make it relatively insensitive to the noise from the simulation model.³³¹

To improve the nested partition method, a modification by Olafsson (2004) was made to introduce a statistical selection mechanism to guide the search, which acts as a control to the noise in the algorithm. A first phase to the algorithm undertakes a small sample of the performance variance in each region. This also allows very poor regions to be screened. Then based on the initial samples a second phase of sampling is conducted to ensure that the correct selection of a region is made, using an indifference zone selection procedure with a specified minimum probability.³³²

Ordinal optimisation

In ordinal optimisation (OO), the objective is relaxed from finding the optimal solution to finding a subset of 'good enough' solutions. This is also known as 'goal softening'. The authors make it clear that this is a retreat from a hard method to a soft method, but it is necessary within the context of simulation optimisation. Exact optimisation may just be too computationally expensive for many simulation optimisation problems. The shift from cardinal to ordinal optimisation methods means looking to maximise the alignment probability of alternative solutions – that is, the probability that competing solutions are ranked correctly.

The aim of the OO algorithm is to find a subset of the search space \mathbf{G} which is 'good enough'. In traditional optimisation, the subset \mathbf{G} is a solution, the optimum (or an optima). However in OO the subset \mathbf{G} is a set of solutions. The OO algorithm uses a selected subset \mathbf{S} of the search space, which again is a set of solutions. In traditional optimisation, \mathbf{S} is also a single solution, with the aim of $\mathbf{S} = \mathbf{G}$, so that the selected solution is the optimum. However in ordinal optimisation, the objective is that \mathbf{G} intersects \mathbf{S} above a user defined alignment probability (e.g. 95%), k . This alignment probability, $Prob[|\mathbf{G} \cap \mathbf{S}| \geq k]$, is the probability that there are k truly good enough designs within \mathbf{S} .

An important component of the method is the Ordered Performance Curve (OPC). This simply ranks the designs from a sample of the search space in order of their performance. Lau *et al.* (1997) identify four general types of OPC, as denoted in Figure AB.3.³³³

The procedure for OO is simple:

1. Using a uniform and random sampling method, sample N designs
2. Estimate the performance of these N designs (using a crude fast model if required)
3. Estimate the OPC type and the noise level within the model used for Step 2. The user specifies the size of the good enough set and the required alignment level k (e.g. 95%)
4. Calculate the s value (the initial size of subset \mathbf{S}) based on Ho *et al.* (2008) tables (data from step 3).³³⁴
5. Select the top s designs of N and specify as the selected set \mathbf{S} .
6. This top set contains at least k truly good enough designs with probability no less than 95%

The OO method introduces a tolerance to the imprecise estimates determined by a stochastic simulation model because the goal has been softened and the user has a high confidence in obtaining a 'good enough' design from a selected set. OO is a method itself, but the idea of ordinal comparison can be integrated into alternative search algorithms to provide a much faster convergence rate.³³⁴

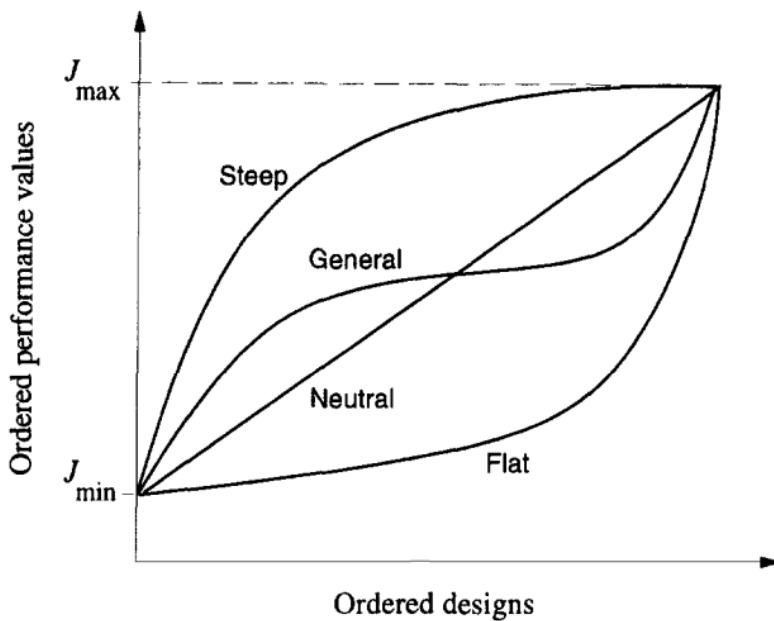


Figure AB.3: Types of Ordered Performance Curve (Lau et al (1997))³³³

Particle swarm optimisation

Particle swarm optimisation (PSO) is a population based method. The algorithm is initialised with a population of random particles (representing different solutions) and each generation of the algorithm sees an updating of the particles, which allows a search for the optima.

Each particle (id) records its previous best position (P_{bestid}) and has a velocity with which the particle travels within the multi-dimensional search space ($x \in X$). At each iteration, the particle with the best fitness (G) and the position vector of the current particle are combined to adjust the velocity. That new updated velocity is then used to compute the new position for the particle. Global best and local best alternatives are possible when looking to influence the direction of the swarm. Two tuning parameters (c_1 and c_2) determine the relative influence of the social and cognition components (learning factors) of the algorithm. With these, the following updating rule is applied within the algorithm:

$$v_{id}^{new} = v_{id}^{old} + c_1 \times rand_1 \times (P_{bestid} - x_{id}) + c_2 \times rand_2 \times (G_{bestid} - x_{id}) \quad [AB.1]$$

$$x_{id}^{new} = x_{id}^{old} + v_{id}^{new} \quad [AB.2]$$

The maximum velocity is traditionally constrained, which controls the ability of the algorithm to explore the search space. However, an inertia component has been added to provide a better balance of exploration and exploitation, and this has rendered maximum velocity redundant.³³⁵ At first, the inertia weight was a constant parameter, however a decreasing rate for the inertia parameter was found to have more potential. This is because the higher weight

enables great exploitation as the beginning to find a good area, and the smaller weight towards the end has the ability to search the local area.

Averaging framework for simulated annealing

The averaging framework for simulated annealing, as developed by Prudius et al. (2012), is adaptive, and uses the information gathered at previous iterations of the algorithm to determine the amount of simulation effort required.²³⁵ Averaging is used to provide estimates of the objective values based on the average of all previously visited solutions.

The authors present a random search variant of the method which incorporates a point-based movement. This allows the random search to iterate between different points within the feasible space, and is particularly applicable to discrete optimisation problems. The authors find that the random search method is globally convergent under mild conditions, and do not explore any potential limitations of the method.

Empirical stochastic branch and bound

The empirical stochastic branch-and-bound method, as developed by Xu et al. (2013), is a combination of nested partitioning and branch and bound. It uses the partitioning structure of stochastic branch and bound to determine subregions of the search space. However, it uses bounds based on the performance of sampled solutions, as per the nested partitions method.²³³ These bounds are determined by maintaining a set of feasible solutions, as well as a set of all solutions. It simulates solutions within the set of feasible solutions and computes the bounds using their estimated performance. In the next iteration, a subset of solutions from the current partition is sampled. The method is memory intensive due to the overhead required to retain and refine the partition structure. The assumptions regarding bounding are key, because they provide the guarantee for convergence. Also, there are four tuning parameters which enable the balance between sampling solutions and running simulations to be adjusted.

APPENDIX B.4: DATA EXTRACTION TABLES

Ahmed MA, Alkhamis TM. Simulation-based optimization using simulated annealing with ranking and selection. Computers & Operations Research. 2002;29(4):387-402. ²²⁰	
Development	
What is the method?	Simulated annealing with ranking and selection
What problem was the method originally developed for?	Solving a discrete stochastic optimisation problem
Has the method been adapted from its original context?	The method combines simulated annealing and ranking and selection procedures.
Was the method designed to address discrete event simulation (DES) optimisation?	Yes
Or is it a general optimisation method that could be suitable for DES?	No, it adapts general optimisation methods for use with a DES
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	That simulated annealing (SA) has been successfully applied to solve a wide range of combinatorial optimisation problems. Ranking and selection (RS) procedures are statistical methods designed to solve discrete stochastic optimisation problems. Combining the two approaches resolves the limitations of each individual approach. The method shows that the configuration that has been visited most often in the first m iterations converges almost surely to a globally optimum solution.
How does the method work	In each iteration of the algorithm, two neighbouring configurations are compared using RS. This procedure explicitly sets the sample size (run length), so to guarantee that the probability of selecting the best configuration is suitably large. <ol style="list-style-type: none"> 1. Obtain initial solution and temperature 2. Choose neighbour candidate based on probability distribution 3. Move to selected candidate 4. Update temperature and repeat
What assumptions does the method require?	The method resembles original SA (it accepts worse neighbouring configurations, with an acceptance probability which tends to zero). This therefore requires a 'generating probability function' for each candidate solution points.
What are the theoretical limitations of the method?	SA tends to need an accurate evaluation of the objective function values, and RS tends to only be efficient when the number of alternatives are small.

	Original SA only guaranteed to converge in probability.
What are the potential biases associated with the method?	None identified by the authors. Standard SA biases (initial temperature, temperature length, cooling rate, final temperature)
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes – an inventory system example. This was a standard infinite horizon single item periodic review inventory model with zero lead times. An (s,S) policy for linear costs – order placed when inventory below s, and order is the difference between S and the inventory position. At each period an order can be placed for any positive quantity of stock. The objective is to minimise the long-run average cost function per period. By varying s and S. The model is developed with exponential demands, so that an analytical solution is possible.
If so, how did it perform? (Speed, optimality, ease of implementation)	Little information given. The method converges almost surely to a global optimal position. Results are very close to analytical solutions in all cases.
If not, are there any suggestions to its practical applicability?	I think by forcing simplistic assumptions (exponential demand) you are simplifying the search space.
Ahmed MA, Alkhamis TM, Hasan M. Optimizing discrete stochastic systems using simulated annealing and simulation. Computers & Industrial Engineering. 1997;32(4):823-36. ²¹⁹	
Development	
What is the method?	Simulated annealing
What problem was the method originally developed for?	Solving a discrete stochastic optimisation problem
Has the method been adapted from its original context?	They have used Simulated Annealing. They have adapted the rejection/acceptance criteria for stochastic constraints. Normally SA uses a penalty system for constraints, but this can result in an “optimal” result being infeasible. Therefore an adaptation to the SA algorithm is used which only accepts feasible transitions.
How does the method work	Initialise the user-defined settings (initial temp, temp length, cooling rate, final temp) Standard SA methods with only feasible transitions accepted
Was the method designed to address discrete event simulation (DES) optimisation?	Yes
Or is it a general optimisation method that could be suitable for DES?	No, it adapts general SA for use with a DES
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	That simulated annealing (SA) has been successfully applied to solve a wide range of combinatorial optimisation problems. Acceptance/rejection is based on the expected output of the simulation model. Rejection/acceptance is modified to take into consideration stochastic system constraints. The method also includes control variate as a variance reduction technique.

How does the method work	Standard SA method, with a couple of modifications: <ol style="list-style-type: none"> 1. SA algorithm only accepts feasible transitions 2. Neighbourhood based on a single perturbation 3. Controlled probability for uphill perturbation move
What assumptions does the method require?	None
What are the theoretical limitations of the method?	SA tends to need an accurate evaluation of the objective function values. Only allows single neighbour perturbation. Inefficient?
What are the potential biases associated with the method?	Standard SA biases (initial temperature, temperature length, cooling rate, final temperature)
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	<ol style="list-style-type: none"> 1. Classical machine repair problem 2. Two-echelon repairable item provisioning system 3. Multi-echelon repairable item provisioning system
If so, how did it perform? (Speed, optimality, ease of implementation)	For 1, the algorithm found the analytical solution. For 2, the algorithm almost found the analytical solution. For 3, not analytical solution is possible – it was tested against a greedy algorithm with the same number of iterations. The proposed algorithm was superior for all 6 cases of Test 3. No further details
If not, are there any suggestions to its practical applicability?	
M.H. Alrefaei, S. Andradoittir, A simulated annealing algorithm with constant temperature for discrete stochastic optimization, Management Science 45 (1999) 748–764.²²¹	
Development	
What is the method?	Modified simulated annealing
What problem was the method originally developed for?	Discrete stochastic optimisation
Has the method been adapted from its original context?	The method differs from the original simulated annealing algorithm by using a constant (rather than decreasing) temperature
Was the method designed to address discrete event simulation (DES) optimisation?	Yes
Or is it a general optimisation method that could be suitable for DES?	No
Theoretical basis	
How does the method address optimising a DES	Because rapidly decreasing temperatures reduce the algorithm time, but the convergence of the algorithm is not guaranteed. The higher the

with a combinatorial problem?	temperature value, the greater the chance of accepting a worse move (hill climbing move).
How does the method work	It is a simple modification of the standard SA method. The method also modifies the criteria to evaluate the optimal solution. Firstly by using the number of visits made to different states, and secondly to use the state that has the best average estimated value.
What assumptions does the method require?	Standard SA assumptions
What are the theoretical limitations of the method?	The paper provides a proof of almost-sure convergence.
What are the potential biases associated with the method?	Standard SA biases
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Discrete queuing systems
If so, how did it perform? (Speed, optimality, ease of implementation)	Compared against GM-GP and FH SA algorithms (using decreasing annealing schedule), the modified algorithms perform well. However, the performance depends on the choice of temperature, neighbourhood, and the number of estimated function values from each iteration. However, the algorithm does not show great sensitivity to the initial temperature. Better overall performance in comparison, with better overall optimal solutions estimated.
If not, are there any suggestions to its practical applicability?	No
Andradottir S, Prudius AA. Balanced Explorative and Exploitative Search with Estimation for Simulation Optimization. Informs Journal on Computing. 2009;21(2):193-208.²¹⁶	
Development	
What is the method?	Balanced Explorative and Exploitative Search with Estimation (BEESE) for simulation optimisation
What problem was the method originally developed for?	Simulation optimisation
Has the method been adapted from its original context?	No
Was the method designed to address discrete event simulation (DES) optimisation?	Yes
Or is it a general optimisation method that could be suitable for DES?	No
Theoretical basis	
How does the method address optimising a DES	The method is a framework which balances exploration for a global search with exploitation for a local search. The advantage is that its

with a combinatorial problem?	numerically efficient when applied to solve problems with little known structure
How does the method work	<p>The concept underpinning the method is that there's an optimal switch point from exploration to exploitation.</p> <p>Deterministic optimisation using R-BEES has a simple probability assigned which is sampled using a uniform distribution to switch between searching either the global space or the local space.</p> <p>Stochastic optimisation using R-BEES, with a probability parameter sampled to switch between the current optimal solution and sampling as before.</p> <p>Adaptive variations of the two methods are possible. These require tracking of the change in optimal solution, and the distance between the top two optimal solutions. The algorithm switches from local to global if the change in optimal solution is small (suggest near local optima). If making good progress, the local search is maintained. A switch from global to local can happen when a promising region is found (small improvement) or large improvement with a short distance (suggesting a focus on the local area).</p>
What assumptions does the method require?	None, the method is a framework to guide the development of search algorithms which more formally consider the balance between exploration and exploitation
What are the theoretical limitations of the method?	None (see above)
What are the potential biases associated with the method?	The method requires user-specified algorithms which deterministically, stochastically or adaptively switch the algorithm between local and global search. These are open to bias.
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	No
If so, how did it perform? (Speed, optimality, ease of implementation)	
If not, are there any suggestions to its practical applicability?	
Azadeh A, Maghsoudi A. Optimization of production systems through integration of computer simulation, design of experiment, and Tabu search: the case of a large steelmaking workshop. International Journal of Advanced Manufacturing Technology. 2010;48(5-8):785-800.²³⁰	
Development	
What is the method?	A combination of design of experiment and Tabu search
What problem was the method originally developed for?	A discrete production system with discrete decision making parameters

Has the method been adapted from its original context?	Both design of experiment, and Tabu search, are established simulation and global metaheuristic optimisation methods.
Was the method designed to address discrete event simulation (DES) optimisation?	Yes
Or is it a general optimisation method that could be suitable for DES?	
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	It uses a metaheuristic method (Tabu search) to optimise the fitted metamodel. This limits the simulations required by fitting a regression model using parameter sets determined by the DOE. The Tabu search is then used to optimise this regression function.
How does the method work	<ol style="list-style-type: none"> 1. Formulate problem 2. Simulation verification and validation 3. Simulated output and parameters 4. Use ANOVA and RSM to establish a design of experiments 5. Fit a polynomial order regression (Least Squares) 6. Identify efficient parameters 7. Undertake a Tabu search
What assumptions does the method require?	Assumptions regarding the functional form of the fitted regression model
What are the theoretical limitations of the method?	Tabu searches are memory intensive.
What are the potential biases associated with the method?	Regression model determinant on the simulation data – appropriate design required
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes a discrete production system with discrete decision making parameters.
If so, how did it perform? (Speed, optimality, ease of implementation)	It claims a global optimum was identified but it's not fully clear how that's proven in this case.
If not, are there any suggestions to its practical applicability?	The method was only tested for 4 factors (converters, slag pockets, mixers, blastfurnaces), which only required 81 experiments to be run.
Ding HW, Benyoucef L, Xie XL. A simulation optimization methodology for supplier selection problem. International Journal of Computer Integrated Manufacturing. 2005;18(2-3):210-24.²²⁶	
Development	
What is the method?	Genetic algorithm
What problem was the method originally	It is designed for a supplier selection problem. This is a supply chain planning problem.

developed for?	
Has the method been adapted from its original context?	Genetic algorithms have been used in numerous optimisation problems.
Was the method designed to address discrete event simulation (DES) optimisation?	No
Or is it a general optimisation method that could be suitable for DES?	Yes
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	<p>The method allows the encoding of design points as chromosomes. Design points are explicitly discrete in a combinatorial problem, and therefore the methodology is naturally aligned with our problem.</p> <p>GA's used to search large, non-linear search spaces where expert knowledge is lacking or difficult to encode and where traditional optimization techniques fall short.</p>
How does the method work	<ol style="list-style-type: none"> 1. Initialise parameters for GA optimizer and sim model 2. Create initial population 3. Check feasibility of each network, and repair if not feasible (this allows 'bad genes' to be removed and only feasible networks evaluated) 4. Create and run DES for each individual 5. Calculated the fitness of each individual according to the fitness definition 6. Select individuals for mating 7. Mate individuals to produce offspring 8. Check stopping criteria
What assumptions does the method require?	<p>GA's work with a population of individual strings (chromosomes). Each string represents a possible solution. In practice, each position in the chromosome may take on one of a finite set of values. Each chromosome is assigned a fitness values according to the result of the simulation. Highly fit chromosomes survive more frequently and are given more opportunities to reproduce.</p> <p>Therefore for realistic problems, GA can often find good(near optimal) solutions in a relatively short search period</p>
What are the theoretical limitations of the method?	Not sure at present. None discussed
What are the potential biases associated with the method?	<p>GA's require user input in key ways:</p> <ol style="list-style-type: none"> 1. How constraints/legal solutions are managed 2. How fitness is evaluated 3. How mutation and reproductions occurs

	4. Stopping criteria
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes – a supply chain simulation model is used. A binary string is used, with segments of the string representing supplier utilisation, assignment weight, and replenishment level. Maximum GA generations set at 500, and each population contains 20 individuals. Roulette wheel selection used, with two-point crossover. Fixed probabilities for mutation and crossover. An elitist strategy is used to preserve the best individuals. A penalty factor is used to avoid ‘missed demands’ – constraint on system
If so, how did it perform? (Speed, optimality, ease of implementation)	Each simulation takes less than one second. The algorithm converges quickly (within 100 generations – so a couple of minutes).
If not, are there any suggestions to its practical applicability?	The case study illustrates the applicability of the methods. The simulation model is coded efficiently which allows a rapid evaluation. There could have been greater evaluation of the performance by altering user-defined parameters of the algorithm. Little is said about these fundamental assumptions.
Ghiani G, Legato P, Musmanno R, Vocaturo F. A combined procedure for discrete simulation-optimization problems based on the simulated annealing framework. Computational Optimization and Applications. 2007;38(1):133-45.²²²	
Development	
What is the method?	Simulated Annealing
What problem was the method originally developed for?	Optimising a discrete stochastic problem
Has the method been adapted from its original context?	Yes
Was the method designed to address discrete event simulation (DES) optimisation?	No
Or is it a general optimisation method that could be suitable for DES?	Yes
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	It uses Simulated Annealing, combined with a statistical procedure for comparing solutions.
How does the method work	Standard SA methodology with a Rinott Procedure to compare the current solution and neighbour by using ‘Indifference Zone Ranking and Selection’ = SARP

What assumptions does the method require?	The Rinott Procedure requires the definition of an indifference zone width, and a confidence level. An objective function gap of less than the indifference zone width is considered negligible. In the first iteration of the algorithm, independent replications from a reference configurations are taken to estimate the sample mean and marginal sample variance, and subsequently solve the Rinott integral.
What are the theoretical limitations of the method?	None reported in the paper. In theory, standard SA limitations
What are the potential biases associated with the method?	None reported in the paper. In theory, standard SA biases
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	No, only test problems
If so, how did it perform? (Speed, optimality, ease of implementation)	-
If not, are there any suggestions to its practical applicability?	-
Haddock J, Mittenthal J. Simulation optimization using simulated annealing. Computers and Industrial Engineering 1992;20(4):87}395.²²³	
Development	
What is the method?	Simulated annealing
What problem was the method originally developed for?	Standard optimisation
Has the method been adapted from its original context?	No
Was the method designed to address discrete event simulation (DES) optimisation?	No
Or is it a general optimisation method that could be suitable for DES?	Yes
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	Standard SA
How does the method work	Uses traditional methods for SA. No changes identified
What assumptions does the method require?	The method resembles original SA (it accepts worse neighbouring configurations, with an acceptance probability which tends to zero). This therefore requires a 'generating probability function' for each candidate solution points.

What are the theoretical limitations of the method?	SA tends to need an accurate evaluation of the objective function values, and RS tends to only be efficient when the number of alternatives are small. Original SA only guaranteed to converge in probability.
What are the potential biases associated with the method?	None identified by the authors. Standard SA biases (initial temperature, temperature length, cooling rate, final temperature)
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	A simple simulation model is used but due to age of the study it is unlikely to be relevant (1992)
If so, how did it perform? (Speed, optimality, ease of implementation)	
If not, are there any suggestions to its practical applicability?	
Ho YC, Cassandras CG, Chen CH, Dai L. Ordinal optimisation and simulation. Journal of the Operational Research Society. 2000;51(4):490-500.²³²	
Development	
What is the method?	Ordinal optimisation
What problem was the method originally developed for?	Optimisation of a complex stochastic simulation model
Has the method been adapted from its original context?	No
Was the method designed to address discrete event simulation (DES) optimisation?	No
Or is it a general optimisation method that could be suitable for DES?	Yes
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	Rather than requiring the computations to make a precise estimate of each compared design (which converges slowly), this method of ordinal comparison can converge exponentially fast. Also the method uses goal softening to ease the computational burden. The key concepts are that order converges exponentially fast, compared to value converging at a much slower rate. It is easier to estimate whether $A > B$ than it is to estimate $A - B = ?$

	Also, goal softening eases the computational burden of finding the optimum
How does the method work	<p>The method assumes that the performance measures are normally distributed. The optimisation problem can then be reduced to determining whether the difference in the means is positive or negative.</p> <p>The method is then to draw independent samples of the compared solutions, and estimate the indifference amount (e.g. the amount of overlap between the distributions of the two compared solutions). Then as the algorithm iterates the simulation will finally produce an estimate of the ordinal difference between the two solutions which is below a defined level of tolerance (the indifference amount).</p> <p>The method also promotes the use of variance reduction techniques (e.g. common random numbers) to reduce the variance in the comparison</p>
What assumptions does the method require?	Normality in the distribution in the uncertainty of the output function
What are the theoretical limitations of the method?	None noted by the authors
What are the potential biases associated with the method?	Normality. The use of an indifference amount causes the possibility to arise that the wrong solution is selected. The concept behind this goal softening approach is that if the solutions are very close, then it doesn't matter if the wrong solution is selected. This implies an explicit bias. It is for the decision maker to determine whether it is acceptable.
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes – a queuing system with N parallel servers. Buffer allocation problem.
If so, how did it perform? (Speed, optimality, ease of implementation)	Very fast convergence to the proven optimum. Little discussion about how easy it is to implement.
If not, are there any suggestions to its practical applicability?	This was a very hypothetical simulation with a proven analytical solution.
Hong, L. Jeff, and Barry L. Nelson. "Discrete optimization via simulation using COMPASS." Operations Research 54.1 (2006): 115-129.²¹⁷	
Development	
What is the method?	COMPASS
What problem was the method originally developed for?	The optimisation of discrete simulation models
Has the method been adapted from its original context?	
Was the method designed to address discrete event simulation (DES) optimisation?	No
Or is it a general optimisation method that could	Yes

be suitable for DES?	
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	It is a random research-based algorithm with a unique neighbourhood structure, termed 'the most promising area'. It is defined at each iteration to help the algorithm focus on a subset of the decision space that is of potential so to reduce the required simulation runs.
How does the method work	Initially, the most promising area is the whole decision space. At every iteration samples of a couple of solutions are made. All visited solutions are collected in a set, and proper runs are assigned by the simulation-allocation rule (SAR) for each solution in the set. As the algorithm proceeds, the set enlarges and more runs required for each solution. The sample average of each solution in the set is updated iteratively. As the sample becomes sufficiently large, the sample average approaches truth, and the algorithm can correctly select the solution that truly has best performance. For each iteration, the algorithm selects the best current solution, and the most-promising area is defined as the set of feasible solutions that are at least as close to the current best as they are to other visited solutions. Therefore as more solutions are sampled the most-promising area shrinks in size.
What assumptions does the method require?	<ol style="list-style-type: none"> 1. It requires that the sample mean of the simulation is a good estimator. If the simulation is independent and identically distributed, then strong law of large numbers applies. 2. The SAR guarantees a converge because the most promising area shrinks in size
What are the theoretical limitations of the method?	Well the method requires a stored set of solutions and require continue recalculation of sample means and variance – likely to be memory intensive.
What are the potential biases associated with the method?	None stated explicitly
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	No
If so, how did it perform? (Speed, optimality, ease of implementation)	
If not, are there any suggestions to its practical applicability?	
Huang C-J, Chang K-H, Lin JT. Optimal vehicle allocation for an Automated Materials Handling System using simulation optimisation. International Journal of Production Research. 2012;50(20):5734-46.²¹⁸	
Development	
What is the method?	An implementation of the 'Convergent Optimisation via Most-Promising-Area Stochastic Search' COMPASS method of Hong and Nelson (2004)

What problem was the method originally developed for?	The optimisation of discrete simulation models
Has the method been adapted from its original context?	No
Was the method designed to address discrete event simulation (DES) optimisation?	Yes
Or is it a general optimisation method that could be suitable for DES?	
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	It is a random research-based algorithm with a unique neighbourhood structure, termed 'the most promising area'. It is defined at each iteration to help the algorithm focus on a subset of the decision space that is of potential so to reduce the required simulation runs.
How does the method work	Initially, the most promising area is the whole decision space. At every iteration samples of a couple of solutions are made. All visited solutions are collected in a set, and proper runs are assigned by the simulation-allocation rule (SAR) for each solution in the set. As the algorithm proceeds, the set enlarges and more runs required for each solution. The sample average of each solution in the set is updated iteratively. As the sample becomes sufficiently large, the sample average approaches truth, and the algorithm can correctly select the solution that truly has best performance. For each iteration, the algorithm selects the best current spolution, and the most-promising area is defined as the set of feasible solutions that are at least as close to the current best as they are to other visited solutions. Therefore as more solutions are sampled the most-promising area shrinks in size.
What assumptions does the method require?	<ol style="list-style-type: none"> 3. It requires that the sample mean of the simulation is a good estimator. If the simulation is independent and identically distributed, then strong law of large numbers applies. 4. The SAR guarantees a converge because the most promising area shrinks in size
What are the theoretical limitations of the method?	Well the method requires a stored set of solutions and require continue recalculation of sample means and variance – likely to be memory intensive.
What are the potential biases associated with the method?	None stated explicitly
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes – a vehicle allocation problem.
If so, how did it perform? (Speed, optimality,	It is not clear from the example if a near optimal result is found, the model does not seem to converge after only 30 iterations, however it does

ease of implementation)	show continual improvement in the objective function
If not, are there any suggestions to its practical applicability?	None stated in the paper.
Jacobson SH, Sullivan KA, Johnson AW. Discrete manufacturing process design optimization using computer simulation and generalized hill climbing algorithms. Engineering Optimization. 1998;31(2):247-60.²¹⁴	
Development	
What is the method?	The whole class of stochastic generalised hill climbing methods for discrete manufacturing design optimisation problems
What problem was the method originally developed for?	For discrete manufacturing design optimisation problems
Has the method been adapted from its original context?	The paper is an overview and application of several GHC algorithms
Was the method designed to address discrete event simulation (DES) optimisation?	No
Or is it a general optimisation method that could be suitable for DES?	Yes
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	
How does the method work	The paper presents a generalised framework for 'generalised hill climbing' methods. These retain the best visited solution and allow the visiting of many inferior designs to search for a globally optimal design. Simulated annealing, threshold accepting, tabu search, monte carlo search, local search and Weibull accepting are all variations of GHC.
What assumptions does the method require?	Penalties are applied to infeasible solutions or broken constraints
What are the theoretical limitations of the method?	The paper doesn't discuss these
What are the potential biases associated with the method?	The paper doesn't discuss these
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes
If so, how did it perform? (Speed, optimality, ease of implementation)	Simulated annealing, threshold accepting, weibull accepting all yielded results superior to those by monte carlo search and local search

	A neighbourhood selection process that balances conservatism with aggression was superior.
If not, are there any suggestions to its practical applicability?	None
Jun Z, Yu-An T, Xue-Lan Z, Jun L. An improved dynamic structure-based neural networks determination approaches to simulation optimization problems. Neural Computing & Applications. 2010;19(6):883-901.²²⁷	
Development	
What is the method?	The method is a hybrid approach using an improved 'orthogonal genetic algorithm with quantization (OGA/Q)' evolutionary algorithm combined with a neural network determination approach.
What problem was the method originally developed for?	Simulation optimisation
Has the method been adapted from its original context?	Yes
Was the method designed to address discrete event simulation (DES) optimisation?	Yes
Or is it a general optimisation method that could be suitable for DES?	
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	The method uses OGA/Q to solve the simulation optimisation problem. At the same time, dynamic structure-based neural networks are applied to learn and replace the known simulation model as an auxiliary method to solve the problem.
How does the method work	<ol style="list-style-type: none"> 1. Start with the generation counter to zero 2. Set initial configurations for the simulator (population initialization) 3. Evaluate fitness of all initial configurations in population by running simulations (population evaluation) 4. Increase generation counter 5. Generate the next simulator configurations by the optimization algorithm, which uses performance measures and search techniques to decide on these configurations (population recombination) 6. Evaluate the new obtained configurations by simulations (population evaluation) 7. Test for termination criterion (number of generations, fitness) and stop or go back to 4.
What assumptions does the method require?	Evolutionary algorithms avoid the shortcomings of SO methods – in particular, sensitivity to local extrema, limitations in addressing problems with mixed numerical and no-numerical variables or high computational load.
What are the theoretical limitations of the method?	EA's are in general quite slow. The proposed method (Orthogonal genetic algorithm with quantization (OGA/Q) designs a new method for generating good initial populations and a new crossover operator.
What are the potential biases associated with	Not explicitly stated in the article.

the method?	
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	No – just test examples
If so, how did it perform? (Speed, optimality, ease of implementation)	The proposed determination approach based on OGA/Q for the DSNN (dynamic structure-based neural network) can find optimal or close-to optimal solutions quickly. The number of evaluations was drastically reduced. Numerous evaluations replaced by the evaluation of trained DSNNs. Although NN training is time-consuming, its training frequency is small.
If not, are there any suggestions to its practical applicability?	No
Kamrani F, Ayani R, Moradi F. A framework for simulation-based optimization of business process models. Simulation-Transactions of the Society for Modeling and Simulation International. 2012;88(7):852-69.²¹⁵	
Development	
What is the method?	The method(s) involve combining the Hungarian algorithm with a hill climbing heuristic method.
What problem was the method originally developed for?	Optimizing a business process model with the objective of finding the most beneficial assignment of tasks to agents. Where assignments of tasks to agents do not affect the flow of the business process, a Hungarian algorithm is applied. Where assignments of tasks to agents DO affect the workflow, depending on who performs them – then the heuristic method is applied to large problem.
Has the method been adapted from its original context?	Yes
Was the method designed to address discrete event simulation (DES) optimisation?	Yes in this case
Or is it a general optimisation method that could be suitable for DES?	These are general methods that have been combined
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	Because there are scenarios where the assignment combinations may change the probabilities that govern the path of the workflow – this means that there is not a unique ‘value added’ matrix. Instead, different assignments result in different matrices.
How does the method work	The method overcomes this by distinguishing between tasks whose assignment may affect the flow of the process, and other tasks. These are called critical and non-critical tasks. Depending on the size of the problem, either the optimal or near-optimal solution can be identified. The method employs the Hungarian algorithm to find the optimal assignment for non-critical tasks, which minimises the complexity of the algorithm. A heuristic method used in the algorithm (similar to hill climbing) is used to find the near optimal solution for the critical tasks.

What assumptions does the method require?	That you can clearly distinguish between critical and non-critical tasks.
What are the theoretical limitations of the method?	The Hungarian algorithm has a time complexity of $O(m^3)$, where m is the number of tasks. Little mention is given to the appropriateness of a hill-climbing method for the heuristic.
What are the potential biases associated with the method?	None explicitly stated.
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	The random search method was applied with three initial solutions. The algorithm reached a near optima after 80 iterations. The relative deviation for a number of problem sizes is less than 0.5% from the optimal value
If so, how did it perform? (Speed, optimality, ease of implementation)	
If not, are there any suggestions to its practical applicability?	
Korytkowski P, Wisniewski T, Rymaszewski S. An evolutionary simulation-based optimization approach for dispatching scheduling. Simulation Modelling Practice and Theory. 2013;35:69-85.²²⁸	
Development	
What is the method?	Genetic algorithm with modified genetic operations (selection and crossover)
What problem was the method originally developed for?	GA's have been used routinely for optimisation, including combinatorial simulation optimisation
Has the method been adapted from its original context?	Yes – the modification of (regularly modified) selection and crossover rules
Was the method designed to address discrete event simulation (DES) optimisation?	No
Or is it a general optimisation method that could be suitable for DES?	Yes
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	The method described is a very standard application of a GA.
How does the method work	
What assumptions does the method require?	Pre-experiments were conducted to established the correct size of the initial population (trade off between speed and avoiding premature convergence). Selection process was modified, with firstly a proportionate selection) draw from an unjust roulette wheel), and then a tournament

	<p>to avoid losing the best genetic material. The algorithm also applied an elitist strategy, so the chromosomes with the best fitness go into the next population.</p> <p>Crossover was two-point.</p> <p>Mutation probability set at 0.05</p> <p>Stop condition was when the best solution stays within a range (E) for a given population size.</p>
What are the theoretical limitations of the method?	Generic GA limitations
What are the potential biases associated with the method?	The fact that GA's in general require a lot of 'tinkering' to set up.
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes – an offset printing problem. Discrete parameters requiring a DES to evaluate
If so, how did it perform? (Speed, optimality, ease of implementation)	The algorithm reached a stop condition after 10.5 hrs (36k replications). For all 4 criteria it converged and found, after an acceptable time, a near optimal result.
If not, are there any suggestions to its practical applicability?	None noted
Kuo, R. J., and C. Y. Yang. "Simulation optimization using particle swarm optimization algorithm with application to assembly line design." Applied Soft Computing 11.1 (2011): 605-613.²³⁴	
Development	
What is the method?	Particle Swarm Optimisation
What problem was the method originally developed for?	Simulation optimisation (assembly line design)
Has the method been adapted from its original context?	No
Was the method designed to address discrete event simulation (DES) optimisation?	No
Or is it a general optimisation method that could be suitable for DES?	Yes
Theoretical basis	
How does the method address optimising a DES	Particle Swarm Optimisation) PSO has many similarities to evolutionary methods such as genetic algorithms. The method uses the simulation

with a combinatorial problem?	<p>system as the fitness function for the algorithm.</p> <p>PSO does not have an evolutionary operator in the algorithm (unlike GA). Therefore there is no crossover or mutation.</p> <p>However, this method in this study uses a modification of PSO which incorporates mutation based on similarity (PSOMS). The concept is based on similarity between the specific particle and the current global best particle in the swarm. The collectivity is used to randomly mutate the position of all the particles so to maintain diversity in the space.</p> <p>In PSO, the potential solutions (particles) move through the search space by following the currently optimum particles.</p>
How does the method work	<p>There are three global variables which are tracked:</p> <ol style="list-style-type: none"> 1. The target value or condition 2. The global best value indicating which particle's data is currently the closest to the target value or condition 3. Stopping value which indicates when the algorithm should stop <p>Each particle consists of:</p> <ol style="list-style-type: none"> 1. Data that represents a possible solution 2. A velocity value which indicates how much that possible solution can be changed 3. A personal best value indicating the closest the particle's data has come to the target <p>All these data are combined in an updating procedure to inform a new set of velocities for the particles.</p>
What assumptions does the method require?	Not clearly explained
What are the theoretical limitations of the method?	Not clearly explained. PSO has a strong theoretical background however
What are the potential biases associated with the method?	Number of particles, weight, number of epochs, inertia rate, rate of mutation. These are all user defined parameters
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes – an assembly line design problem
If so, how did it perform? (Speed, optimality, ease of implementation)	It was fully validated and tested in a good comprehensive study. There were 30 tests conducted. PSO has better regional searching ability and required approximately 150 iterations to achieve satisfactory convergence results. It was compared to GA's and the PSOMS was found to have the best problem-solving effect.
If not, are there any suggestions to its practical	None mentioned

applicability?	
Lacksonen, T., Empirical comparison of search algorithms for discrete event simulation. Computers & Industrial Engineering, 2001, 40, 133–148.²²⁴	
Development	
What is the method?	Four methods are tested in an empirical study: <ol style="list-style-type: none"> 1. Genetic algorithm 2. Pattern search 3. Simulated annealing 4. Simplex method
What problem was the method originally developed for?	These were tested on two integer problems – standard buffer problem, and distribution models.
Has the method been adapted from its original context?	No – the algorithms appear to be implementations of standard algorithms
Was the method designed to address discrete event simulation (DES) optimisation?	No
Or is it a general optimisation method that could be suitable for DES?	Yes
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	See other papers for each method
How does the method work	See other papers for each method
What assumptions does the method require?	The methods required coding of all parameters as integers. A decision regarding the initial solution is required for the SA, pattern search. Simplex requires N+1 initial solutions for N variables.
What are the theoretical limitations of the method?	Pattern search and simplex are local search techniques using one-at a time direct search. However, SA and GA are global search techniques.
What are the potential biases associated with the method?	SA and GA require tuning of several parameters (see other papers)
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes
If so, how did it perform? (Speed, optimality, ease of implementation)	Problem size had a significant effect on the accuracy on all methods apart from GA. The other three methods had poor results when the problem size grew. However, GA required significantly more replications to achieve the better results.

If not, are there any suggestions to its practical applicability?	Clear trade-off between accuracy and speed. However, although the GA is slow, it is robust and found good results for all factors tested. Good solutions with the GA were usually found within 1000 replications.
Prudius AA, Andradottir S. Averaging frameworks for simulation optimization with applications to simulated annealing. Naval Research Logistics. 2012;59(6):411-29.²³⁵	
Development	
What is the method?	Adaptive random search
What problem was the method originally developed for?	General optimisation methods
Has the method been adapted from its original context?	Yes
Was the method designed to address discrete event simulation (DES) optimisation?	Yes
Or is it a general optimisation method that could be suitable for DES?	No
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	The method is adaptive, so it used information gathered at previous iterations to decide on how much simulation effort is expended in the current iteration. Also, averaging is used
How does the method work	The method is a random search. Also presented is a random search method using point-based methods, where there is iterative movement between points within the feasible region.
What assumptions does the method require?	The method is fundamentally an adaptation of established search methods which uses the feedback of information from previous iterations to determine the next iteration.
What are the theoretical limitations of the method?	The methods are only globally convergent under mild conditions
What are the potential biases associated with the method?	None mentioned in the paper
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes. Not a fully combinatorial problem but a discrete three-stage buffer allocation problem using a discrete event simulation model.
If so, how did it perform? (Speed, optimality, ease of implementation)	Not much information is provided for this. The global algorithms perform better than the local algorithms, but they also show that averaging alone isn't necessarily beneficial. The numerical examples involving two variants of SA demonstrated that averaging alone may either help or hurt performance relative to no averaging. But that averaging together with adaptiveness in expending simulation effort appears to be effective.
If not, are there any suggestions to its practical applicability?	None mentioned.

Rosen SL, Harmonosky CM. An improved simulated annealing simulation optimization method for discrete parameter stochastic systems. Computers & Operations Research. 2005;32(2):343-58. ²²⁵	
Development	
What is the method?	SA based simulation optimisation method
What problem was the method originally developed for?	developed to improve the performance of SA for discrete variable simulation optimisation
Has the method been adapted from its original context?	Yes
Was the method designed to address discrete event simulation (DES) optimisation?	No
Or is it a general optimisation method that could be suitable for DES?	Yes – Simulated Annealing
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	<p>Simulated Annealing has an established based for discrete parameters simulation optimisation when there is a large solution space. The method has two phases. The first phase is a search process using linear approximations. Phase two is an exploration of a small subset of the feasible region around the phase 1 solution to locate a solution of possible higher quality.</p> <p>Therefore Phase 1 is possible local optima solutions, and SA works well for phase 2 because it has the ability to move away from a local optima.</p> <p>The technique of constructing linear model approximations and searching along the direction of improvement related to the linear model is similar to response surface methods. This method is a modification of RSM because it is applicable to discrete decision space.</p>
How does the method work	The main concept is to search out different high quality local optima. Starting points are generated in all areas of the feasible region, but not generated within a close area of the final phase 2 solution (which would trigger convergence at an already found solution.
What assumptions does the method require?	The model assumes that if a solution is of high quality, that other high quality solutions could exist in adjacent and nearby neighbourhoods and therefore these are searched.
What are the theoretical limitations of the method?	
What are the potential biases associated with the method?	As with any SA method, tuning is required to established the algorithm. Also a user defined termination criteria.
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes
If so, how did it perform? (Speed, optimality,	It required substantially fewer (~17%) simulation runs to optimise.

ease of implementation)	
If not, are there any suggestions to its practical applicability?	None. However, a positive is that it allows better control over the convergence rate of the algorithm by varying the number of searches completed. It can be used as either a quick local optima method, or a more thorough method to find a global optima with high probability.
L. Shi, S. Olafsson, Nested partitions method for stochastic optimization, Methodology and Computing in Applied Probability 2(3) (2000) 271–291.²⁷⁵	
Development	
What is the method?	Nested partitions
What problem was the method originally developed for?	Simulation optimisation
Has the method been adapted from its original context?	No
Was the method designed to address discrete event simulation (DES) optimisation?	Yes – optimisation of discrete stochastic problems
Or is it a general optimisation method that could be suitable for DES?	No
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	<p>Nested partitions is a global sampling strategy for optimisation of a large but finite space which is constantly adapted via the partitioning of the search space.</p> <p>The NP method naturally suits parallelisation, which is a benefit for computation.</p> <p>It naturally combines global and local search</p>
How does the method work?	<p>At each iteration of the algorithm, we assume there is a sub-region of the search space which is the ‘most promising’. This most promising region is partitioned into subregions and the entire surrounding region is aggregated into one region. Therefore at each iteration a disjoint subset of the feasible region is searched. Each region is sampled using a random sampling scheme and a ‘promising index’ is calculated for each region.</p> <p>The promising indices are used to compare all regions and identify the most promising for the next iteration.</p> <p>If one subregion is best, then it becomes the most promising region</p> <p>If the surrounding region is best, a region of less depth than the current region = most promising</p> <p>The partitioning and sampling is repeated</p>

What assumptions does the method require?	Each point in the region must have a positive probability of being selected to ensure convergence
What are the theoretical limitations of the method?	None explored in this paper
What are the potential biases associated with the method?	There are 5 important considerations: <ol style="list-style-type: none"> 1. How to partition the search space 2. How to obtain sampled points 3. How to select a 'promising index' 4. How to backtrack 5. How to select the initial region of most promise
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes – a stochastic Travelling Salesman Problem
If so, how did it perform? (Speed, optimality, ease of implementation)	Performance is relatively good for large problems.
If not, are there any suggestions to its practical applicability?	The convergence and efficiency of the NP method relies on partitioning being performed such that good solutions are clustered in a particular sub region. If this holds – the method works well.
Xu WL, Nelson BL. Empirical stochastic branch-and-bound for optimization via simulation. IIE Transactions. 2013;45(7):685-98.²³⁶	
Development	
What is the method?	The method is empirical stochastic branch and bound (ESB&B).
What problem was the method originally developed for?	'large scale complicated stochastic optimisation'
Has the method been adapted from its original context?	Yes
Was the method designed to address discrete event simulation (DES) optimisation?	Yes
Or is it a general optimisation method that could be suitable for DES?	It's a combination of nested partitioning and branch and bound
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	
How does the method work	It uses the partitioning structure of stochastic branch and bound, however uses bounds based on the performance of sampled solutions – as per nested partitions.

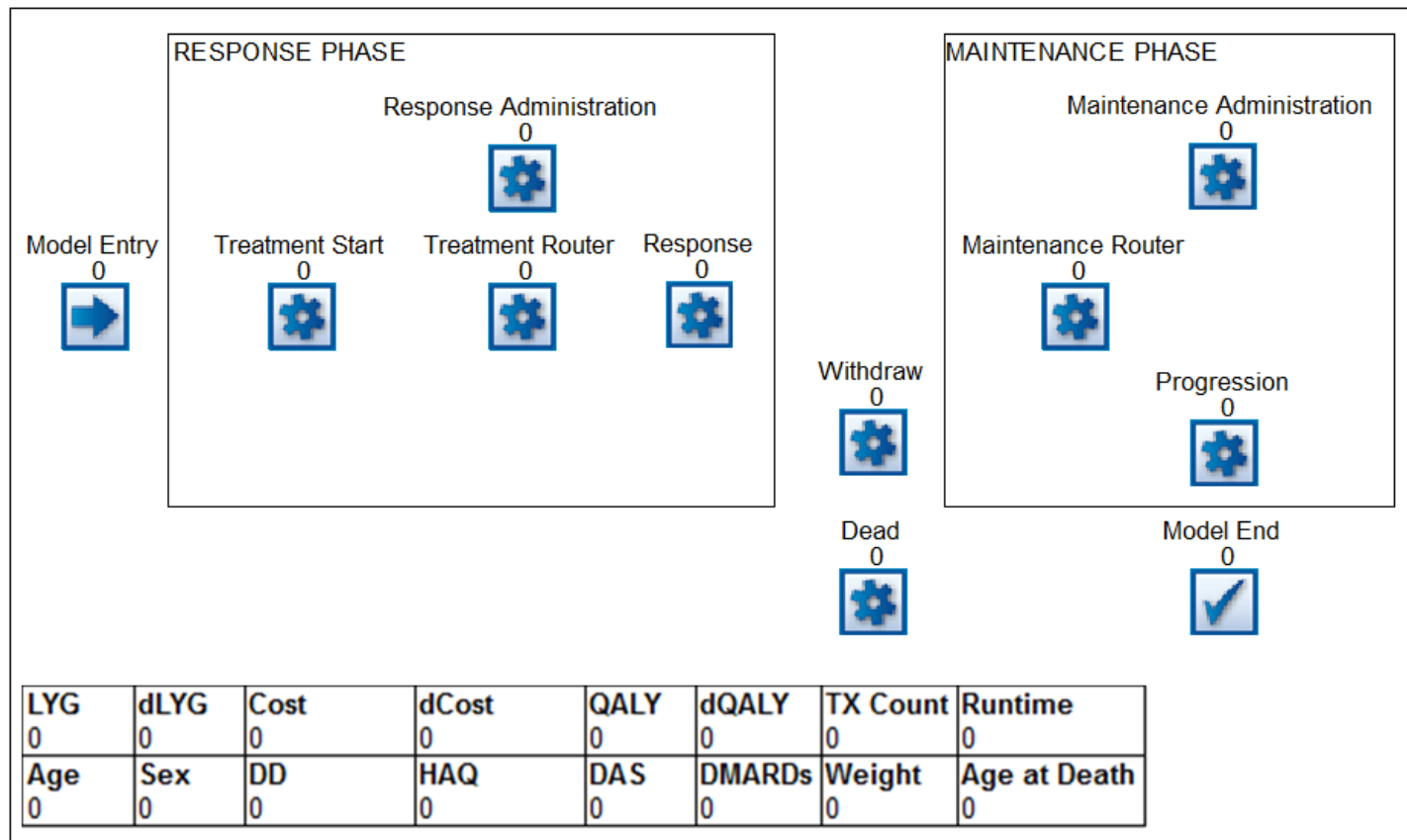
	The algorithm bounds as per NP by maintaining a set of feasible solutions. It also maintains a set of all solutions. It simulates solutions within the set of feasible solutions and computes the bounds using their estimated performance. In the next iteration, it allocates a subset of solutions to be sampled from the current partition. At each iteration, the best estimated performance is the current best solution
What assumptions does the method require?	Not fully clear
What are the theoretical limitations of the method?	There is an overhead needed to retain and refine a larger partition structure.
What are the potential biases associated with the method?	The bounding assumptions are key – these provide the ‘convergence estimators’. Sampling and simulation can be balanced by the adjusting of four parameters: <ol style="list-style-type: none"> 1. Number of samples for current best region 2. Number of samples for other regions 3. Initial number of simulations for new samples 4. Incremental number of simulations for re-sampled solutions
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes – a three-stage flow line with finite buffer storage.
If so, how did it perform? (Speed, optimality, ease of implementation)	Converges asymptotically to the global optimum. Shows that ESB&B outperforms nested partitioning in general. Advantages are maximised when the problem is noisy or significant interactions between decision variables. Normal probability based sample allocation scheme offers the most potential
If not, are there any suggestions to its practical applicability?	
Yang T, Kuo Y, Chang I. Tabu-search simulation optimization approach for flow-shop scheduling with multiple processors - a case study. International Journal of Production Research. 2004;42(19):4015-30.²³¹	
Development	
What is the method?	Tabu search simulation optimisation
What problem was the method originally developed for?	TS is a local search-based optimisation method that has been successfully applied to solve many difficult combinatorial optimisation problems
Has the method been adapted from its original context?	Not really – it seems a fairly straight forward implementation of TS to a DES model
Was the method designed to address discrete event simulation (DES) optimisation?	No

Or is it a general optimisation method that could be suitable for DES?	Yes
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	There isn't a great deal of discussion regarding the influence of a DES evaluating the objective. Only that there needs to be a balance between precision and efficiency.
How does the method work	Initial solution using a heuristic approach. Then a move is initiated (pair-wise exchange/swap) is often used as a move to construct a neighbourhood solution in a permutation-type problem (glover 1995). The neighbourhood size has an impact on the process for selecting a neighbour (e.g. searching whole neighbourhood for best solution may not be feasible). The TS algorithm searches for non-tabu moves except if it's exhausted all non-tabu options and cannot improve. If the best search result from the tabu list outperforms the best solution, then the aspiration criterion overrules the tabu rule. A long-term memory structure of intensification and diversification are used to generate the initial search sequence and restart the TS procedure.
What assumptions does the method require?	An assumption for the Tabu tenure size. This determines the tabu list for the recent past (preventing the search from repeating moves).
What are the theoretical limitations of the method?	Tabu searches are memory intensive.
What are the potential biases associated with the method?	No real discussion. There are many assumptions regarding the memory and neighbourhood selection that can lead to bias, but not discussed.
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes – flow shop problem
If so, how did it perform? (Speed, optimality, ease of implementation)	The solution is significantly better than that found by a simple descent algorithm. Also in most instances it's quicker.
If not, are there any suggestions to its practical applicability?	No
Yang T, Fu H-P, Yang K-Y. An evolutionary-simulation approach for the optimization of multi-constant work-in-process strategy - A case study. International Journal of Production Economics. 2007;107(1):104-14.²²⁹	
Development	
What is the method?	Evolutionary algorithm (using just mutation)
What problem was the method originally developed for?	For a simulation optimisation
Has the method been adapted from its original context?	Yes – from the original EA methods

Was the method designed to address discrete event simulation (DES) optimisation?	Yes
Or is it a general optimisation method that could be suitable for DES?	Yes – it's an adaptation of EAs
Theoretical basis	
How does the method address optimising a DES with a combinatorial problem?	The method requires an appropriate coding of the solution.
How does the method work	The method uses a vector to code the chromosomes. A population of solutions is generated, and roulette wheel is used for selection (not clear if elitism contained). Recombination is conducted via two randomly selecting two intersecting points and the ranges between the two points are exchanged. Mutation is via the exchange of two randomly chosen genes
What assumptions does the method require?	Little – the EA appears fairly robust to the problem type
What are the theoretical limitations of the method?	None discussed
What are the potential biases associated with the method?	None discussed. In particular, regarding the appropriate number of evaluations, and the number of algorithm runs and stopping rule.
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	Yes – a parallel-machine scheduling problem
If so, how did it perform? (Speed, optimality, ease of implementation)	Reported as effective and robust. However, the computational efficiency was a concern as the problem size increased.
If not, are there any suggestions to its practical applicability?	Only the computational efficiency. The heuristic could potentially be more efficient which would allow problems of a greater size to be solved
Practical applicability	
Has the method been used to optimise a DES with a combinatorial problem?	
If so, how did it perform? (Speed, optimality, ease of implementation)	
If not, are there any suggestions to its practical applicability?	

APPENDIX C: SIMUL8 MODEL

APPENDIX C.1: MODEL



APPENDIX C.2: FULL MODEL CODE

SIMUL8 Documentation for: C:\Users\User\Google Drive\FELLOWSHIP\MODEL\model8.s8 at time 16/05/2015 11:48:39 Version: 21.0.0.3122

Created by: Jon Tosh
Last opened by: Jon Tosh

General Simulation Information

Warm Up Time: 0 Results Collection Time: 2000 (Minutes)
Start of day: 540 Length of day: 480, Days per week: 5
Current Random Stream Set: 2
Travel Time between objects: automatically set up
Data display when simulation stopped: Work Item Count

Distributions

s_mod
Named Distribution
Distribution Detail:
Log Normal 4.85656 2.98837 1.41979 0.56665
s_good
Named Distribution
Distribution Detail:
Log Normal 9.97923 5.72325 2.15832 0.53326

Labels

tta
(Number)
haq
(Number)
ttt
(Number)
t
(Number)
ttw
(Number)

tth
(Number)
router
(Number)
tc1
(Number)
tc2
(Number)
dCost
(Number)
Cost
(Number)
timex
(Number)
tx
(Text)
tx_cost
(Number)
tx_count
(Number)
tx_tta
(Number)
ID
(Number)
ttr
(Number)
t_p
(Number)
tx_class
(Text)
admin_count
(Number)
u2
(Number)
lhaq
(Number)
c_age
(Number)
u1
(Number)

tq1
(Number)
b_age
(Number)
tq2
(Number)
prog_count
(Number)
hhaq
(Number)
QALY
(Number)
dQALY
(Number)
age_death
(Number)
loop_count
(Number)
tx_discrete
(Number)
tx_response
(Text)
t_haq
(Number)
c_dmards
(Number)
tx_r_t
(Number)
tx_good
(Number)
tx_w_t
(Number)
response_rand
(Number)
tx_s_t
(Number)
c_bdmards
(Number)
tx_mod
(Number)

loop_check
(Number)
t_haq_entry
(Number)
response_cost
(Number)
tx_s_haq
(Number)
tx_r_haq
(Number)
pre_response_haq
(Number)
gainhaq
(Number)
tx_w_haq
(Number)
failed2dmards
(Number)
t_dec
(Number)
b_dd
(Number)
LYG
(Number)
b_sex
(Number)
tx_weight
(Number)
phaq
(Number)
age_onset
(Number)
c_dd
(Number)
eular_mod
(Number)
b_haq
(Number)
dead_tx
(Number)
dLYG
(Number)

bio_prog_mod_s
(Number)
bio_prog_mod_i_age
(Number)
bio_prog_mod_i
(Number)
bio_prog_mod_i_pgen
(Number)
bio_prog_mod_i_dd
(Number)
bio_prog_mod_i_das
(Number)
bio_prog_mod_i_dmards
(Number)
bio_prog_mod_s_age
(Number)
bio_prog_mod_s_pgen
(Number)
bio_prog_mod_s_dd
(Number)
bio_prog_mod_s_das
(Number)
bio_prog_mod_s_dmards
(Number)
eular_good
(Number)
bio_prog_good_s_age
(Number)
bio_prog_good_i_age
(Number)
bio_prog_good_i
(Number)
bio_prog_good_s
(Number)
bio_prog_good_i_pgen
(Number)
bio_prog_good_i_dd
(Number)
bio_prog_good_i_das
(Number)
bio_prog_good_i_dmards
(Number)

bio_prog_good_s_pgen
(Number)
bio_prog_good_s_dd
(Number)
bio_prog_good_s_das
(Number)
bio_prog_good_s_dmards
(Number)
i_d_qaly
(Number)
b_das
(Number)
b_weight
(Number)
b_dmards
(Number)
entry_t_p
(Number)
eular_none
(Number)
i_d_cost
(Number)
dead_flag
(Number)
bio_prog_mod_xt5
(Number)
bio_prog_good_xt5
(Number)
bio_prog_good_xt2
(Number)
bio_prog_good_rho2
(Number)
bio_prog_mod_xt2
(Number)
bio_prog_mod_rho2
(Number)
bio_prog_mod_xt3
(Number)
bio_prog_mod_rho3
(Number)
bio_prog_mod_xt4
(Number)

bio_prog_mod_rho4
(Number)
bio_prog_mod_rho5
(Number)
bio_prog_good_rho5
(Number)
bio_prog_good_xt3
(Number)
bio_prog_good_rho4
(Number)
bio_prog_good_xt4
(Number)
bio_prog_good_rho3
(Number)
bio_prog_good_rho1
(Number)
bio_prog_mod_rho1
(Number)
bio_prog_mod_rho6
(Number)
bio_prog_good_rho6
(Number)
r_tx
(Number)
debug1
(Number)
tx_c
(Number)
c_das
(Number)
tx_resp
(Text)
s_first_mod_g_age
(Number)
s_first_mod_g_age2
(Number)
s_first_mod_g_pgen
(Number)
s_first_mod_g_dd
(Number)
s_first_mod_g_dd2
(Number)

s_first_mod_g_das
(Number)
s_first_mod_g_dmards
(Number)
s_first_mod_g_haq
(Number)
s_first_mod_g_cons
(Number)
s_first_mod_g Insg
(Number)
s_first_mod_g_kappa
(Number)
s_first_mod_g_sigma
(Number)
s_first_good ln_age
(Number)
s_first_good ln_age2
(Number)
s_first_good ln_pgen
(Number)
s_first_good ln_dd
(Number)
s_first_good ln_dd2
(Number)
s_first_good ln_das
(Number)
s_first_good ln_dmards
(Number)
s_first_good ln_haq
(Number)
s_first_good ln_cons
(Number)
s_first_good ln Insg
(Number)
s_first_good ln_sigma
(Number)
loopexit
(Number)
temp_haq
(Number)
testg
(Number)

tx1
(Text)
tx2
(Text)
tx3
(Text)
tx4
(Text)
tx5
(Text)
tx6
(Text)
tx7
(Text)
tx8
(Text)
tx9
(Text)
tx10
(Text)
tx11
(Text)
tx12
(Text)
tx13
(Text)
tx14
(Text)
loopcount
(Number)
block
(Number)
trip
(Number)
Images
Default Image Entry
Width: 32 Height: 32
Transparent Color: 12945930
Default Image Queue
Width: 32 Height: 32
Transparent Color: 12748546

Default Image Activity
 Width: 32 Height: 32
 Transparent Color: 12946445
 Default Image Exit
 Width: 32 Height: 32
 Transparent Color: 13078035
 Default Image Resource
 Width: 32 Height: 32
 Transparent Color: 12880393
 Default Image Conveyor
 Width: 32 Height: 32
 Transparent Color: 16777215
 Default Image Tank
 Width: 32 Height: 32
 Transparent Color: 16777215
 Default Image Rotz
 Width: 32 Height: 32
 Transparent Color: 16777215
 Default Image Process
 Width: 32 Height: 32
 Transparent Color: 16777215
 Default Image Loader
 Width: 32 Height: 32
 Transparent Color: 16777215
 Default Image Vehicle
 Width: 32 Height: 32
 Transparent Color: 16777215
 Default Image Component
 Width: 32 Height: 32
 Transparent Color: 255
 Default Image 3D Light
 Width: 32 Height: 32
 Transparent Color: 16777215
 Default Image 3D Object
 Width: 32 Height: 32
 Transparent Color: 16777215
 Redb
 Width: 8 Height: 8
 Transparent Color: 16777215
 Image 2
 Width: 32 Height: 32
 Transparent Color: 16777215

SIMUL8 Windows and Sub Processes

 Open
 Icon Location X:960 Y:585 W:32 H:32
 Window Location X:0 Y:1174 W:1902 H:960
 Color 16777215

Work Item Types

 Main Work Item Type

Image: Redb
 Length 1
 Attached Labels:
 ID
 LYG
 dLYG
 Cost
 dCost
 QALY
 dQALY
 b_age
 b_sex
 b_dd
 b_haq
 b_das
 b_dmards
 b_weight
 age_death
 tx_count
 timex
 ttd
 tta
 ttw
 tth
 ttr
 prog_count
 haq
 lhaq
 temp_haq
 hhaq
 t

tx_s_haq
 loopexit
 router
 dead_tx
 dead_flag
 r_tx
 c_age
 c_dd
 s_first_mod_g_age
 s_first_mod_g_age2
 s_first_mod_g_pgen
 s_first_mod_g_dd
 s_first_mod_g_dd2
 s_first_mod_g_das
 s_first_mod_g_dmards
 s_first_mod_g_haq
 s_first_mod_g_cons
 s_first_mod_g Insg
 s_first_mod_g_kappa
 s_first_mod_g_sigma
 s_first_good_ln_age
 s_first_good_ln_age2
 s_first_good_ln_pgen
 s_first_good_ln_dd
 s_first_good_ln_dd2
 s_first_good_ln_das
 s_first_good_ln_dmards
 s_first_good_ln_haq
 s_first_good_ln_cons
 s_first_good_ln Insg
 s_first_good_ln_sigma
 c_dmards
 testg
 tx
 tx_class
 tx_discrete
 tx_weight
 response_cost
 tx_cost
 tc2
 tc1
 i_d_cost

```

i_d_qaly
bio_prog_mod_i
bio_prog_mod_s
bio_prog_mod_i_age
bio_prog_mod_i_pgen
bio_prog_mod_i_dd
bio_prog_mod_i_das
bio_prog_mod_i_dmards
bio_prog_mod_s_age
bio_prog_mod_s_pgen
bio_prog_mod_s_dd
bio_prog_mod_s_das
bio_prog_mod_s_dmards
bio_prog_mod_rho1
bio_prog_mod_xt2
bio_prog_mod_rho2
bio_prog_mod_xt3
bio_prog_mod_rho3
bio_prog_mod_xt4
bio_prog_mod_rho4
bio_prog_mod_xt5
bio_prog_mod_rho5
bio_prog_mod_rho6
bio_prog_good_i
bio_prog_good_s
bio_prog_good_i_age
bio_prog_good_i_pgen
bio_prog_good_i_dd
bio_prog_good_i_das
bio_prog_good_i_dmards
bio_prog_good_s_age
bio_prog_good_s_pgen
bio_prog_good_s_dd
bio_prog_good_s_das
bio_prog_good_s_dmards
bio_prog_good_rho1
bio_prog_good_xt2
bio_prog_good_rho2
bio_prog_good_xt3
bio_prog_good_rho3
bio_prog_good_xt4
bio_prog_good_rho4

```

```

bio_prog_good_xt5
bio_prog_good_rho5
bio_prog_good_rho6
tx_tta
admin_count
tq2
u2
tq1
u1
tx_c
tx_mod
tx_good
response_rand
tx_response
tx_r_t
tx_w_t
c_das
tx_resp
phaq
age_onset
failed2dmards
t_p
tx_s_t
tx_r_haq
tx_w_haq
eular_none
eular_mod
eular_good
pre_response_haq
gainhaq
debug1
t_haq_entry
t_haq
entry_t_p
loop_check
loop_count
c_bdmards
t_dec
trip

```

Simulation Objects

```

-----
Start
Model Entry
-----
Display Parameters 4
X:30 Y:145 W:32 H:32
Xinc -10 Yinc 0
Show Title
Show Count
Show Image
Work Item Type: Main Work Item Type
Inter-arrival time
Distribution Detail:
Fixed 0.00001 0 0 0
Route Out Objects
Treatment Start
On Label Action Visual Logic:
VL SECTION: Model Entry Logic , LOCALDATA:
u:[NUMBER], q:[NUMBER], S:[NUMBER]
'Set Results labels = 0
'(These capture the output for each patient)
SET LYG = 0
SET dLYG = 0
SET Cost = 0
SET dCost = 0
SET QALY = 0
SET dQALY = 0
(Disabled) SET trip = 0
'HAQout is a record of HAQ progression data for the last
person to be simulated. At the start of each patient run, this sheet
is cleared first.
Clear Sheet Area haqout[2,2] , 200 , 200
'Set instantaneous discounting parameters
SET i_d_cost = LOG[1+d_cost]
SET i_d_qaly = LOG[1+d_qaly]
'Set timers (timex is for loops, t = time in model, tc1 and tq1
are 'increment' timestamps for discrete QALY and cost calcs)
SET timex = 0
SET t = 0
SET tc1 = t
SET tq1 = t

```

```

'Set any random numbers to be used so not resampled every
time
SET r_tx = RANDOM[0]
'Set the break code so that the model stops once the sim_nbr
set of patients have been run
IF Model Entry.Arrived Count = sim_nbr
  SET Model Entry.Interarrival Time = large_nbr
'Set patient characteristics at baseline
SET b_age = data[2,4+ID]
SET b_sex = data[3,4+ID]
SET b_dd = data[4,4+ID]
SET b_haq = data[5,4+ID]
SET b_das = data[6,4+ID]
SET b_weight = data[8,4+ID]
SET age_death = data[9,4+ID]
SET ttd = age_death-b_age
IF [b_age-b_dd] < 1
  SET age_onset = 1
ELSE
  SET age_onset = b_age-b_dd
SET haq = b_haq
'HAQadjust forces the HAQ value to be an eligible discrete
HAQ score
CALL haqadjust
'Set the current patient characteristics (currently equal to
baseline)
SET c_age = b_age
SET c_dd = b_dd
SET c_dmards = b_dmards
SET c_bdmards = 0
SET failed2dmards = 0
'SET initial Utility
SET u1 =
[[[0.91002]+[haq*NEG[0.20418]]+[haq^2]*NEG[0.01291]]]+[c_a
ge*NEG[0.00097]]]+[[c_age^2]*NEG[0.00001]]
'debug option
(Disabled) SET u1 = 1
'Set the treatment counter
SET tx_count = 0
'Set Biologic progression model parameters
SET bio_prog_mod_i = data[37,4+ID]
SET bio_prog_mod_s = data[38,4+ID]

```

```

SET bio_prog_mod_i_age = data[39,4+ID]
SET bio_prog_mod_i_pgen = data[40,4+ID]
SET bio_prog_mod_i_dd = data[41,4+ID]
SET bio_prog_mod_i_das = data[42,4+ID]
SET bio_prog_mod_i_dmards = data[43,4+ID]
SET bio_prog_mod_s_age = data[44,4+ID]
SET bio_prog_mod_s_pgen = data[45,4+ID]
SET bio_prog_mod_s_dd = data[46,4+ID]
SET bio_prog_mod_s_das = data[47,4+ID]
SET bio_prog_mod_s_dmards = data[48,4+ID]
SET bio_prog_mod_rho1 = data[49,4+ID]
SET bio_prog_mod_xt2 = data[50,4+ID]
SET bio_prog_mod_rho2 = data[51,4+ID]
SET bio_prog_mod_xt3 = data[52,4+ID]
SET bio_prog_mod_rho3 = data[53,4+ID]
SET bio_prog_mod_xt4 = data[54,4+ID]
SET bio_prog_mod_rho4 = data[55,4+ID]
SET bio_prog_mod_xt5 = data[56,4+ID]
SET bio_prog_mod_rho5 = data[57,4+ID]
SET bio_prog_mod_rho6 = data[58,4+ID]
SET bio_prog_good_i = data[59,4+ID]
SET bio_prog_good_s = data[60,4+ID]
SET bio_prog_good_i_age = data[61,4+ID]
SET bio_prog_good_i_pgen = data[62,4+ID]
SET bio_prog_good_i_dd = data[63,4+ID]
SET bio_prog_good_i_das = data[64,4+ID]
SET bio_prog_good_i_dmards = data[65,4+ID]
SET bio_prog_good_s_age = data[66,4+ID]
SET bio_prog_good_s_pgen = data[67,4+ID]
SET bio_prog_good_s_dd = data[68,4+ID]
SET bio_prog_good_s_das = data[69,4+ID]
SET bio_prog_good_s_dmards = data[70,4+ID]
SET bio_prog_good_rho1 = data[71,4+ID]
SET bio_prog_good_xt2 = data[72,4+ID]
SET bio_prog_good_rho2 = data[73,4+ID]
SET bio_prog_good_xt3 = data[74,4+ID]
SET bio_prog_good_rho3 = data[75,4+ID]
SET bio_prog_good_xt4 = data[76,4+ID]
SET bio_prog_good_rho4 = data[77,4+ID]
SET bio_prog_good_xt5 = data[78,4+ID]
SET bio_prog_good_rho5 = data[79,4+ID]
SET bio_prog_good_rho6 = data[80,4+ID]

```

```

'Set EULAR HAQ response changes
SET eular_none = data[81,4+ID]
SET eular_mod = data[82,4+ID]
SET eular_good = data[83,4+ID]
SET t_dec = seq[7,2]
Label Actions
ID
Unique
Financial Information
  Capital: 10
  Per Unit: 1
Carbon Emissions Information
  Carbon Footprint: 10
  Per Unit: 1
Activity
Treatment Router
-----
  Display Parameters 4
  X:231 Y:145 W:32 H:32
  Xinc -10 Yinc 0
  Show Title
  Show Count
  Show Image
Replicate 1
Priority 50
Routing In
Priority
  Route In Objects
  Response Administration
  Treatment Start
  Require resources before collecting any work items
Routing Out
Label
On label: router
Preference only
  Route Out Objects
  Response
  Dead
  Response Administration
  Release resources as soon as task complete
Operation Time

```

```

Distribution Detail:
  Fixed 0 0 0 0
On End Visual Logic:
  VL SECTION: Treatment Router Work Complete Logic ,
LOCALDATA: u:[NUMBER] , q:[NUMBER] , S:[NUMBER]
  'debug option
  (Disabled) SET tx_discrete = 0
  IF ttd = 1
    SET ttd = ttd
  'SET MODEL COSTS HERE (HAQ RELATED AND DISCRETE
TREATMENT)
  CALL haqcost
  SET timex = 0
  SET admin_count = admin_count+1
  IF admin_count = 1
    IF tx_discrete = 1
      SET tta = 0
  CALL min_response
  SET t = t+timex
  IF router = 1
    SET ttd = ttd-ttr
    SET tta = tta-ttr
    IF tx_discrete = 0
      CALL cresponsecost
  IF router = 2
    CALL haqcost
    SET dead_tx = tx_count
    SET dead_flag = 1
    SET ttd = 0
  IF router = 3
    SET ttd = ttd-tta
    SET ttr = ttr-tta
  IF t > [age_death-b_age]
    SET t = t
  (Disabled) SET tx_discrete = seq[4,tx_count+1]
Financial Information
  Capital: 10
  Per Unit: 1
Carbon Emissions Information
  Carbon Footprint: 10
  Per Unit: 1

```

```

Activity
Response
-----
  Display Parameters 4
  X:314 Y:144 W:32 H:32
  Xinc -10 Yinc 0
  Show Title
  Show Count
  Show Image
  Replicate 1
  Priority 50
  Routing In
  Priority
    Route In Objects
    Treatment Router
  Require resources before collecting any work items
  Routing Out
  Label
  On label: router
  Preference only
  Route Out Objects
  Withdraw
  Maintenance Router
  Release resources as soon as task complete
  Operation Time
  Distribution Detail:
  Fixed 0 0 0 0
  On End Visual Logic:
  VL SECTION: Response Work Complete Logic
  (Disabled) CALL qaly
  'set timers (t_p is the timer index for cDMARD progression)
  SET t_p = 0.5
  SET tth = 0.5
  SET pre_response_haq = haq
  SET response_rand = RANDOM[0]
  IF tx_class = "PALLIATIVE CARE"
    SET router = 2
    SET tx_response = "PALLIATIVE CARE"
  SET ttw = large_nbr
  ELSE
  IF response_rand < [1-tx_good]-tx_mod
    SET router = 1

```

```

  SET tx_response = "NONE"
  SET haq = haq+eular_none
  ELSE IF response_rand > 1-tx_good
    SET router = 2
    SET tx_response = "GOOD"
    SET ttw = SAMPLE["Lognormal,9.979226,5.723247"]
    SET haq = haq+eular_good
    CALL haqadjust
  ELSE
    SET router = 2
    SET tx_response = "MOD"
    SET ttw = SAMPLE["lognormal,4.856564,2.988371"]
    SET haq = haq+eular_mod
    CALL haqadjust
  'set trackers and output values
  SET tx_r_haq = haq
  SET gainhaq = pre_response_haq-haq
  SET tx_r_t = t
  SET haqout[3,3*tx_count] = tx_r_haq
  SET haqout[2,3*tx_count] = tx_r_t
  SET haqout[6,3*tx_count] = tx_response
  CALL qaly
Financial Information
  Capital: 10
  Per Unit: 1
Carbon Emissions Information
  Carbon Footprint: 10
  Per Unit: 1
End
Model End
-----
  Display Parameters 4
  X:561 Y:279 W:32 H:32
  Xinc -10 Yinc 0
  Show Title
  Show Count
  Show Image
  Input Objects
  Dead
  Financial Information
  Per Unit: 100

```

Carbon Emissions Information
Per Unit: 100

Activity
Dead

Display Parameters 4
X:396 Y:279 W:32 H:32
Xinc -10 Yinc 0
Show Title
Show Count
Show Image
Replicate 1
Priority 50
Routing In
Priority
Route In Objects
Treatment Router
Maintenance Router
Require resources before collecting any work items
Routing Out
Circulate
Preference only
Route Out Objects
Model End
Release resources as soon as task complete
Operation Time
Distribution Detail:
Fixed 0 0 0 0
On End Visual Logic:
VL SECTION: Dead Work Complete Logic , LOCALDATA:
tester:[NUMBER]
(Disabled) SET tx_count = tx_count+1
CALL qaly
SET t = t
SET ttd = ttd
SET tester = age_death-b_age
SET age_death = age_death
SET tx_s_t = tx_s_t
SET tx_r_t = tx_r_t
SET tx_w_t = tx_w_t
SET tx_r_t = 999

IF tx_w_t < tx_s_t
SET tx_w_t = 999
SET haqout[7,2] = age_death-b_age
SET haqout[8,2] = dead_tx
(Disabled) IF trip = 1
(Disabled) SET t = t-0.01
(Disabled) SET trip = 0
SET LYG = t
'debug
(Disabled) SET LYG = age_death-b_age
SET dLYG = LYG/[[1+i_d_qaly]^LYG]
SET output[1,4+ID] = ID
SET output[2,4+ID] = LYG
SET output[3,4+ID] = Cost
SET output[4,4+ID] = QALY
SET output[5,4+ID] = dLYG
SET output[6,4+ID] = dCost
SET output[7,4+ID] = dQALY
SET output[8,4+ID] = b_age
SET output[9,4+ID] = b_sex
SET output[10,4+ID] = b_dd
SET output[11,4+ID] = b_haq
SET output[12,4+ID] = b_das
SET output[13,4+ID] = b_dmards
SET output[14,4+ID] = b_weight
SET output[15,4+ID] = age_death
SET output[16,4+ID] = dead_tx
SET output[17,4+ID] = output[2,4+ID]
SET output[18,4+ID] = [age_death-b_age]
SET output[19,4+ID] = [age_death-b_age]-t
SET output[20,4+ID] = t
IF ID = sim_nbr
Sum Sheet Area output[2,5] , 1 , sim_nbr , output[2,3]
Sum Sheet Area output[3,5] , 1 , sim_nbr , output[3,3]
Sum Sheet Area output[4,5] , 1 , sim_nbr , output[4,3]
Sum Sheet Area output[5,5] , 1 , sim_nbr , output[5,3]
Sum Sheet Area output[6,5] , 1 , sim_nbr , output[6,3]
Sum Sheet Area output[7,5] , 1 , sim_nbr , output[7,3]
Sum Sheet Area output[8,5] , 1 , sim_nbr , output[8,3]
Sum Sheet Area output[9,5] , 1 , sim_nbr , output[9,3]
Sum Sheet Area output[10,5] , 1 , sim_nbr , output[10,3]
Sum Sheet Area output[11,5] , 1 , sim_nbr , output[11,3]

Sum Sheet Area output[12,5] , 1 , sim_nbr , output[12,3]
Sum Sheet Area output[13,5] , 1 , sim_nbr , output[13,3]
Sum Sheet Area output[14,5] , 1 , sim_nbr , output[14,3]
Sum Sheet Area output[15,5] , 1 , sim_nbr , output[15,3]
Sum Sheet Area output[16,5] , 1 , sim_nbr , output[16,3]
SET output[2,4] = output[2,3]/sim_nbr
SET output[3,4] = output[3,3]/sim_nbr
SET output[4,4] = output[4,3]/sim_nbr
SET output[5,4] = output[5,3]/sim_nbr
SET output[6,4] = output[6,3]/sim_nbr
SET output[7,4] = output[7,3]/sim_nbr
SET output[8,4] = output[8,3]/sim_nbr
SET output[9,4] = output[9,3]/sim_nbr
SET output[10,4] = output[10,3]/sim_nbr
SET output[11,4] = output[11,3]/sim_nbr
SET output[12,4] = output[12,3]/sim_nbr
SET output[13,4] = output[13,3]/sim_nbr
SET output[14,4] = output[14,3]/sim_nbr
SET output[15,4] = output[15,3]/sim_nbr
SET output[16,4] = output[16,3]/sim_nbr
SET m_LYG_nbr = output[2,4]
SET m_Cost_nbr = output[3,4]
SET m_QALY_nbr = output[4,4]
SET m_dLYG_nbr = output[5,4]
SET m_dCost_nbr = output[6,4]
SET m_dQALY_nbr = output[7,4]
SET m_age_nbr = output[8,4]
SET m_sex_nbr = output[9,4]
SET m_dd_nbr = output[10,4]
SET m_haq_nbr = output[11,4]
SET m_das_nbr = output[12,4]
SET m_dmards_nbr = output[13,4]
SET m_weight_nbr = output[14,4]
SET m_age_death_nbr = output[15,4]
SET m_tx_nbr = output[16,4]
Financial Information
Capital: 10
Per Unit: 1
Carbon Emissions Information
Carbon Footprint: 10
Per Unit: 1

Activity

Withdraw

Display Parameters 4
X:396 Y:204 W:32 H:32
Xinc -10 Yinc 0
Show Title
Show Count
Show Image
Replicate 1
Priority 50
Routing In
Priority
Route In Objects
Response
Maintenance Router
Require resources before collecting any work items
Routing Out
Circulate
Preference only
Route Out Objects
Treatment Start
Release resources as soon as task complete
Operation Time
Distribution Detail:
Fixed 0 0 0 0
On End Visual Logic:
VL SECTION: Withdraw Work Complete Logic
SET tx_w_t = t
SET tx_w_haq = haq

SET haqout[3,[3*tx_count]+1] = tx_w_haq
SET haqout[2,[3*tx_count]+1] = tx_w_t
SET haqout[6,3*tx_count] = tx_response
(Disabled) CALL qaly
IF tx_discrete = 0
CALL continuouscost
(Disabled) SET ttd = ttd-ttw
(Disabled) SET tx_count = tx_count+1
(Disabled) SET tx = seq[2,tx_count+2]
(Disabled) SET tx_class = seq[3,tx_count+2]
(Disabled) SET tx_discrete = seq[4,tx_count+2]

(Disabled) SET tx_weight = seq[5,tx_count+2]
(Disabled) IF tx_discrete = 1
(Disabled) IF tx = "ABT"
(Disabled) SET tx_tta = costs[4,2]
(Disabled) IF tx = "IFX"
(Disabled) SET tx_tta = costs[4,8]
(Disabled) IF tx = "RTX"
(Disabled) SET tx_tta = costs[4,9]
(Disabled) IF tx = "TCZ"
(Disabled) SET tx_tta = costs[4,10]
(Disabled) ELSE
(Disabled) SET tx_tta = large_nbr
(Disabled) ELSE
(Disabled) SET tx_tta = large_nbr
(Disabled) SET tta = tx_tta
SET haq = haq+gainhaq
CALL haqadjust
SET t = t+withdraw_time_nbr
SET ttd = ttd-withdraw_time_nbr
CALL qaly
Financial Information
Capital: 10
Per Unit: 1
Carbon Emissions Information
Carbon Footprint: 10
Per Unit: 1

Activity

Maintenance Router

Display Parameters 4
X:491 Y:145 W:32 H:32
Xinc -10 Yinc 0
Show Title
Show Count
Show Image
Replicate 1
Priority 50
Routing In
Priority
Route In Objects
Progression

Maintenance Administration

Response

Require resources before collecting any work items
Routing Out
Label
On label: router
Preference only
Route Out Objects
Progression
Withdraw
Dead
Maintenance Administration
Release resources as soon as task complete
Operation Time
Distribution Detail:
Fixed 0 0 0 0
On End Visual Logic:
VL SECTION: Maintenance Router Work Complete Logic ,
LOCALDATA: u:[NUMBER] , q:[NUMBER] , S:[NUMBER]
CALL haqcost
(Disabled) SET tx_discrete = 0
SET timex = 0
CALL min_maintenance
SET t = t+timex
IF router = 3
CALL haqcost
SET dead_tx = tx_count
SET dead_flag = 1
SET ttd = 0
IF router = 4
SET ttd = ttd-tta
SET ttw = ttw-tta
SET tth = tth-tta
IF router = 1
SET ttd = ttd-tth
SET ttw = ttw-tth
SET tta = tta-tth
IF router = 2
SET ttd = ttd-ttw
(Disabled) SET tx_discrete = seq[4,tx_count+1]
Financial Information
Capital: 10

Per Unit: 1
Carbon Emissions Information
Carbon Footprint: 10
Per Unit: 1

Activity

Progression

Display Parameters 4
X:561 Y:209 W:32 H:32
Xinc -10 Yinc 0
Show Title
Show Count
Show Image
Replicate 1
Priority 50
Routing In
Priority
Route In Objects
Maintenance Router
Require resources before collecting any work items
Routing Out
Circulate
Preference only
Route Out Objects
Maintenance Router
Release resources as soon as task complete
Operation Time
Distribution Detail:
Fixed 0 0 0 0
On End Visual Logic:
VL SECTION: Progression Work Complete Logic
CALL haqprog
CALL qaly
Financial Information
Capital: 10
Per Unit: 1
Carbon Emissions Information
Carbon Footprint: 10
Per Unit: 1

Activity

Response Administration

Display Parameters 4
X:231 Y:80 W:32 H:32
Xinc -10 Yinc 0
Show Title
Show Count
Show Image
Replicate 1
Priority 50
Routing In
Priority
Route In Objects
Treatment Router
Require resources before collecting any work items
Routing Out
Circulate
Preference only
Route Out Objects
Treatment Router
Release resources as soon as task complete
Operation Time
Distribution Detail:
Fixed 0 0 0 0
On End Visual Logic:
VL SECTION: Response Administration Work Complete Logic
CALL discretectost
(Disabled) CALL qaly
CALL haqcost
SET tta = tx_tta
Financial Information
Capital: 10
Per Unit: 1
Carbon Emissions Information
Carbon Footprint: 10
Per Unit: 1

Activity

Maintenance Administration

Display Parameters 4
X:560 Y:79 W:32 H:32

Xinc -10 Yinc 0
Show Title
Show Count
Show Image
Replicate 1
Priority 50
Routing In
Priority
Route In Objects
Maintenance Router
Require resources before collecting any work items
Routing Out
Circulate
Preference only
Route Out Objects
Maintenance Router
Release resources as soon as task complete
Operation Time
Distribution Detail:
Fixed 0 0 0 0
On End Visual Logic:
VL SECTION: Maintenance Administration Work Complete Logic
CALL discretectost
CALL haqcost
SET tta = tx_tta
Financial Information
Capital: 10
Per Unit: 1
Carbon Emissions Information
Carbon Footprint: 10
Per Unit: 1

Activity

Treatment Start

Display Parameters 4
X:127 Y:145 W:32 H:32
Xinc -10 Yinc 0
Show Title
Show Count
Show Image
Replicate 1

```

Priority 50
Routing In
Priority
  Route In Objects
  Model Entry
  Withdraw
Require resources before collecting any work items
Routing Out
Label
On label: router
Preference only
  Route Out Objects
  Treatment Router
Release resources as soon as task complete
Operation Time
Distribution Detail:
  Fixed 0 0 0 0
On End Visual Logic:
  VL SECTION: Treatment Start Work Complete Logic ,
LOCALDATA: u:[NUMBER], q:[NUMBER], S:[NUMBER]
  'Increment the treatment counter
  SET tx_count = tx_count+1
  SET admin_count = 0
  SET prog_count = 0
  SET c_dmards = c_dmards+1
  SET tx_s_haq = haq
  SET tx_s_t = t
  IF c_dmards > 2
    SET failed2dmards = 1
  'Pickup the treatment parameters and values
  IF seq[2,tx_count+1] = 1
    SET tx = "ABT"
  IF seq[2,tx_count+1] = 2
    SET tx = "ABTS"
  IF seq[2,tx_count+1] = 3
    SET tx = "ADA"
  IF seq[2,tx_count+1] = 4
    SET tx = "CTZ"
  IF seq[2,tx_count+1] = 5
    SET tx = "ETN"
  IF seq[2,tx_count+1] = 6
    SET tx = "GOL"

```

```

IF seq[2,tx_count+1] = 7
  SET tx = "HCQ"
IF seq[2,tx_count+1] = 8
  SET tx = "IFX"
IF seq[2,tx_count+1] = 9
  SET tx = "MTX"
IF seq[2,tx_count+1] = 10
  SET tx = "RTX"
IF seq[2,tx_count+1] = 11
  SET tx = "PC"
IF seq[2,tx_count+1] = 12
  SET tx = "SSZ"
IF seq[2,tx_count+1] = 13
  SET tx = "TCZ"
IF seq[2,tx_count+1] = 14
  SET tx = "TICORA"
SET tx_weight = seq[5,tx_count+1]
SET tx_discrete = seq[4,tx_count+1]
IF seq[3,tx_count+1] = 1
  SET tx_class = "bDMARD"
IF seq[3,tx_count+1] = 2
  SET tx_class = "cDMARD"
IF seq[3,tx_count+1] = 3
  SET tx_class = "PALLIATIVE CARE"
  SET tx_discrete = 0
IF tx_class = "PALLIATIVE CARE"
  SET tx_mod = 0
  SET tx_good = 0
  'Set treatment administration time (large number if a
  continuous tx)
  IF tx_discrete = 1
    IF tx = "ABT"
      SET tx_tta = costs[4,2]
    IF tx = "IFX"
      SET tx_tta = costs[4,8]
    IF tx = "RTX"
      SET tx_tta = costs[4,9]
    IF tx = "TCZ"
      SET tx_tta = costs[4,10]
  ELSE
    SET tx_tta = large_nbr
  'Set timing labels

```

```

SET tta = tx_tta
SET ttr = response_time_nbr
'Set Costs
IF tx = "ABT"
  SET response_cost = data[11,4+ID]
  SET tx_cost = data[24,4+ID]
IF tx = "ABTS"
  SET response_cost = data[12,4+ID]
  SET tx_cost = data[25,4+ID]
IF tx = "ADA"
  SET response_cost = data[13,4+ID]
  SET tx_cost = data[26,4+ID]
IF tx = "CTZ"
  SET response_cost = data[14,4+ID]
  SET tx_cost = data[27,4+ID]
IF tx = "ETN"
  SET response_cost = data[15,4+ID]
  SET tx_cost = data[28,4+ID]
IF tx = "GOL"
  SET response_cost = data[16,4+ID]
  SET tx_cost = data[29,4+ID]
IF tx = "IFX"
  SET response_cost = data[17,4+ID]
  SET tx_cost = data[30,4+ID]
IF tx = "RTX"
  SET response_cost = data[18,4+ID]
  SET tx_cost = data[31,4+ID]
IF tx = "TCZ"
  SET response_cost = data[19,4+ID]
  SET tx_cost = data[32,4+ID]
IF tx = "HCQ"
  SET response_cost = data[20,4+ID]
  SET tx_cost = data[33,4+ID]
IF tx = "MTX"
  SET response_cost = data[21,4+ID]
  SET tx_cost = data[34,4+ID]
IF tx = "SSZ"
  SET response_cost = data[22,4+ID]
  SET tx_cost = data[35,4+ID]
IF tx = "TICORA"
  SET response_cost = data[23,4+ID]
  SET tx_cost = data[36,4+ID]

```

```

IF tx_class = "bDMARD"
  SET c_bdmards = c_bdmards+1
'Set router to Treatment Router
SET router = 1
IF tx_count = 1
  IF tx = "ADA"
    SET tx_mod = data[84,4+ID]
    SET tx_good = data[92,4+ID]
  IF tx = "ETN"
    SET tx_mod = data[85,4+ID]
    SET tx_good = data[93,4+ID]
  IF tx = "GOL"
    SET tx_mod = data[86,4+ID]
    SET tx_good = data[94,4+ID]
  IF tx = "IFX"
    SET tx_mod = data[87,4+ID]
    SET tx_good = data[95,4+ID]
  IF tx = "HCQ"
    SET tx_mod = data[88,4+ID]
    SET tx_good = data[96,4+ID]
  IF tx = "MTX"
    SET tx_mod = data[89,4+ID]
    SET tx_good = data[97,4+ID]
  IF tx = "SSZ"
    SET tx_mod = data[90,4+ID]
    SET tx_good = data[98,4+ID]
  IF tx = "TICORA"
    SET tx_mod = data[91,4+ID]
    SET tx_good = data[99,4+ID]
IF tx_count = 2
  IF tx = "ABT"
    SET tx_mod = data[100,4+ID]
    SET tx_good = data[113,4+ID]
  IF tx = "ABTS"
    SET tx_mod = data[101,4+ID]
    SET tx_good = data[114,4+ID]
  IF tx = "ADA"
    SET tx_mod = data[102,4+ID]
    SET tx_good = data[115,4+ID]
  IF tx = "CTZ"
    SET tx_mod = data[103,4+ID]
    SET tx_good = data[116,4+ID]

```

```

IF tx = "ETN"
  SET tx_mod = data[104,4+ID]
  SET tx_good = data[117,4+ID]
IF tx = "GOL"
  SET tx_mod = data[105,4+ID]
  SET tx_good = data[118,4+ID]
IF tx = "IFX"
  SET tx_mod = data[106,4+ID]
  SET tx_good = data[119,4+ID]
IF tx = "RTX"
  SET tx_mod = data[107,4+ID]
  SET tx_good = data[120,4+ID]
IF tx = "TCZ"
  SET tx_mod = data[108,4+ID]
  SET tx_good = data[121,4+ID]
IF tx = "HCQ"
  SET tx_mod = data[109,4+ID]
  SET tx_good = data[122,4+ID]
IF tx = "MTX"
  SET tx_mod = data[110,4+ID]
  SET tx_good = data[124,4+ID]
IF tx = "SSZ"
  SET tx_mod = data[111,4+ID]
  SET tx_good = data[124,4+ID]
IF tx = "TICORA"
  SET tx_mod = data[112,4+ID]
  SET tx_good = data[125,4+ID]
IF tx_count > 2
  IF tx = "ABT"
    SET tx_mod = data[100,4+ID]*[1-[t_dec*c_bdmards]]
    SET tx_good = data[113,4+ID]*[1-[t_dec*c_bdmards]]
  IF tx = "ABTS"
    SET tx_mod = data[101,4+ID]*[1-[t_dec*c_bdmards]]
    SET tx_good = data[114,4+ID]*[1-[t_dec*c_bdmards]]
  IF tx = "ADA"
    SET tx_mod = data[102,4+ID]*[1-[t_dec*c_bdmards]]
    SET tx_good = data[115,4+ID]*[1-[t_dec*c_bdmards]]
  IF tx = "CTZ"
    SET tx_mod = data[103,4+ID]*[1-[t_dec*c_bdmards]]
    SET tx_good = data[116,4+ID]*[1-[t_dec*c_bdmards]]
  IF tx = "ETN"
    SET tx_mod = data[104,4+ID]*[1-[t_dec*c_bdmards]]

```

```

  SET tx_good = data[117,4+ID]*[1-[t_dec*c_bdmards]]
IF tx = "GOL"
  SET tx_mod = data[105,4+ID]*[1-[t_dec*c_bdmards]]
  SET tx_good = data[118,4+ID]*[1-[t_dec*c_bdmards]]
IF tx = "IFX"
  SET tx_mod = data[106,4+ID]*[1-[t_dec*c_bdmards]]
  SET tx_good = data[119,4+ID]*[1-[t_dec*c_bdmards]]
IF tx = "RTX"
  SET tx_mod = data[107,4+ID]*[1-[t_dec*c_bdmards]]
  SET tx_good = data[120,4+ID]*[1-[t_dec*c_bdmards]]
IF tx = "TCZ"
  SET tx_mod = data[108,4+ID]*[1-[t_dec*c_bdmards]]
  SET tx_good = data[121,4+ID]*[1-[t_dec*c_bdmards]]
IF tx = "HCQ"
  SET tx_mod = data[109,4+ID]*[1-[t_dec*c_bdmards]]
  SET tx_good = data[122,4+ID]*[1-[t_dec*c_bdmards]]
IF tx = "MTX"
  SET tx_mod = data[110,4+ID]*[1-[t_dec*c_bdmards]]
  SET tx_good = data[124,4+ID]*[1-[t_dec*c_bdmards]]
IF tx = "SSZ"
  SET tx_mod = data[111,4+ID]*[1-[t_dec*c_bdmards]]
  SET tx_good = data[124,4+ID]*[1-[t_dec*c_bdmards]]
IF tx = "TICORA"
  SET tx_mod = data[112,4+ID]*[1-[t_dec*c_bdmards]]
  SET tx_good = data[125,4+ID]*[1-[t_dec*c_bdmards]]
IF tx_mod < 0
  SET tx_mod = 0
IF tx_good < 0
  SET tx_good = 0
'Set haqout values
SET haqout[3,[3*tx_count]-1] = tx_s_haq
SET haqout[2,[3*tx_count]-1] = tx_s_t
SET haqout[4,[3*tx_count]-1] = tx
SET haqout[5,[3*tx_count]-1] = tx_class
(Disabled) 'Debug option (set common effectiveness for all
treatments)
(Disabled) SET tx_mod = 0.5
(Disabled) SET tx_good = 0.2
(Disabled) 'debug option (set common costs for all treatments)
(Disabled) SET tx_cost = 100
(Disabled) SET response_cost = 100

```

(Disabled) 'debug option (set discrete tx's to the same administration time)

(Disabled) IF tx_discrete = 1
(Disabled) SET tx_tta = 1
(Disabled) SET tta = tx_tta
(Disabled) IF tx_class = "PALLIATIVE CARE"
(Disabled) SET tx_mod = 0
(Disabled) SET tx_good = 0
(Disabled) IF tx_class = "bDMARD"
(Disabled) SET tx_mod = 0.6
(Disabled) SET tx_good = 0.3
(Disabled) IF tx_class = "cDMARD"
(Disabled) SET tx_mod = 0.4
(Disabled) SET tx_good = 0.1

Financial Information

Capital: 10
Per Unit: 1
Carbon Emissions Information
Carbon Footprint: 10
Per Unit: 1

Information Store

Simulation Time

SIMUL8 Data
Current Value 0

Warm Up Period

SIMUL8 Data
Current Value 0

Results Collection Period

SIMUL8 Data
Current Value 2000

Current Work Item

SIMUL8 Data

Current Value 0

Overhead Cost

SIMUL8 Data
Current Value 0

Overhead Revenue

SIMUL8 Data
Current Value 0

Graph Sync Interval

SIMUL8 Data
Current Value 5

large_nbr

Number
Current Value 1000000
Reset Value 1000000

sim_nbr

Number
Current Value 47
Reset Value 47

data

Spreadsheet

output

Spreadsheet

m_age_death_nbr

Number
Current Value 0
Reset Value 0

m_LYG_nbr

Number
Current Value 0
Reset Value 0

m_dLYG_nbr

Number
Current Value 0
Reset Value 0

m_Cost_nbr

Number
Current Value 0
Reset Value 0

m_dCost_nbr

Number
Current Value 0
Reset Value 0

m_QALY_nbr

Number
Current Value 0
Reset Value 0

m_dQALY_nbr

Number
Current Value 0
Reset Value 0

m_age_nbr

Number
Current Value 0
Reset Value 0

m_sex_nbr

Number
Current Value 0
Reset Value 0

m_dd_nbr

Number
Current Value 0
Reset Value 0

m_haq_nbr

Number
Current Value 0
Reset Value 0

m_das_nbr

Number
Current Value 0
Reset Value 0

m_dmards_nbr

Number
Current Value 0
Reset Value 0

m_weight_nbr

Number
Current Value 0
Reset Value 0

response_time_nbr

Number
Current Value 0.5
Reset Value 0.5

haqllook

Spreadsheet

m_tx_nbr

Number
Current Value 0
Reset Value 0

withdraw_time_nbr

Number
Current Value 0.5
Reset Value 0.5

seq

Spreadsheet

costs

Spreadsheet

d_cost

Number
Current Value 0.035
Reset Value 0.035

d_qaly

Number
Current Value 0.035
Reset Value 0.035

m_age

Number
Current Value 51.4
Reset Value 51.4

m_sex

Number
Current Value 0.73
Reset Value 0.73

m_das

Number
Current Value 6.5
Reset Value 6.5

m_dmards

Number
Current Value 0
Reset Value 0

m_dd

Number
Current Value 0.75
Reset Value 0.75

iDR_c_nbr

Number
Current Value -0.0344
Reset Value -0.0344

iDR_q_nbr

Number
Current Value -0.0344
Reset Value -0.0344

dmard_prog_i1

Number
Current Value 0.238
Reset Value 0.238

dmard_prog_i2

Number
Current Value 1.084
Reset Value 1.084

dmard_prog_i3

Number
Current Value 0.728
Reset Value 0.728

dmard_prog_i4

Number
Current Value 0
Reset Value 0

dmard_prog_s1

Number
Current Value -0.257
Reset Value -0.257

dmard_prog_s2

Number
Current Value 0.023
Reset Value 0.023

dmard_prog_s3

Number
Current Value -0.153
Reset Value -0.153

dmard_prog_s4

Number
Current Value 0
Reset Value 0

dmard_prog_q1

Number
Current Value 0.037
Reset Value 0.037

dmard_prog_q2

Number
Current Value 0.011
Reset Value 0.011

dmard_prog_q3

Number
Current Value 0.031
Reset Value 0.031

dmard_prog_q4

Number
Current Value 0
Reset Value 0

dmard_prog_c1

Number
Current Value -0.001
Reset Value -0.001

dmard_prog_c2

Number
Current Value -0.001
Reset Value -0.001

dmard_prog_c3

Number
Current Value -0.001
Reset Value -0.001

dmard_prog_c4

Number
Current Value 0
Reset Value 0

dmard_prog_int1

Number
Current Value 11.718
Reset Value 11.718

dmard_prog_int2

Number
Current Value 5.611
Reset Value 5.611

dmard_prog_int3

Number
Current Value 8.874
Reset Value 8.874

dmard_prog_age1

Number
Current Value -0.088
Reset Value -0.088

dmard_prog_age2

Number
Current Value -0.042
Reset Value -0.042

dmard_prog_age3

Number
Current Value -0.063
Reset Value -0.063

dmard_prog_sex1

Number
Current Value -1.78
Reset Value -1.78

dmard_prog_sex2

Number
Current Value -0.479
Reset Value -0.479

dmard_prog_sex3

Number
Current Value -1.195
Reset Value -1.195

dmard_prog_das1

Number
Current Value -0.731
Reset Value -0.731

dmard_prog_das2

Number
Current Value -0.287
Reset Value -0.287

dmard_prog_das3

Number
Current Value -0.474
Reset Value -0.474

dmard_prog_dd1

Number
Current Value -0.021
Reset Value -0.021

dmard_prog_dd2

Number
Current Value 0.011
Reset Value 0.011

dmard_prog_dd3

Number
Current Value -0.003
Reset Value -0.003

dprog

Spreadsheet

dmard_prog_dep1

Number
Current Value -1.333
Reset Value -1.333

dmard_prog_dep2

Number
Current Value -0.547
Reset Value -0.547

dmard_prog_dep3

Number
Current Value -0.56
Reset Value -0.56

dmard_prog_rf1

Number
Current Value -0.204
Reset Value -0.204

dmard_prog_rf2

Number
Current Value 0.11
Reset Value 0.11

dmard_prog_rf3

Number
Current Value -0.209
Reset Value -0.209

dmard_prog_acr1

Number
Current Value -1.036
Reset Value -1.036

dmard_prog_acr2

Number
Current Value -0.576
Reset Value -0.576

dmard_prog_acr3

Number
Current Value -0.753
Reset Value -0.753

dmard_prog_f2d1

Number
Current Value -2.361
Reset Value -2.361

dmard_prog_f2d2

Number
Current Value -0.181
Reset Value -0.181

dmard_prog_f2d3

```
-----
Number
Current Value -1.148
Reset Value -1.148
```

```
dmard_prog_r6m1
```

```
-----
Number
Current Value 2.543
Reset Value 2.543
```

```
dmard_prog_r6m2
```

```
-----
Number
Current Value 0.965
Reset Value 0.965
```

```
dmard_prog_r6m3
```

```
-----
Number
Current Value 1.392
Reset Value 1.392
```

```
m_dep
```

```
-----
Number
Current Value 0.49
Reset Value 0.49
```

```
m_rf
```

```
-----
Number
Current Value 0.73
Reset Value 0.73
```

```
m_acr
```

```
-----
Number
Current Value 1
Reset Value 1
```

```
m_r6m
```

```
-----
Number
Current Value 1
Reset Value 1
```

```
haqout
```

```
-----
Spreadsheet
```

```
Var86
```

```
-----
Number
Current Value 0
Reset Value 0
```

```
Var87
```

```
-----
Number
Current Value 0
Reset Value 0
```

```
start_time
```

```
-----
Number
Current Value 0
Reset Value 0
```

```
end_time
```

```
-----
Number
Current Value 0
Reset Value 0
```

```
runtime
```

```
-----
Number
Current Value 0
Reset Value 0
```

```
end
```

```
---
Spreadsheet
```

```
Start Run Visual Logic:
```

```
VL SECTION: Start Run Logic
```

```
'Warning: This code can run at any simulation time when the
user click the RUN button
```

```
'load sequence details
```

```
Get PC Clock start_time
```

```
SET sim_nbr = Results Collection Period
```

```
SET sim_nbr = sim_nbr
```

```
End Run Visual Logic:
```

```
VL SECTION: End Run Logic , LOCALDATA: u:[NUMBER] ,
q:[NUMBER], S:[NUMBER]
```

```
'Obeyed when the simulation reaches end of "Results Collection
Period"
```

```
SET Model Entry.Interarrival Time = 0.00001
```

```
Get PC Clock end_time
```

```
SET runtime = [end_time-start_time]*100000
```

```
(Disabled) Call COM Event "Finished"
```

```
SET seq[6,2] = 1
```

```
Other Visual Logic:
```

```
VL SECTION: Reset Logic
```

```
Get Throughput ????, 0, 0, ????
```

```
SET Model Entry.Interarrival Time = 0.00001
```

```
Get from EXCEL data[2,5], "[test.XLS]Export", 2, 5, 8, 100
```

```
Other Visual Logic:
```

```
VL SECTION: haqadjust , LOCALDATA: RESULT:[NUMBER]
```

```
IF haq < 0
```

```
SET haq = 0
```

```
ELSE IF haq > 3
```

```
SET haq = 3
```

```
ELSE
```

```
SET lhaq = [TRUNC[haq/0.125]*0.125]
```

```
SET hhaq = lhaq+0.125
```

```
IF RANDOM[0] >= [[haq-lhaq]/[hhaq-haq]]
```

```
SET haq = lhaq
```

```
ELSE
```

```
SET haq = hhaq
```

```
Other Visual Logic:
```

```
VL SECTION: discretecost , LOCALDATA: c:[NUMBER]
```

```
SET c = Cost
```

```
SET Cost = c+tx_cost
```

```
SET c = dCost
```



```

SET dCost = c+[tx_cost/[1+d_cost]^t]
Other Visual Logic:
VL SECTION: cresponsecost , LOCALDATA: c:[NUMBER] ,
temp:[NUMBER]
IF tx = "PC"
SET response_cost = 200*6
SET c = Cost
SET Cost = c+response_cost
SET temp = response_time_nbr/2
SET c = dCost
SET dCost = c+[response_cost/[1+d_cost]^(temp+t)]
Other Visual Logic:
VL SECTION: haqcost , LOCALDATA: haq_temp:[NUMBER] ,
c:[NUMBER] , DiscTemp1:[NUMBER] , DiscTemp2:[NUMBER]
IF haq <= 3
SET haq_temp = 2687.97
IF haq <= 2.5
SET haq_temp = 1246.26
IF haq <= 2
SET haq_temp = 523.68
IF haq <= 1.5
SET haq_temp = 364.68
IF haq <= 1
SET haq_temp = 102.54
IF haq <= 0.5
SET haq_temp = 167.41
SET tc2 = t
SET c = Cost
SET Cost = c+[haq_temp*(tc2-tc1)]
SET c = dCost
SET DiscTemp2 = EXP[iDR_c nbr*tc2]
SET DiscTemp1 = EXP[iDR_c nbr*tc1]
SET dCost = c+[[1/iDR_c nbr]*[DiscTemp2-
DiscTemp1]]*haq_temp]
SET tc1 = t
Other Visual Logic:
VL SECTION: haqprog , LOCALDATA: p1:[NUMBER] ,
p2:[NUMBER] , p3:[NUMBER] , p4:[NUMBER] , y1:[NUMBER] ,
y2:[NUMBER] , y3:[NUMBER] , y4:[NUMBER] , cdf1:[NUMBER] ,
cdf2:[NUMBER] , cdf3:[NUMBER] , cdf4:[NUMBER] ,
pdf1:[NUMBER] , pdf2:[NUMBER] , pdf3:[NUMBER] ,
pdf4:[NUMBER] , haq1:[NUMBER] , haq2:[NUMBER] ,

```

```

haq3:[NUMBER] , haq4:[NUMBER] , dm_age:[NUMBER] ,
dm_sex:[NUMBER] , dm_dd:[NUMBER] , dm_das:[NUMBER] ,
dm_dmards:[NUMBER] , i:[NUMBER] , S:[NUMBER] ,
prob1:[NUMBER] , prob2:[NUMBER] , prob3:[NUMBER] ,
prob4:[NUMBER] , Local_t_p:[NUMBER]
IF tx_class = "bDMARD"
IF tx_response = "MOD"
SET i =
[[[bio_prog_mod_i+[bio_prog_mod_i_age*dm_age]]+[bio_prog_
mod_i_pgen*dm_sex]]+[bio_prog_mod_i_dd*dm_dd]]+[bio_prog_
_mod_i_das*dm_das]]+[bio_prog_mod_i_dmards*dm_dmards]
SET S =
[[[bio_prog_mod_s+[bio_prog_mod_s_age*dm_age]]+[bio_prog_
mod_s_pgen*dm_sex]]+[bio_prog_mod_s_dd*dm_dd]]+[bio_prog_
_mod_s_das*dm_das]]+[bio_prog_mod_s_dmards*dm_dmards]
IF prog_count = 1
SET haq =
haq+[[[i+[S*bio_prog_mod_xt3]]+[bio_prog_mod_rho3*haq]]-
[[i+[S*bio_prog_mod_xt2]]+[bio_prog_mod_rho2*phaq]]]
SET tth = 0.5
ELSE IF prog_count = 2
SET haq =
haq+[[[i+[S*bio_prog_mod_xt4]]+[bio_prog_mod_rho4*haq]]-
[[i+[S*bio_prog_mod_xt3]]+[bio_prog_mod_rho3*phaq]]]
SET tth = 0.5
ELSE IF prog_count = 3
SET haq =
haq+[[[i+[S*bio_prog_mod_xt5]]+[bio_prog_mod_rho5*haq]]-
[[i+[S*bio_prog_mod_xt4]]+[bio_prog_mod_rho4*phaq]]]
SET tth = 0.5
ELSE IF prog_count = 4
SET haq =
haq+[[[i+[S*bio_prog_mod_xt5]]+[bio_prog_mod_rho6*haq]]-
[[i+[S*bio_prog_mod_xt5]]+[bio_prog_mod_rho5*phaq]]]
SET tth = 0.5
ELSE
SET tth = large_nbr
IF tx_class = "cDMARD"
SET p1 =
EXP[[[[[[[dmard_prog_int1+[dmard_prog_age1*age_onset]]+[dm
ard_prog_sex1*b_sex]]+[dmard_prog_dep1*m_dep]]+[dmard_pro
g_dd1*c_dd]]+[dmard_prog_rf1*m_rf]]+[dmard_prog_acr1*m_acr
]]+[dmard_prog_f2d1*failed2dmards]]+[dmard_prog_r6m1*m_r6
m]]]
SET p2 =
EXP[[[[[[[dmard_prog_int2+[dmard_prog_age2*age_onset]]+[dm
ard_prog_sex2*b_sex]]+[dmard_prog_dep2*m_dep]]+[dmard_pro
g_dd2*c_dd]]+[dmard_prog_rf2*m_rf]]+[dmard_prog_acr2*m_acr
]]+[dmard_prog_f2d2*failed2dmards]]+[dmard_prog_r6m2*m_r6
m]]]
SET p3 =
EXP[[[[[[[dmard_prog_int3+[dmard_prog_age3*age_onset]]+[dm

```

```

SET S =
[[[bio_prog_good_s+[bio_prog_good_s_age*dm_age]]+[bio_prog_
_good_s_pgen*dm_sex]]+[bio_prog_good_s_dd*dm_dd]]+[bio_pr
og_good_s_das*dm_das]]+[bio_prog_good_s_dmards*dm_dmard
s]
IF prog_count = 1
SET haq =
haq+[[[i+[S*bio_prog_good_xt3]]+[bio_prog_good_rho3*haq]]-
[[i+[S*bio_prog_good_xt2]]+[bio_prog_good_rho2*phaq]]]
SET tth = 0.5
ELSE IF prog_count = 2
SET haq =
haq+[[[i+[S*bio_prog_good_xt4]]+[bio_prog_good_rho4*haq]]-
[[i+[S*bio_prog_good_xt3]]+[bio_prog_good_rho3*phaq]]]
SET tth = 0.5
ELSE IF prog_count = 3
SET haq =
haq+[[[i+[S*bio_prog_good_xt5]]+[bio_prog_good_rho5*haq]]-
[[i+[S*bio_prog_good_xt4]]+[bio_prog_good_rho4*phaq]]]
SET tth = 0.5
ELSE IF prog_count = 4
SET haq =
haq+[[[i+[S*bio_prog_good_xt5]]+[bio_prog_good_rho6*haq]]-
[[i+[S*bio_prog_good_xt5]]+[bio_prog_good_rho5*phaq]]]
SET tth = 0.5
ELSE
SET tth = large_nbr
IF tx_class = "cDMARD"
SET p1 =
EXP[[[[[[[dmard_prog_int1+[dmard_prog_age1*age_onset]]+[dm
ard_prog_sex1*b_sex]]+[dmard_prog_dep1*m_dep]]+[dmard_pro
g_dd1*c_dd]]+[dmard_prog_rf1*m_rf]]+[dmard_prog_acr1*m_acr
]]+[dmard_prog_f2d1*failed2dmards]]+[dmard_prog_r6m1*m_r6
m]]]
SET p2 =
EXP[[[[[[[dmard_prog_int2+[dmard_prog_age2*age_onset]]+[dm
ard_prog_sex2*b_sex]]+[dmard_prog_dep2*m_dep]]+[dmard_pro
g_dd2*c_dd]]+[dmard_prog_rf2*m_rf]]+[dmard_prog_acr2*m_acr
]]+[dmard_prog_f2d2*failed2dmards]]+[dmard_prog_r6m2*m_r6
m]]]
SET p3 =
EXP[[[[[[[dmard_prog_int3+[dmard_prog_age3*age_onset]]+[dm

```

```

ard_prog_sex3*b_sex]]+[dmard_prog_dep3*m_dep]]+[dmard_prog
g_dd3*c_dd]]+[dmard_prog_rf3*m_rf]]+[dmard_prog_acr3*m_acr
]]+[dmard_prog_f2d3*failed2dmards]]+[dmard_prog_r6m3*m_r6
m]]
SET p4 = 1
SET prob1 = p1/[[[p1+p2]+p3]+p4]
SET prob2 = p2/[[[p1+p2]+p3]+p4]
SET prob3 = p3/[[[p1+p2]+p3]+p4]
SET prob4 = p4/[[[p1+p2]+p3]+p4]
SET haq1 = dprog[6,[[t_p*2]+3]]
SET haq2 = dprog[7,[[t_p*2]+3]]
SET haq3 = dprog[8,[[t_p*2]+3]]
SET haq4 = dprog[9,[[t_p*2]+3]]
SET t_haq_entry =
[[[prob1*haq1]+[prob2*haq2]]+[prob3*haq3]]+[prob4*haq4]
SET t_haq = t_haq_entry
SET entry_t_p = t_p
SET loop_check = 0
SET loop_count = 0
WHILE loop_check = 0
SET haq1 = dprog[6,[[t_p*2]+3]]
SET haq2 = dprog[7,[[t_p*2]+3]]
SET haq3 = dprog[8,[[t_p*2]+3]]
SET haq4 = dprog[9,[[t_p*2]+3]]
SET t_haq =
[[[prob1*haq1]+[prob2*haq2]]+[prob3*haq3]]+[prob4*haq4]
SET loop_count = loop_count+1
SET t_p = t_p+0.5
IF loop_count > [26-t_p]
SET loop_check = 1
IF [ABS[t_haq-t_haq_entry]] >= [0.125]
SET loop_check = 1
SET haq = [haq+t_haq]-t_haq_entry
SET tth = t_p-entry_t_p
IF loop_count > [26-t_p]
SET tth = large_nbr
(Disabled) SET haq = haq+0.125
(Disabled) SET tth = 3
'fix this!!!
IF tx_class = "PALLIATIVE CARE"
SET tth = [1/0.045]*0.125
SET haq = haq+0.125

```

```

CALL haqadjust
SET prog_count = prog_count+1
SET phaqa = haq
IF prog_count > 10
SET prog_count = prog_count
Other Visual Logic:
VL SECTION: continuouscost , LOCALDATA: C:[NUMBER] ,
DiscTemp1:[NUMBER] , DiscTemp2:[NUMBER]
SET C = Cost
SET Cost = C+[[tx_w_t-tx_r_t]*[12*tx_cost]]
SET C = dCost
SET DiscTemp2 = EXP[iDR_c nbr*tx_w_t]
SET DiscTemp1 = EXP[iDR_c nbr*tx_r_t]
SET dCost = C+[[[1/iDR_c nbr]*[DiscTemp2-
DiscTemp1]]*tx_cost*12]]
(Disabled) SET dCost =
C+[[[tx_c*12]*[[[1/NEG[i_d_cost]]*EXP[NEG[i_d_cost]*tx_w_t]]]-
[[1/NEG[i_d_cost]]*EXP[NEG[i_d_cost]*tx_r_t]]]]
SET admin_count = admin_count+1
Other Visual Logic:
VL SECTION: qaly , LOCALDATA: q:[NUMBER] , u:[NUMBER] ,
DiscTemp2:[NUMBER] , DiscTemp1:[NUMBER] , h:[NUMBER] ,
h2:[NUMBER] , a:[NUMBER] , a2:[NUMBER]
SET c_age = b_age+t
SET tq2 = t
SET u2 =
[[[0.91002]+[haq*NEG[0.20418]]]+[haq^2]*NEG[0.01291]]]+[c_a
ge*NEG[0.00097]]]+[[c_age^2]*NEG[0.00001]]
'debug option
(Disabled) SET u2 = 1
SET q = QALY
SET QALY = [[[[u2+u1]*[tq2-tq1]]]/2]+q
SET q = dQALY
SET DiscTemp2 = EXP[iDR_q nbr*tq2]
SET DiscTemp1 = EXP[iDR_q nbr*tq1]
SET dQALY = q+[[[1/iDR_q nbr]*[DiscTemp2-
DiscTemp1]]*[u2+u1]/2]]
'debug option
(Disabled) SET dQALY = QALY
SET u1 = u2
SET tq1 = t
Other Visual Logic:

```

```

VL SECTION: min_maintenance
'this avoids competing events with the same time, which causes
issues for routing
IF ttd = ttw
SET ttw = ttw+0.0001
IF ttd = tth
SET tth = tth+0.0001
IF ttd = tta
SET tta = tta+0.0001
IF ttw = tth
SET tth = tth+0.0001
IF ttw = tta
SET tta = tta+0.0001
IF tta = tth
SET tth = tth+0.0001
'this returns the minimum of (ttw, ttd, tth, tta)
IF [ttd < tta] & [ttd < tth] & [ttd < ttw] = 1
SET timex = ttd
SET router = 3
IF [tth < tta] & [tth < ttd] & [tth < ttw] = 1
SET timex = tth
SET router = 1
IF [tta < ttd] & [tta < tth] & [tta < ttw] = 1
SET timex = tta
SET router = 4
IF [ttw < ttd] & [ttw < tth] & [ttw < tta] = 1
SET timex = ttw
SET router = 2
Other Visual Logic:
VL SECTION: haqresponse , LOCALDATA: dm_age:[NUMBER] ,
dm_sex:[NUMBER] , dm_dd:[NUMBER] , dm_das:[NUMBER] ,
dm_dmards:[NUMBER] , i:[NUMBER] , S:[NUMBER]
'this is the haqchange given a response to a tx (cDMARD or
bDMARD)
IF prog_count = 0
SET c_age = b_age+t
SET c_dd = b_dd+t
SET c_das = b_das
SET dm_age = [c_age-m_age]/10
SET dm_sex = b_sex
SET dm_dd = c_dd-m_dd
SET dm_das = c_das-m_das

```

```

SET dm_dmards = c_dmards-m_dmards
SET phaqa = haq
IF tx_response = "GOOD"
  SET i =
  [(((bio_prog_good_i+(bio_prog_good_i*dm_age))+[bio_prog_
good_i_pgen*dm_sex])+[bio_prog_good_i_dd*dm_dd])+[bio_prog
_good_i_das*dm_das])+[bio_prog_good_i_dmards*dm_dmards]
  SET S =
  [(((bio_prog_good_s+[bio_prog_good_s_age*dm_age])+[bio_prog
_good_s_pgen*dm_sex])+[bio_prog_good_s_dd*dm_dd])+[bio_pr
og_good_s_das*dm_das])+[bio_prog_good_s_dmards*dm_dmard
s]
  SET haq =
  haq+[[[i+[S*bio_prog_good_xt2]]+[bio_prog_good_rho2*haq]]-
[[i+[S*0]]+[bio_prog_good_rho1*phaq]]]
  IF tx_response = "MOD"
    SET i =
    [(((bio_prog_mod_i+[bio_prog_mod_i_age*dm_age])+[bio_prog_
mod_i_pgen*dm_sex])+[bio_prog_mod_i_dd*dm_dd])+[bio_prog
_mod_i_das*dm_das])+[bio_prog_mod_i_dmards*dm_dmards]
    SET S =
    [(((bio_prog_mod_s+[bio_prog_mod_s_age*dm_age])+[bio_prog_
mod_s_pgen*dm_sex])+[bio_prog_mod_s_dd*dm_dd])+[bio_prog
_mod_s_das*dm_das])+[bio_prog_mod_s_dmards*dm_dmards]
    SET haq =
    haq+[[[i+[S*bio_prog_mod_xt2]]+[bio_prog_mod_rho2*haq]]-
[[i+[S*0]]+[bio_prog_mod_rho1*phaq]]]
    IF tx_response = "NONE"
      SET haq = haq
      SET tth = large_nbr
      IF tx_class = "PALLIATIVE CARE"
        SET tth = [1/0.045]*0.125
        SET haq = haq
      CALL haqadjust
      SET prog_count = prog_count+1
      SET phaqa = haq
Other Visual Logic:
VL SECTION: min_response , LOCALDATA: RESULT:[NUMBER] ,
temp1:[NUMBER]
'this avoids having competing events with the same times, which
causes issues for routing
IF ttd = ttr

```

```

SET ttr = ttr+0.0001
IF ttd = tta
  SET tta = tta+0.0001
IF ttd < 0
  SET ttd = ttd
'this returns the minimum of (ttd, ttr, tta)
IF [ttd < tta] & [ttd < ttr] = 1
  SET timex = ttd
  SET router = 2
IF [ttr < tta] & [ttr < ttd] = 1
  SET timex = ttr
  SET router = 1
IF [tta < ttd] & [tta < ttr] = 1
  SET timex = tta
  SET router = 3

```

SIMUL8 Profit Financial Information

Currency: £
Fixed Cost: 0
Fixed Revenue: 0

Carbon Emissions Information

Carbon Footprint Unit: CO2e
Fixed Carbon Footprint: 0
Fixed Carbon Offset: 0

APPENDIX D: SIMULATION OPTIMISATION MODEL

APPENDIX D.1: MODEL

(page intentionally blank)

APPENDIX D.2: MODEL CODE

CALL SIMUL8 MODEL

```
Sub runmodel(i, numsim)

Dim tabl As Range
Dim max_runs, lambda, evaluate_switch As Integer
Dim t_dec As Double
Dim s As Integer
Calculate
t_dec = Range("t_dec").Value
Set tabl = ThisWorkbook.Names("tabl").RefersToRange
max_runs = Range("max_runs").Value
lambda = Range("lambda").Value
evaluate_switch = Range("evaluate_switch").Value

Range("S20").Value = i
s = Range("S20").Value

If evaluate_switch = 1 Then

    Set MYSIMUL8 = GetObject("", "SIMUL8.S8Simulation")

    MYSIMUL8.Open "C:\model8.s8"

    line = 1

    MYSIMUL8.GDValueDouble("seq", 6, 2) = -1
    MYSIMUL8.GDValueDouble("seq", 7, 2) = t_dec

    Do Until line > 14
        tx = Worksheets("Frontsheet").Range("B35").Offset(i - 1, line - 1).Value

        If tx <> "" Then
            MYSIMUL8.GDValueDouble("seq", 2, line + 1) = Application.VLookup(tx, tabl, 11, 0) 'tx
            MYSIMUL8.GDValueDouble("seq", 3, line + 1) = Application.VLookup(tx, tabl, 8, 0) 'class
            MYSIMUL8.GDValueDouble("seq", 4, line + 1) = Application.VLookup(tx, tabl, 9, 0) 'discrete
            MYSIMUL8.GDValueDouble("seq", 5, line + 1) = Application.VLookup(tx, tabl, 10, 0) 'weight
```

```
End If

    line = line + 1
Loop

MYSIMUL8.RunSim numsim

'this makes the VBA wait until the model completes and then SIMUL8 makes the cell equal 1 on the run
end. this is the trigger for VBA to progress.
Do Until MYSIMUL8.GDValueDouble("seq", 6, 2) = 1
    Sleep (100)
Loop

Worksheets("Frontsheet").Range("Q35:Q35").Offset(s - 1, 0).Value = MYSIMUL8.GDValue("output", 6, 4)
Worksheets("Frontsheet").Range("R35:R35").Offset(s - 1, 0).Value = MYSIMUL8.GDValue("output", 7, 4)

'extra debugging outputs from SIMUL8 model
' Worksheets("Frontsheet").Range("AE35:AE35").Offset(s - 1, 0).Value = MYSIMUL8.GDValue("output", 2, 4)
' Worksheets("Frontsheet").Range("AF35:AF35").Offset(s - 1, 0).Value = MYSIMUL8.GDValue("output", 18, 5)

nmb = (MYSIMUL8.GDValue("output", 7, 4) * lambda - MYSIMUL8.GDValue("output", 6, 4))
Worksheets("Frontsheet").Range("P35:P35").Offset(s - 1, 0).Value = nmb
MYSIMUL8.Save "C:\model8.s8"
End If
```

End Sub

SIMULATION OPTIMISATION CODE

```
Private Sub CommandButton1_Click()

Dim numsim, i, max_runs, export_switch, gen_export_switch, exp_cooling_active, linear_cooling_active, _
lambda, evaluate_switch, init_sol, currentseqlength, cache_search_active, temp_algorithm_start,
email_notification, _
mintempstop, min_init_length, max_length, min_length, acceptworsecount, worsecount, restart_rule, _
restarts, programme, line, sequence_generation, repetition_scheduledstopsim, worseloop,
end_run_search, _
stopsimreason, failed_attempts, accept, min_temp_stop_active, temp_algorithm_increment,
end_run_search_active, _
```

```

max_failed_attempts, max_failed_attempts_active, evaluationsetnum, duplicate, duplicate_index,
total_i, _
overallbestseq_index, overallbestseqlength, currentruns, totalruns, percentcomplete,
lundy_cooling_active As Integer

```

```

Dim r_seq, r_seq2, start_time, end_time, eval_start_time, eval_end_time, best_cost, best_qaly,
temperature, _
initial_temperature, current_nmb, current_cost, current_qaly, best_nmb, p_accept_target, p_accept, _
best_nmb_index, exp_cooling, linear_cooling, previous_nmb, p_remove, p_add, bp_remove, _
bp_add, p_add2, p_remove2, bp_add2, bp_remove2, overall_best_nmb, lundy_cooling As Double

```

```
Dim currentseq, prevseq, bestseq, overallbestseq
```

```
Dim tabl, rng As Range
```

```
Dim tx, pc_name As String
```

```

evaluate_switch = Range("evaluate_switch").Value
init_sol = Range("init_sol").Value
bp_remove = Range("p_remove").Value
bp_add = Range("p_add").Value
bp_remove2 = Range("p_remove2").Value
bp_add2 = Range("p_add2").Value
min_init_length = Range("min_init_length").Value
sequence_generation = Range("sequence_generation").Value
numsim = Range("numsim").Value
export_switch = Range("export_switch").Value
gen_export_switch = Range("gen_export_switch").Value
lambda = Range("lambda").Value
max_runs = Range("max_runs").Value
restarts = Range("restarts").Value
initial_temperature = Range("init_temp").Value
programme = Range("programme").Value
repetition_schedule = Range("rep_sched").Value
exp_cooling = Range("exp_cooling").Value
exp_cooling_active = Range("exp_cooling_active").Value
linear_cooling = Range("linear_cooling").Value
linear_cooling_active = Range("linear_cooling_active").Value
lundy_cooling = Range("lundy_cooling").Value
lundy_cooling_active = Range("lundy_cooling_active").Value
end_run_search_active = Range("end_run_search_active").Value
end_run_search = Range("end_run_search").Value
restart_rule = Range("restart_rule").Value
email_notification = Range("email_notification").Value

```

```
pc_name = Range("pc_name").Value
```

```

min_temp_stop = Range("min_temp_stop").Value
min_temp_stop_active = Range("min_temp_stop_active").Value
max_failed_attempts = Range("max_failed_attempts").Value
max_failed_attempts_active = Range("max_failed_attempts_active").Value
cache_search_active = Range("cache_search_active").Value
p_accept_target = Range("p_accept_target").Value
temp_algorithm_increment = Range("temp_algorithm_increment").Value
temp_algorithm_start = Range("temp_algorithm_start").Value
worseloop = Range("worseloop").Value
totalruns = max_runs * restarts

```

```
tabl = ThisWorkbook.Names("tabl").RefersToRange
```

```

max_length = 14
min_length = 1
duplicate = 0
acceptworsecount = 0
worsecount = 0
p_accept = 0
overall_best_nmb = 0
total_i = 1
'prevrestartbestseq_index = 0
'overallbestseq_index = 0
'Range("V20:V21").ClearContents
If programme = 5 Then
    max_runs = 10000
    Range("X23:X32").ClearContents
    end_run_search_active = 0
End If

```

```

If lundy_cooling_active = 1 Then
    repetition_schedule = 1
End If

```

```

Calculate
ReDim currentseq(0 To 0)

```

```

If evaluate_switch = 1 Then      'Debug option to turn off evaluation
    Set MYSIMUL8 = GetObject("", "SIMUL8.S8Simulation")
End If

```

```

Range("B35:AJ10034").ClearContents
Range("B35:AJ10034").Interior.ColorIndex = 0
Range("B35:AJ10034").Font.Bold = False
Range("B23:U33").ClearContents
Range("U20").ClearContents
Worksheets("Cache").Range("B2:R10002").ClearContents
Worksheets("R1").Range("B2:AJ10002").ClearContents
Worksheets("R2").Range("B2:AJ10002").ClearContents
Worksheets("R3").Range("B2:AJ10002").ClearContents
Worksheets("R4").Range("B2:AJ10002").ClearContents
Worksheets("R5").Range("B2:AJ10002").ClearContents
Worksheets("R6").Range("B2:AJ10002").ClearContents
Worksheets("R7").Range("B2:AJ10002").ClearContents
Worksheets("R8").Range("B2:AJ10002").ClearContents
Worksheets("R9").Range("B2:AJ10002").ClearContents
Worksheets("R10").Range("B2:AJ10002").ClearContents

evaluationsetnum = 1000 - Application.WorksheetFunction.CountBlank( _
    Worksheets("Evaluation Set").Range("A2:A1002")) + 1
start_time = Now()

If gen_export_switch = 1 Then      'Generate a fresh patient dataset
    Call gen_export
End If

If exp_cooling_active + linear_cooling_active = 2 Then 'check a cooling schedule is selected
    MsgBox "ERROR: Define the cooling schedule properly"
    End
End If

If exp_cooling_active + lundy_cooling_active = 2 Then 'check a cooling schedule is selected
    MsgBox "ERROR: Define the cooling schedule properly"
    End
End If

If lundy_cooling_active + linear_cooling_active = 2 Then 'check a cooling schedule is selected
    MsgBox "ERROR: Define the cooling schedule properly"
    End
End If

If lundy_cooling_active + linear_cooling_active + exp_cooling_active = 3 Then 'check a cooling schedule is
selected

```

```

MsgBox "ERROR: Define the cooling schedule properly"
End
End If

If export_switch = 1 Then      'export the dataset to SIMUL8
    Dim length, width, L, W, x, y As Integer
    L = 1
    W = 1
    x = 1
    y = 1
    length = numsim
    width = 124

    MYSIMUL8.Open "C:\model8.s8"

    Do Until x > width
        Do Until L > length
            W = L
            y = x
            MYSIMUL8.GDValueDouble("data", x + 1, L + 4) = _
                Worksheets("Export").Cells(W + 4, y + 1).Value
            Application.StatusBar = _
                "Row " & L & " out of " & length & _
                ". Column " & x & " out of " & width
            L = L + 1
            Calculate
        Loop
        x = x + 1
        L = 1
    Loop

    MYSIMUL8.Save "C:\model8.s8"

End If

temploop = 0
b = 1

Do While b <= restarts      '##### main 'restart' loop
    Range("B35:AJ10034").ClearContents
    Range("B35:AJ10034").Interior.ColorIndex = 0
    Range("B35:AJ10034").Font.Bold = False

```



```

If programme <> 5 Then
  If mintempstop > initial_temperature Then 'check that the initial temp > min temp
    MsgBox "ERROR: Minimum Temperature value > Initial Temperature."
  End
End If
End If

i = 0

If programme = 4 Then 'Runs a set of sequences as defined in "Evaluation Set"
  Worksheets("Evaluation Set").Range("O2:Q20002").ClearContents

  Do
    i = i + 1
    total_i = total_i + 1

    If Worksheets("Evaluation Set").Range("A2:A2").Offset(i - 1, 0) = "" Then
      end_time = Now()
      Range("S21").Value = SecondsToDateTimeSerial(DateDiff("s", start_time, end_time))
      i = i - 1
      MsgBox i & " Evaluations. Model Run Time = " _
        & SecondsToDateTimeSerial(DateDiff("s", start_time, end_time))

      Exit Do
    Else
      Worksheets("Frontsheet").Range("B21:O21").Value = _
        Worksheets("Evaluation Set").Range("A2:N2").Offset(i - 1, 0).Value
      Application.StatusBar = "Running evaluation programme: Iteration " _
        & i & " out of " & evaluationsetnum
      insert_length = 14 - Application.WorksheetFunction.CountBlank(Range("B21:O21"))
      ReDim currentseq(1 To insert_length)
      currentseqlength = insert_length
      a = 1
      Do Until a > insert_length
        currentseq(a) = Range("B21").Offset(0, a - 1).Value
        Range("B35").Offset(i - 1, a - 1).Value = currentseq(a)
        a = a + 1
      Loop
      Call eligible(currentseq, currentseqlength, seqfault)
    End If
  End Do
End If

```

```

duplicate = 0
duplicate_index = 0
eval_start_time = 0
eval_end_time = 0

If total_i > 1 Then
  If cache_search_active = 1 Then
    eval_start_time = Now()
    Call cache_search(total_i, i, duplicate, duplicate_index, currentseqlength)
    eval_end_time = Now()
  End If
End If

If duplicate = 1 Then
  Worksheets("Frontsheet").Range("P35:P35").Offset(i - 1, 0).Value =
Worksheets("Cache").Range("P2").Offset(duplicate_index - 1, 0).Value
  Worksheets("Frontsheet").Range("Q35:Q35").Offset(i - 1, 0).Value =
Worksheets("Cache").Range("Q2").Offset(duplicate_index - 1, 0).Value
  Worksheets("Frontsheet").Range("R35:R35").Offset(i - 1, 0).Value =
Worksheets("Cache").Range("R2").Offset(duplicate_index - 1, 0).Value
  Worksheets("Frontsheet").Range("AE35:AE35").Offset(i - 1, 0).Value = "DUPLICATE"
  Worksheets("Frontsheet").Range("AF35:AF35").Offset(i - 1, 0).Value =
SecondsToDateTimeSerial(DateDiff("s", eval_start_time, eval_end_time))
Else
  eval_start_time = Now()
  Call runmodel(i, numsim) 'run model to evaluate
  If Worksheets("Frontsheet").Range("AC35:AC35").Offset(i - 1, 0).Value = 1 Then
    nmb = -999999999
    Worksheets("Frontsheet").Range("P35:P35").Offset(i - 1, 0).Value = nmb
  End If

  eval_end_time = Now()
  Worksheets("Frontsheet").Range("AE35:Ae35").Offset(i - 1, 0).Value = "EVALUATE"
  Worksheets("Frontsheet").Range("AF35:AF35").Offset(i - 1, 0).Value =
SecondsToDateTimeSerial(DateDiff("s", eval_start_time, eval_end_time))
End If

Worksheets("Evaluation Set").Range("O2:Q2").Offset(i - 1, 0).Value = _
Worksheets("Frontsheet").Range("P35:R35").Offset(i - 1, 0).Value

Loop
End 'Once the programme is run, exit the simulation.

```

```

End If

i = 1

stopsim = 0
failed_attempts = 0
acceptworsecount = 0
Do While stopsim = 0      '##### main 'iteration' loop

    Calculate

    Range("S21").Value = SecondsToDateTimeSerial(DateDiff("s", start_time, Now()))

    currentruns = ((b - 1) * max_runs) + i
    percentcomplete = Int(currentruns / totalruns * 100)

    Application.StatusBar = "Restart " & b & " out of " _
        & restarts & ": Evaluation " & i & " out of " & max_runs & " (" & percentcomplete & "%)"

    If programme = 5 Then
        If i = 1 Then
            temperature = temp_algorithm_start
            p_accept = 0
            prevseqlength = currentseqlength
        End If
    Else
        If i = 1 Then      'set temperature
            temperature = initial_temperature
            prevseqlength = currentseqlength
        Else
            temperature = temperature
        End If

        If temploop = repetition_schedule Then 'cooling schedule temp reduction

            If exp_cooling_active = 1 Then
                temperature = temperature * exp_cooling
                temploop = 0
            End If

            If linear_cooling_active = 1 Then
                temperature = temperature - linear_cooling

```

```

        temploop = 0
    End If

    If lundy_cooling_active = 1 Then
        temperature = temperature / (1 + (lundy_cooling * temperature))
        temploop = 0
    End If

    End If
    If temperature < 0 Then
        MsgBox "MAJOR ERROR: TEMPERATURE NEGATIVE"
    End
    End If
End If

temploop = temploop + 1

If i = 1 Then      '##### FIRST ITERATION
    If b = 1 Then      'First Restart
        If init_sol = 1 Then      'Random Seq to start
            Call gen_random_seq(i, currentseq, currentseqlength)
        Else
            'or insert Seq to start
            '(to check performance with same starting sequence)
            insert_length = 14 - Application.WorksheetFunction.CountBlank(Range("B21:O21"))
            ReDim currentseq(1 To insert_length)
            currentseqlength = insert_length
            a = 1
            Do Until a > insert_length
                currentseq(a) = Range("B21").Offset(0, a - 1).Value
                Range("B35").Offset(i - 1, a - 1).Value = currentseq(a)
                a = a + 1
            Loop
            Call eligible(currentseq, currentseqlength, seqfault)
            If seqfault = 1 Then
                MsgBox "MAJOR ERROR: INITIAL SEQUENCE NOT ELIGIBLE"
            End
            End If
        End If
    Else
        'next restarts
        Select Case restart_rule
            Case 1      'random restart
                If init_sol = 1 Then      'Random Seq to start

```

```

    Call gen_random_seq(i, currentseq, currentseqlength)
Else
    'or insert Seq to start
    '(to check performance with same starting sequence)
    insert_length = 14 - Application.WorksheetFunction.CountBlank(Range("B21:O21"))
    ReDim currentseq(1 To insert_length)
    currentseqlength = insert_length
    a = 1
    Do Until a > insert_length
        currentseq(a) = Range("B21").Offset(0, a - 1).Value
        Range("B35").Offset(i - 1, a - 1).Value = currentseq(a)
        a = a + 1
    Loop
    Call eligible(currentseq, currentseqlength, seqfault)
    If seqfault = 1 Then
        MsgBox "MAJOR ERROR: INITIAL SEQUENCE NOT ELIGIBLE"
    End
    End If
End If
Case 2
    'restart with overall best seq
    insert_length = 14
Application.WorksheetFunction.CountBlank(Range("B23:O23").Offset(overallbestseq_index - 1, 0))
    ReDim currentseq(1 To insert_length)
    currentseqlength = insert_length
    a = 1
    Do Until a > insert_length
        currentseq(a) = Range("B23").Offset(overallbestseq_index - 1, a - 1).Value
        Range("B35").Offset(0, a - 1).Value = currentseq(a)
        a = a + 1
    Loop
    Call eligible(currentseq, currentseqlength, seqfault)
    If seqfault = 1 Then
        MsgBox "MAJOR ERROR: INITIAL SEQUENCE NOT ELIGIBLE"
    End
    End If
End Select
prevseqlength = currentseqlength
End If

Else
    '##### SUBSEQUENT ITERATIONS
    If sequence_generation = 1 Then 'Generate a random sequence
        Call gen_random_seq(i, currentseq, currentseqlength)
    ElseIf sequence_generation = 2 Then 'Pairwise swap

```

```

'Call pairswap(i, prevseq, currentseq, currentseqlength, prevseqlength)

Select Case currentseqlength
    'addition/remove 1 or 2 tx,
    'dependent on probs & length of sequence
Case 1 To 2
    p_add = 1
    p_add2 = bp_add2
    p_remove = 0
    p_remove2 = 0
    r_seq = Rnd
    r_seq2 = Rnd
    If r_seq > 1 - p_remove Then
        If r_seq2 < p_remove2 / p_remove Then
            Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
        Else
            Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
        End If
    ElseIf r_seq < p_add Then
        If r_seq2 < p_add2 / p_add Then
            Call addseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
        Else
            Call addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
        End If
    Else
        Call pairswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
Case 3
    p_add = bp_add
    p_add2 = bp_add2
    p_remove = bp_remove
    p_remove2 = 0
    r_seq = Rnd
    r_seq2 = Rnd
    If r_seq > 1 - p_remove Then
        If r_seq2 < p_remove2 / p_remove Then
            Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
        Else
            Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
        End If
    ElseIf r_seq < p_add Then
        If r_seq2 < p_add2 / p_add Then
            Call addseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)

```

```

Else
    Call addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
End If
Else
    Call pairswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
End If
Case 4 To 10
    p_add = bp_add
    p_add2 = bp_add2
    p_remove = bp_remove
    p_remove2 = bp_remove2
    r_seq = Rnd
    r_seq2 = Rnd
    If r_seq > 1 - p_remove Then
        If r_seq2 < p_remove2 / p_remove Then
            Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
        Else
            Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
        End If
    ElseIf r_seq < p_add Then
        If r_seq2 < p_add2 / p_add Then
            Call addseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
        Else
            Call addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
        End If
    Else
        Call pairswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
Case 11 To 12
    p_add = bp_add
    p_add2 = 0
    p_remove = bp_remove
    p_remove2 = bp_remove2
    r_seq = Rnd
    r_seq2 = Rnd
    If r_seq > 1 - p_remove Then
        If r_seq2 < p_remove2 / p_remove Then
            Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
        Else
            Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
        End If
    ElseIf r_seq < p_add Then

```

```

        If r_seq2 < p_add2 / p_add Then
            Call addseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
        Else
            Call addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
        End If
    Else
        Call pairswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
Case 13 To 14
    p_add = 0
    p_add2 = 0
    p_remove = bp_remove
    p_remove2 = bp_remove2
    r_seq = Rnd
    r_seq2 = Rnd
    If r_seq > 1 - p_remove Then
        If r_seq2 < p_remove2 / p_remove Then
            Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
        Else
            Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
        End If
    Else
        Call pairswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
End Select
Elseif sequence_generation = 3 Then 'Random swap + Add/Subtract (DEFAULT)
    Select Case currentseqlength 'addition/remove 1 or 2 tx,
        'dependent on probs & length of sequence
    Case 1 To 2
        p_add = 1
        p_add2 = bp_add2
        p_remove = 0
        p_remove2 = 0
        r_seq = Rnd
        r_seq2 = Rnd
        If r_seq > 1 - p_remove Then
            If r_seq2 < p_remove2 / p_remove Then
                Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
            Else
                Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
            End If
        ElseIf r_seq < p_add Then

```

```

If r_seq2 < p_add2 / p_add Then
  Call addseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
Else
  Call addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
End If
Else
  Call randomswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
End If
Case 3
p_add = bp_add
p_add2 = bp_add2
p_remove = bp_remove
p_remove2 = 0
r_seq = Rnd
r_seq2 = Rnd
If r_seq > 1 - p_remove Then
  If r_seq2 < p_remove2 / p_remove Then
    Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
  Else
    Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
  End If
Elseif r_seq < p_add Then
  If r_seq2 < p_add2 / p_add Then
    Call addseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
  Else
    Call addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
  End If
Else
  Call randomswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
End If
Case 4 To 10
p_add = bp_add
p_add2 = bp_add2
p_remove = bp_remove
p_remove2 = bp_remove2
r_seq = Rnd
r_seq2 = Rnd
If r_seq > 1 - p_remove Then
  If r_seq2 < p_remove2 / p_remove Then
    Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
  Else
    Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)

```

```

End If
Elseif r_seq < p_add Then
  If r_seq2 < p_add2 / p_add Then
    Call addseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
  Else
    Call addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
  End If
Else
  Call randomswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
End If
Case 11 To 12
p_add = bp_add
p_add2 = 0
p_remove = bp_remove
p_remove2 = bp_remove2
r_seq = Rnd
r_seq2 = Rnd
If r_seq > 1 - p_remove Then
  If r_seq2 < p_remove2 / p_remove Then
    Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
  Else
    Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
  End If
Elseif r_seq < p_add Then
  If r_seq2 < p_add2 / p_add Then
    Call addseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
  Else
    Call addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
  End If
Else
  Call randomswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
End If
Case 13 To 14
p_add = 0
p_add2 = 0
p_remove = bp_remove
p_remove2 = bp_remove2
r_seq = Rnd
r_seq2 = Rnd
If r_seq > 1 - p_remove Then
  If r_seq2 < p_remove2 / p_remove Then
    Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)

```

```

Else
    Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
End If
Else
    Call randomswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
End If
End Select
End If
End If

duplicate = 0
duplicate_index = 0
eval_start_time = 0
eval_end_time = 0

If total_i > 1 Then
    If cache_search_active = 1 Then
        eval_start_time = Now()
        Call cache_search(total_i, i, duplicate, duplicate_index, currentseqlength)
        eval_end_time = Now()
    End If
End If

If duplicate = 1 Then
    Worksheets("Frontsheet").Range("P35:P35").Offset(i - 1, 0).Value =
Worksheets("Cache").Range("P2").Offset(duplicate_index - 1, 0).Value
    Worksheets("Frontsheet").Range("Q35:Q35").Offset(i - 1, 0).Value =
Worksheets("Cache").Range("Q2").Offset(duplicate_index - 1, 0).Value
    Worksheets("Frontsheet").Range("R35:R35").Offset(i - 1, 0).Value =
Worksheets("Cache").Range("R2").Offset(duplicate_index - 1, 0).Value
    Worksheets("Frontsheet").Range("AE35:AE35").Offset(i - 1, 0).Value = "DUPLICATE"
    Worksheets("Frontsheet").Range("AF35:AF35").Offset(i - 1, 0).Value =
SecondsToDateTimeSerial(DateDiff("s", eval_start_time, eval_end_time))
    Worksheets("Frontsheet").Range("AG35:AG35").Offset(i - 1, 0).Value = duplicate_index
Else
    eval_start_time = Now()
    Range("U20").Value = total_i
    Call runmodel(i, numsim) 'run model to evaluate
    If Worksheets("Frontsheet").Range("AC35:AC35").Offset(i - 1, 0).Value = 1 Then
        nmb = -999999999
        Worksheets("Frontsheet").Range("P35:P35").Offset(i - 1, 0).Value = nmb
    End If

```

```

eval_end_time = Now()
Worksheets("Frontsheet").Range("AE35:Ae35").Offset(i - 1, 0).Value = "EVALUATE"
Worksheets("Frontsheet").Range("AF35:AF35").Offset(i - 1, 0).Value =
SecondsToDateTimeSerial(DateDiff("s", eval_start_time, eval_end_time))
End If

current_nmb = Worksheets("Frontsheet").Range("P35:P35").Offset(i - 1, 0).Value
current_cost = Worksheets("Frontsheet").Range("Q35:Q35").Offset(i - 1, 0).Value
current_qaly = Worksheets("Frontsheet").Range("R35:R35").Offset(i - 1, 0).Value

Worksheets("Cache").Range("B2:R2").Offset(total_i - 1, 0).Value =
Worksheets("Frontsheet").Range("B35:R35").Offset(i - 1, 0).Value

If i = 1 Then 'update current best results
    best_nmb = current_nmb
    best_cost = current_cost
    best_qaly = current_qaly
    previous_nmb = current_nmb
    bestseq = currentseq
    Worksheets("Frontsheet").Range("T35:T35").Value = best_nmb
    best_nmb_index = i
    Worksheets("Frontsheet").Range("S35:S35").Value = best_nmb_index
Else
    If current_nmb > best_nmb Then 'improvement. update current best
        best_nmb = current_nmb
        best_cost = current_cost
        best_qaly = current_qaly
        bestseq = currentseq

        Worksheets("Frontsheet").Range("T35:T35").Offset(i - 1, 0).Value = best_nmb
        best_nmb_index = i
        Worksheets("Frontsheet").Range("S35:S35").Offset(i - 1, 0).Value = i
    Else 'otherwise retain the current best
        Worksheets("Frontsheet").Range("T35:T35").Offset(i - 1, 0).Value = best_nmb
        Worksheets("Frontsheet").Range("S35:S35").Offset(i - 1, 0).Value = best_nmb_index
    End If
End If

Worksheets("Frontsheet").Range("W35:W35").Offset(i - 1, 0).Value = temperature

If temperature = 0 Then

```

```

    temperature = 1
End If

If programme = 1 Then      'random search (every sequence accepted)
    If i = 1 Then
        prevseq = currentseq
        previous_nmb = current_nmb
        prevseqlength = currentseqlength
    Else
        If current_nmb > previous_nmb Then
            Worksheets("Frontsheet").Range("U35:U35").Offset(i - 1, 0).Value = 1
            Worksheets("Frontsheet").Range("V35:V35").Offset(i - 1, 0).Value = 1
            prevseq = currentseq
            prevseqlength = currentseqlength
            previous_nmb = current_nmb
        Else
            Worksheets("Frontsheet").Range("U35:U35").Offset(i - 1, 0).Value = 0
            Worksheets("Frontsheet").Range("V35:V35").Offset(i - 1, 0).Value = 1
            prevseq = currentseq
            prevseqlength = currentseqlength
            previous_nmb = current_nmb
            acceptworsecount = acceptworsecount + 1
        End If
        prevseq = currentseq
    End If
ElseIf programme = 2 Then  '##### SIMULATED ANNEALING
    If i = 1 Then
        prevseq = currentseq
        previous_nmb = current_nmb
        prevseqlength = currentseqlength
    Else
        If current_nmb > previous_nmb Then      'Improvement = Accept
            Worksheets("Frontsheet").Range("U35:U35").Offset(i - 1, 0).Value = 1
            Worksheets("Frontsheet").Range("V35:V35").Offset(i - 1, 0).Value = 1
            prevseq = currentseq
            prevseqlength = currentseqlength
            previous_nmb = current_nmb
            failed_attempts = 0
        Else
            r_sa = Rnd
            If r_sa < Exp((current_nmb - previous_nmb) / temperature) Then      'Accept
                Worksheets("Frontsheet").Range("U35:U35").Offset(i - 1, 0).Value = 0

```

```

                Worksheets("Frontsheet").Range("V35:V35").Offset(i - 1, 0).Value = 1
                prevseq = currentseq
                prevseqlength = currentseqlength
                previous_nmb = current_nmb
                failed_attempts = 0
                acceptworsecount = acceptworsecount + 1
            Else
                Worksheets("Frontsheet").Range("U35:U35").Offset(i - 1, 0).Value = 0      'Reject
                Worksheets("Frontsheet").Range("V35:V35").Offset(i - 1, 0).Value = 0
                prevseq = prevseq
                prevseqlength = prevseqlength
                previous_nmb = previous_nmb
                failed_attempts = failed_attempts + 1
            End If
        End If
    End If
ElseIf programme = 3 Then  'Local search (only improving moves accepted)
    If i = 1 Then
        prevseq = currentseq
        previous_nmb = current_nmb
        prevseqlength = currentseqlength
    Else
        If current_nmb > previous_nmb Then      'Improvement = Accept
            Worksheets("Frontsheet").Range("U35:U35").Offset(i - 1, 0).Value = 1
            Worksheets("Frontsheet").Range("V35:V35").Offset(i - 1, 0).Value = 1
            prevseq = currentseq
            prevseqlength = currentseqlength
            previous_nmb = current_nmb
            failed_attempts = 0
        Else
            'Reject
            Worksheets("Frontsheet").Range("U35:U35").Offset(i - 1, 0).Value = 0
            Worksheets("Frontsheet").Range("V35:V35").Offset(i - 1, 0).Value = 0
            prevseq = prevseq
            prevseqlength = prevseqlength
            previous_nmb = previous_nmb
            failed_attempts = failed_attempts + 1
        End If
    End If
ElseIf programme = 5 Then  'INITIAL TEMP SETTING ALGORITHM
    If i = 1 Then
        prevseq = currentseq
        previous_nmb = current_nmb

```

```

    prevseqlength = currentseqlength
Else
    If current_nmb > previous_nmb Then 'Improvement = Accept
        Worksheets("Frontsheet").Range("U35:U35").Offset(i - 1, 0).Value = 1
        Worksheets("Frontsheet").Range("V35:V35").Offset(i - 1, 0).Value = 1
        prevseq = currentseq
        prevseqlength = currentseqlength
        previous_nmb = current_nmb
        failed_attempts = 0
    Else
        worsecount = worsecount + 1
        r_sa = Rnd
        If r_sa < Exp((current_nmb - previous_nmb) / temperature) Then 'Accept
            Worksheets("Frontsheet").Range("U35:U35").Offset(i - 1, 0).Value = 0
            Worksheets("Frontsheet").Range("V35:V35").Offset(i - 1, 0).Value = 1
            prevseq = currentseq
            prevseqlength = currentseqlength
            previous_nmb = current_nmb
            failed_attempts = 0
            acceptworsecount = acceptworsecount + 1
        Else
            Worksheets("Frontsheet").Range("U35:U35").Offset(i - 1, 0).Value = 0 'Reject
            Worksheets("Frontsheet").Range("V35:V35").Offset(i - 1, 0).Value = 0
            prevseq = prevseq
            prevseqlength = prevseqlength
            previous_nmb = previous_nmb
            failed_attempts = failed_attempts + 1
        End If
    End If
End If

If temperature = 1 Then
    temperature = 0
End If

#####STOPPING RULE
If i = max_runs Then 'Maximum runs reached = stop
    stopsim = 1
    stopsimreason = 1
End If

If min_temp_stop_active = 1 Then 'Minimum temperature reached

```

```

    If temperature < min_temp_stop Then
        stopsim = 1
        stopsimreason = 2
    End If
End If

If max_failed_attempts_active = 1 Then
    If failed_attempts >= max_failed_attempts Then 'Max consecutive failed attempts reached
        stopsim = 1
        stopsimreason = 3
    End If
End If

If programme = 5 Then
    If worsecount > 0 Then
        p_accept = acceptworsecount / worsecount
    End If

    If worsecount = worseloop Then
        temperature = temperature + temp_algorithm_increment

        If p_accept >= p_accept_target Then
            stopsim = 1
            Worksheets("Frontsheet").Range("X23:X23").Offset(b - 1, 0).Value = temperature
        Else
            acceptworsecount = 0
            worsecount = 0
        End If
    End If
    Worksheets("Frontsheet").Range("AH35:AH35").Offset(i - 1, 0).Value = p_accept
    Worksheets("Frontsheet").Range("AI35:AI35").Offset(i - 1, 0).Value = acceptworsecount
    Worksheets("Frontsheet").Range("AJ35:AJ35").Offset(i - 1, 0).Value = worsecount
End If

i = i + 1
total_i = total_i + 1
ActiveWindow.ScrollRow = i
Loop

Calculate

a = 1

```



```

Do Until a > UBound(bestseq)           'Update the best sequence in the spreadsheet
    Range("B23").Offset(b - 1, a - 1).Value = bestseq(a)
    a = a + 1
Loop
' prevrestartbestseq = bestseq
' prevrestartbestseq_index = b

If b = 1 Then
    overall_best_nmb = best_nmb
    overallbestseq = bestseq
    overallbestseq_index = b
Else
    overall_best_nmb = best_nmb
    overallbestseq = bestseq
    overallbestseq_index = b
End If
Range("V20").Value = overallbestseq_index
' Range("V21").Value = prevrestartbestseq_index

Worksheets("Frontsheet").Range("P23:P23").Offset(b - 1, 0).Value = best_nmb
Worksheets("Frontsheet").Range("Q23:Q23").Offset(b - 1, 0).Value = best_cost
Worksheets("Frontsheet").Range("R23:R23").Offset(b - 1, 0).Value = best_qaly
Worksheets("Frontsheet").Range("S23:S23").Offset(b - 1, 0).Value = best_nmb_index
Worksheets("Frontsheet").Range("T23:T23").Offset(b - 1, 0).Value = i - 1
Worksheets("Frontsheet").Range("U23:U23").Offset(b - 1, 0).Value = stopsimreason

Select Case b
Case 1
    Sheets("R1").Range("B2:AJ10002").Value = Sheets("Frontsheet").Range("B35:AJ10035").Value
Case 2
    Sheets("R2").Range("B2:AJ10002").Value = Sheets("Frontsheet").Range("B35:AJ10035").Value
Case 3
    Sheets("R3").Range("B2:AJ10002").Value = Sheets("Frontsheet").Range("B35:AJ10035").Value
Case 4
    Sheets("R4").Range("B2:AJ10002").Value = Sheets("Frontsheet").Range("B35:AJ10035").Value
Case 5
    Sheets("R5").Range("B2:AJ10002").Value = Sheets("Frontsheet").Range("B35:AJ10035").Value
Case 6
    Sheets("R6").Range("B2:AJ10002").Value = Sheets("Frontsheet").Range("B35:AJ10035").Value
Case 7
    Sheets("R7").Range("B2:AJ10002").Value = Sheets("Frontsheet").Range("B35:AJ10035").Value
Case 8

```

```

    Sheets("R8").Range("B2:AJ10002").Value = Sheets("Frontsheet").Range("B35:AJ10035").Value
Case 9
    Sheets("R9").Range("B2:AJ10002").Value = Sheets("Frontsheet").Range("B35:AJ10035").Value
Case 10
    Sheets("R10").Range("B2:AJ10002").Value = Sheets("Frontsheet").Range("B35:AJ10035").Value
End Select

    b = b + 1           'Next Restart (if relevant)
Loop

If programme = 5 Then
    stopsimreason = 4
End If

If end_run_search_active = 1 Then
    Range("B35:AJ999999").ClearContents
    i = 1

    Do Until i > end_run_search
        Calculate
        Application.StatusBar = "End Run Evaluation " & i & " out of " & end_run_search

        If i = 1 Then
            a = 1
            overallbestseqlength = UBound(overallbestseq)
            Do Until a > UBound(overallbestseq)           'Update the best sequence in the spreadsheet
                Range("B35").Offset(i - 1, a - 1).Value = overallbestseq(a)
                a = a + 1
            Loop
            prevseq = overallbestseq
        Else
            duplicate = 0
            duplicate_index = 0
            eval_start_time = 0
            eval_end_time = 0

            If sequence_generation = 1 Then 'Generate a random sequence
                Call gen_random_seq(i, currentseq, currentseqlength)
            ElseIf sequence_generation = 2 Then 'Pairwise swap
                Call pairswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
            ElseIf sequence_generation = 3 Then 'Random swap + Add/Subtract (DEFAULT)
                Select Case currentseqlength 'addition/remove 1 or 2 tx,

```

```

'dependent on probs & length of sequence
Case 1 To 2
  p_add = 1
  p_add2 = bp_add2
  p_remove = 0
  p_remove2 = 0
  r_seq = Rnd
  r_seq2 = Rnd
  If r_seq > 1 - p_remove Then
    If r_seq2 < p_remove2 / p_remove Then
      Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
    Else
      Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
  ElseIf r_seq < p_add Then
    If r_seq2 < p_add2 / p_add Then
      Call addseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
    Else
      Call addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
  Else
    Call randomswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
  End If
Case 3
  p_add = bp_add
  p_add2 = bp_add2
  p_remove = bp_remove
  p_remove2 = 0
  r_seq = Rnd
  r_seq2 = Rnd
  If r_seq > 1 - p_remove Then
    If r_seq2 < p_remove2 / p_remove Then
      Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
    Else
      Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
  ElseIf r_seq < p_add Then
    If r_seq2 < p_add2 / p_add Then
      Call addseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
    Else
      Call addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
  End If

```

```

Else
  Call randomswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
End If
Case 4 To 10
  p_add = bp_add
  p_add2 = bp_add2
  p_remove = bp_remove
  p_remove2 = bp_remove2
  r_seq = Rnd
  r_seq2 = Rnd
  If r_seq > 1 - p_remove Then
    If r_seq2 < p_remove2 / p_remove Then
      Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
    Else
      Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
  ElseIf r_seq < p_add Then
    If r_seq2 < p_add2 / p_add Then
      Call addseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
    Else
      Call addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
  Else
    Call randomswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
  End If
Case 11 To 12
  p_add = bp_add
  p_add2 = 0
  p_remove = bp_remove
  p_remove2 = bp_remove2
  r_seq = Rnd
  r_seq2 = Rnd
  If r_seq > 1 - p_remove Then
    If r_seq2 < p_remove2 / p_remove Then
      Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
    Else
      Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
  ElseIf r_seq < p_add Then
    If r_seq2 < p_add2 / p_add Then
      Call addseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
    Else

```

```

        Call addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
Else
    Call randomswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
End If
Case 13 To 14
    p_add = 0
    p_add2 = 0
    p_remove = bp_remove
    p_remove2 = bp_remove2
    r_seq = Rnd
    r_seq2 = Rnd
    If r_seq > 1 - p_remove Then
        If r_seq2 < p_remove2 / p_remove Then
            Call removeseq2(i, prevseq, currentseq, currentseqlength, prevseqlength)
        Else
            Call removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
        End If
    Else
        Call randomswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
    End If
End Select
End If

duplicate = 0
duplicate_index = 0
eval_start_time = 0
eval_end_time = 0

If total_i > 1 Then
    If cache_search_active = 1 Then
        eval_start_time = Now()
        Call cache_search(total_i, i, duplicate, duplicate_index, currentseqlength)
        eval_end_time = Now()
    End If
End If

If duplicate = 1 Then
    Worksheets("Frontsheet").Range("P35:P35").Offset(i - 1, 0).Value =
Worksheets("Cache").Range("P2").Offset(duplicate_index - 1, 0).Value

```

```

        Worksheets("Frontsheet").Range("Q35:Q35").Offset(i - 1, 0).Value =
Worksheets("Cache").Range("Q2").Offset(duplicate_index - 1, 0).Value
        Worksheets("Frontsheet").Range("R35:R35").Offset(i - 1, 0).Value =
Worksheets("Cache").Range("R2").Offset(duplicate_index - 1, 0).Value
        Worksheets("Frontsheet").Range("AE35:AE35").Offset(i - 1, 0).Value = "DUPLICATE"
        Worksheets("Frontsheet").Range("AF35:AF35").Offset(i - 1, 0).Value =
SecondsToDateTimeSerial(DateDiff("s", eval_start_time, eval_end_time))
        Worksheets("Frontsheet").Range("AG35:AG35").Offset(i - 1, 0).Value = duplicate_index
    Else
        eval_start_time = Now()
        Range("U20").Value = total_i
        Call runmodel(i, numsim) 'run model to evaluate
        If Worksheets("Frontsheet").Range("AC35:AC35").Offset(i - 1, 0).Value = 1 Then
            nmb = -999999999
            Worksheets("Frontsheet").Range("P35:P35").Offset(i - 1, 0).Value = nmb
        End If

        eval_end_time = Now()
        Worksheets("Frontsheet").Range("AE35:Ae35").Offset(i - 1, 0).Value = "EVALUATE"
        Worksheets("Frontsheet").Range("AF35:AF35").Offset(i - 1, 0).Value =
SecondsToDateTimeSerial(DateDiff("s", eval_start_time, eval_end_time))
    End If

    current_nmb = Worksheets("Frontsheet").Range("P35:P35").Offset(i - 1, 0).Value
    current_cost = Worksheets("Frontsheet").Range("Q35:Q35").Offset(i - 1, 0).Value
    current_qaly = Worksheets("Frontsheet").Range("R35:R35").Offset(i - 1, 0).Value
    Sheets("Cache").Range("B2:R2").Offset(total_i - 1, 0).Value =
Worksheets("Frontsheet").Range("B35:R35").Offset(i - 1, 0).Value
    If i = 1 Then 'update current best results
        best_nmb = current_nmb
        best_cost = current_cost
        best_qaly = current_qaly
        previous_nmb = current_nmb
        bestseq = currentseq
        Worksheets("Frontsheet").Range("T35:T35").Value = best_nmb
        best_nmb_index = i
        Worksheets("Frontsheet").Range("S35:S35").Value = best_nmb_index
    Else
        If current_nmb > best_nmb Then 'improvement. update current best
            best_nmb = current_nmb
            best_cost = current_cost
            best_qaly = current_qaly

```

```

bestseq = currentseq

Worksheets("Frontsheet").Range("T35:T35").Offset(i - 1, 0).Value = best_nmb
best_nmb_index = i
Worksheets("Frontsheet").Range("S35:S35").Offset(i - 1, 0).Value = i
Else 'otherwise retain the current best
Worksheets("Frontsheet").Range("T35:T35").Offset(i - 1, 0).Value = best_nmb
Worksheets("Frontsheet").Range("S35:S35").Offset(i - 1, 0).Value = best_nmb_index
End If
Worksheets("Frontsheet").Range("W35:W35").Offset(i - 1, 0).Value = "LOCAL"
End If

If i = 1 Then
currentseq = bestseq
prevseq = currentseq
previous_nmb = current_nmb
currentseqlength = UBound(currentseq)
prevseqlength = currentseqlength
Else
If current_nmb > previous_nmb Then 'Improvement = Accept
Worksheets("Frontsheet").Range("U35:U35").Offset(i - 1, 0).Value = 1
Worksheets("Frontsheet").Range("V35:V35").Offset(i - 1, 0).Value = 1
prevseq = currentseq
prevseqlength = currentseqlength
previous_nmb = current_nmb
failed_attempts = 0
Else 'Reject
Worksheets("Frontsheet").Range("U35:U35").Offset(i - 1, 0).Value = 0
Worksheets("Frontsheet").Range("V35:V35").Offset(i - 1, 0).Value = 0
prevseq = prevseq
prevseqlength = prevseqlength
previous_nmb = previous_nmb
failed_attempts = failed_attempts + 1
End If
End If

i = i + 1
total_i = total_i + 1
ActiveWindow.ScrollRow = i
Loop

If best_nmb > overall_best_nmb Then

```

```

overall_best_nmb = best_nmb
overallbestseq = bestseq
End If

a = 1
Do Until a > UBound(overallbestseq) 'Update the best sequence in the spreadsheet
Range("B33").Offset(0, a - 1).Value = bestseq(a)
a = a + 1
Loop

Worksheets("Frontsheet").Range("P33:P33").Value = best_nmb
Worksheets("Frontsheet").Range("Q33:Q33").Value = best_cost
Worksheets("Frontsheet").Range("R33:R33").Value = best_qaly
Worksheets("Frontsheet").Range("S33:S33").Value = best_nmb_index

End If 'End of simulation

end_time = Now()
Range("S21").Value = SecondsToDateTimeSerial(DateDiff("s", start_time, end_time))
Set MYSIMUL8 = Nothing

If email_notification = 1 Then

Dim iMsg As Object
Dim iConf As Object
Dim strbody As String
Dim flds As Variant

Set iMsg = CreateObject("CDO.Message")
Set iConf = CreateObject("CDO.Configuration")

iConf.Load -1 ' CDO Source Defaults
Set flds = iConf.Fields
With flds
.Item("http://schemas.microsoft.com/cdo/configuration/smtpusessl") = True
.Item("http://schemas.microsoft.com/cdo/configuration/smtpauthenticate") = 1
.Item("http://schemas.microsoft.com/cdo/configuration/sendusername") = ██████████
.Item("http://schemas.microsoft.com/cdo/configuration/sendpassword") = ██████████
.Item("http://schemas.microsoft.com/cdo/configuration/smtpserver") = "smtp.gmail.com"

.Item("http://schemas.microsoft.com/cdo/configuration/sendusing") = 2
.Item("http://schemas.microsoft.com/cdo/configuration/smtpserverport") = 465

```

```

.Update
End With

strbody = "Hi there" & vbCrLf & vbCrLf & _
"Simulation is complete. MESSAGE FROM " & pc_name & vbCrLf & _
"Number of restarts: " & b - 1 & vbCrLf & _
"Number of evaluations: " & i & vbCrLf & _
"Best NMB: " & overall_best_nmb

With iMsg
Set .Configuration = iConf
.To = ██████████
.CC = ██████████
.BCC = ""
' Note: The reply address is not working if you use this Gmail example
' It will use your Gmail address automatic. But you can add this line
' to change the reply address .ReplyTo = "Reply@something.nl"
.From = """"TOSHALERT ██████████
.Subject = "Simulation complete"
.TextBody = strbody
.Send
End With

End If

Select Case stopsimreason
Case 1
MsgBox "REASON FOR STOPPING = MAXIMUM RUNS REACHED. " & i - 1 & _
" Evaluations. Model Run Time = " & SecondsToDateTimeSerial(DateDiff("s", start_time, end_time))
Case 2
MsgBox "REASON FOR STOPPING = MINIMUM TEMPERATURE REACHED. " & i - 1 & _
" Evaluations. Model Run Time = " & SecondsToDateTimeSerial(DateDiff("s", start_time, end_time))
Case 3
MsgBox "REASON FOR STOPPING = MAXIMUM NUMBER OF FAILED ATTEMPTS REACHED. " & _
i - 1 & " Evaluations. Model Run Time = " & SecondsToDateTimeSerial(DateDiff("s", start_time,
end_time))
Case 4
MsgBox "REASON FOR STOPPING = INITIAL TEMPERATURE SETTING ALGORITHM COMPLETE."
End Select

End Sub

```

GENERATE RANDOM SEQUENCE

```

Sub gen_random_seq(i, currentseq, currentseqlength)

Dim t As String
Dim c1_flag, c2_flag, c3_flag, c4_flag, first_line_flag, _
line, tx, shift, flag, bestseqnum, min_init_length, sim, _
numsim, loopcount, init_sol As Integer
Dim c1_flag_arr, c2_flag_arr, c3_flag_arr, c4_flag_arr As Variant
Dim tabl As Range
Dim checklist, bestseq
Dim p_short, p_length, random As Double

Sheets("Frontsheet").Activate
Set tabl = ThisWorkbook.Names("tabl").RefersToRange
Set c1_flag_arr = ThisWorkbook.Names("_c1_array").RefersToRange
Set c2_flag_arr = ThisWorkbook.Names("_c2_array").RefersToRange
Set c3_flag_arr = ThisWorkbook.Names("_c3_array").RefersToRange
Set c4_flag_arr = ThisWorkbook.Names("_c4_array").RefersToRange

seqfault = 1
min_init_length = Range("min_init_length").Value

Do Until seqfault = 0
ReDim checklist(1 To 1)
ReDim currentseq(1 To 1)

If i = 1 Then
Range("B35:V35").ClearContents
End If

tx = 0
c4_flag = 0
first_line_flag = 0
c1_flag = 0
c2_flag = 0
c3_flag = 0
c4_flag = 0
flag = 0
shift = 0
random = Rnd

```

```

Do Until c4_flag = 1
    tx = tx + 1
    loopcount = 0

    Do
        Calculate
        t = Application.WorksheetFunction.Index(Range("R4:R17"), _
            Application.WorksheetFunction.Rank(Range("S4"), Range("S4:S17")))

        flag = Application.VLookup(t, tabl, 3 + shift, 0)

        'loop until eligible treatment is selected based on option matrix (flag)
        'which is offset depending on previous class flags (see later)

        loopcount = loopcount + 1

        If IsError(Application.Match(t, checklist, 0)) Then
            Exit Do
        End If
    Loop

    If tx = 1 Then
        first_line_flag = 1
    Else
        first_line_flag = 0
    End If

    'UPDATE FLAGS (if it's an error it means that a match hasn't been found)
    If IsError(Application.Match(t, c4_flag_arr, 0)) Then
        Else
            c4_flag = 1
            shift = 4
        End If

        If IsError(Application.Match(t, c3_flag_arr, 0)) Then
            Else
                c3_flag = 1
                shift = 3
            End If

        If IsError(Application.Match(t, c2_flag_arr, 0)) Then

```

```

        Else
            c2_flag = 1
            shift = 2
        End If

        If IsError(Application.Match(t, c1_flag_arr, 0)) Then
            Else
                c1_flag = 1
                shift = 1
            End If

        If c1_flag = 1 Then
            shift = 1
            If c2_flag = 1 Then
                shift = 2
                If c3_flag = 1 Then
                    shift = 3
                    If c4_flag = 1 Then
                        shift = 4
                    End If
                End If
            End If
        End If

        'UPDATE WORKSHEET
        Range("B35").Offset(i - 1, tx - 1).Value = t

        ReDim Preserve checklist(1 To tx)
        ReDim Preserve currentseq(1 To tx)
        checklist(tx) = t
        If tx = 1 Then
            If t = "PC" Then
                t = t
            End If
        End If

    Loop

    currentseq = checklist
    currentseqlength = tx
    Call eligible(currentseq, currentseqlength, seqfault)

```

```

If i = 1 Then
  If currentseqlength < min_init_length Then
    seqfault = 1
  End If
End If
Loop

End Sub

```

CHECK SEQUENCE ELIGIBILITY

```

Sub eligible(currentseq, currentseqlength, seqfault)
  Dim t, fault As String
  Dim loopexit, loopcount, tx, shift, flag As Integer
  Dim c1_flag, c2_flag, c3_flag, c4_flag, first_line_flag As Integer
  Dim c1_flag_arr, c2_flag_arr, c3_flag_arr, c4_flag_arr As Variant
  Dim line As Integer
  Dim tabl As Range

  Set tabl = ThisWorkbook.Names("tabl").RefersToRange
  Set c1_flag_arr = ThisWorkbook.Names("_c1_array").RefersToRange
  Set c2_flag_arr = ThisWorkbook.Names("_c2_array").RefersToRange
  Set c3_flag_arr = ThisWorkbook.Names("_c3_array").RefersToRange
  Set c4_flag_arr = ThisWorkbook.Names("_c4_array").RefersToRange

  loopcount = 0
  shift = 0
  seqfault = 0

  Do Until loopcount = currentseqlength

    loopcount = loopcount + 1
    t = currentseq(loopcount)

    flag = Application.VLookup(t, tabl, 3 + shift, 0)

    If flag = 1 Then
      fault = "No"
    Else
      fault = "Yes"
      seqfault = 1
    End If
  Loop

```

```

End If

If IsError(Application.Match(t, c4_flag_arr, 0)) Then
Else
  c4_flag = 1
  shift = 4
End If

If IsError(Application.Match(t, c3_flag_arr, 0)) Then
Else
  c3_flag = 1
  shift = 3
End If

If IsError(Application.Match(t, c2_flag_arr, 0)) Then
Else
  c2_flag = 1
  shift = 2
End If

If IsError(Application.Match(t, c1_flag_arr, 0)) Then
Else
  c1_flag = 1
  shift = 1
End If

If c1_flag = 1 Then
  shift = 1
  If c2_flag = 1 Then
    shift = 2
    If c3_flag = 1 Then
      shift = 3
      If c4_flag = 1 Then
        shift = 4
      End If
    End If
  End If
End If

If loopcount > 1 Then
  a = 1
  Do Until a = loopcount

```

```

    If t = currentseq(a) Then
        seqfault = 1
    End If

    a = a + 1
Loop
End If

Loop

End Sub

```

PAIRWISE SWAP

```

Sub pairswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
Dim sequence
Dim selection_index1, selection_index2, throwaway, a As Integer
Dim selection_tx1, selection_tx2 As String
Dim r_num As Double
seqfault = 1
throwaway = 0
Do Until seqfault = 0
    r_num = Rnd
    sequence = prevseq

    If prevseqlength > 2 Then

        selection_index1 = Int(((prevseqlength - 2) - 1 + 1) * r_num + 1)

        selection_tx1 = sequence(selection_index1)
        selection_index2 = selection_index1 + 1
        selection_tx2 = sequence(selection_index2)

        Range("Y35").Offset(i - 1, 0).Value = selection_index1
        Range("Z35").Offset(i - 1, 0).Value = selection_tx1
        Range("AA35").Offset(i - 1, 0).Value = selection_index2
        Range("AB35").Offset(i - 1, 0).Value = selection_tx2

        If currentseqlength > 3 Then 'swap elements
            temp = sequence(selection_index1)
            sequence(selection_index1) = sequence(selection_index2)

```

```

        sequence(selection_index2) = temp
    End If

End If

a = 1
Do Until a > prevseqlength
    Range("B35").Offset(i - 1, a - 1).Value = sequence(a)
    a = a + 1
Loop

currentseq = sequence
currentseqlength = prevseqlength
Call eligible(currentseq, currentseqlength, seqfault)
Range("AC35").Offset(i - 1, 0).Value = seqfault
Range("AD35").Offset(i - 1, 0).Value = throwaway

throwaway = throwaway + 1
If throwaway > 200 Then
    throwaway = throwaway
    currentseq = prevseq
    seqfault = 0
End If

Loop
Range("X35").Offset(i - 1, 0).Value = "PAIRSWAP"

End Sub

```

RANDOM SWAP

```

Sub randomswap(i, prevseq, currentseq, currentseqlength, prevseqlength)
Dim sequence
Dim selection_index1, selection_index2, loopcount, a, throwaway As Integer
Dim selection_tx1, selection_tx2 As String
Dim r_num1, r_num2 As Double
seqfault = 1
throwaway = 0

Do Until seqfault = 0
    sequence = prevseq

```



```

selection_index1 = 0
selection_index2 = 0
If prevseqlength > 2 Then
    loopcount = 1

    Do Until selection_index1 <> selection_index2

        r_num1 = Rnd
        r_num2 = Rnd

        selection_index1 = Int(r_num1 * (prevseqlength - 1)) + 1
        selection_index2 = Int(r_num2 * (prevseqlength - 1)) + 1
        selection_tx1 = sequence(selection_index1)
        selection_tx2 = sequence(selection_index2)

        Range("Y35").Offset(i - 1, 0).Value = selection_index1
        Range("Z35").Offset(i - 1, 0).Value = selection_tx1
        Range("AA35").Offset(i - 1, 0).Value = selection_index2
        Range("AB35").Offset(i - 1, 0).Value = selection_tx2

        If loopcount = 1000 Then
            Exit Do
        End If
        loopcount = loopcount + 1

    Loop

    If prevseqlength >= 3 Then 'swap elements
        temp = sequence(selection_index1)
        sequence(selection_index1) = sequence(selection_index2)
        sequence(selection_index2) = temp
    End If

End If

a = 1
Do Until a > prevseqlength
    Range("B35").Offset(i - 1, a - 1).Value = sequence(a)
    a = a + 1
Loop

currentseq = sequence

```

```

currentseqlength = prevseqlength
Call eligible(currentseq, currentseqlength, seqfault)
Range("AC35").Offset(i - 1, 0).Value = seqfault
Range("AD35").Offset(i - 1, 0).Value = throwaway

throwaway = throwaway + 1
If throwaway > 200 Then
    throwaway = throwaway
    currentseq = prevseq
    seqfault = 0
End If

Loop
'Range("B35").Offset(i - 1, selection_index1 - 1).Font.Bold = True
'Range("B35").Offset(i - 1, selection_index2 - 1).Font.Bold = True
'Range("B35").Offset(i - 2, selection_index1 - 1).Font.Bold = True
'Range("B35").Offset(i - 2, selection_index2 - 1).Font.Bold = True
Range("X35").Offset(i - 1, 0).Value = "SWAP"

End Sub

```

REMOVAL

```

Sub removeseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
Dim sequence
Dim selection_index, a, throwaway As Integer
Dim selection_tx As String
Dim r_num As Double
ReDim sequence(1 To prevseqlength) As String
seqfault = 1
throwaway = 0
Do Until seqfault = 0
    sequence = prevseq
    r_num = Rnd
    selection_index = Int(r_num * (prevseqlength - 1)) + 1
    selection_tx = sequence(selection_index)
    Range("Y35").Offset(i - 1, 0).Value = selection_index
    Range("Z35").Offset(i - 1, 0).Value = selection_tx

    If selection_index > 1 Then
        ReDim arr1(1 To selection_index - 1) As String
    End If

```

```

ReDim arr2(1 To prevseqlength - selection_index) As String
ReDim newseq(1 To prevseqlength - 1) As String

For n = 1 To UBound(arr1)
    newseq(n) = sequence(n)
Next n

For n = 1 To UBound(arr2)
    newseq(n + selection_index - 1) = sequence(n + selection_index)
Next n

Else
ReDim newseq(1 To prevseqlength - 1) As String
For n = 1 To UBound(newseq)
    newseq(n) = sequence(n + 1)
Next n
End If

a = 1
Do Until a > prevseqlength - 1
    Range("B35").Offset(i - 1, a - 1).Value = newseq(a)
    a = a + 1
Loop
currentseq = newseq
currentseqlength = prevseqlength - 1

Call eligible(currentseq, currentseqlength, seqfault)
Range("AC35").Offset(i - 1, 0).Value = seqfault
Range("AD35").Offset(i - 1, 0).Value = throwaway
throwaway = throwaway + 1
If throwaway > 200 Then
    throwaway = throwaway
    currentseq = prevseq
    seqfault = 0
End If
Loop

Range("X35").Offset(i - 1, 0).Value = "REMOVAL"
'Range("B35").Offset(i - 2, selection_index - 1).Interior.ColorIndex = 3

End Sub

```

ADDITION

```

Sub addseq(i, prevseq, currentseq, currentseqlength, prevseqlength)
Dim sequence
Dim selection_index, insert_selection_index, a, throwaway As Integer
Dim selection_tx, insert_selection_tx As String
Dim r_num, r_num2 As Double
ReDim sequence(1 To prevseqlength) As String
Dim txlist(1 To 14) As String
txlist(1) = "ABT"
txlist(2) = "ABTS"
txlist(3) = "ADA"
txlist(4) = "CTZ"
txlist(5) = "ETN"
txlist(6) = "GOL"
txlist(7) = "HCQ"
txlist(8) = "IFX"
txlist(9) = "MTX"
txlist(10) = "RTX"
txlist(11) = "PC"
txlist(12) = "SSZ"
txlist(13) = "TCZ"
txlist(14) = "TICORA"
seqfault = 1
throwaway = 0

Do Until seqfault = 0
    sequence = prevseq
    r_num = Rnd
    selection_index = Int(r_num * (prevseqlength - 1)) + 1
    selection_tx = sequence(selection_index)
    Range("Y35").Offset(i - 1, 0).Value = selection_index + 1

    If selection_index > 1 Then

        ReDim arr1(1 To selection_index) As String
        ReDim arr2(1 To prevseqlength - selection_index) As String
        ReDim newseq(1 To prevseqlength + 1) As String

        For n = 1 To UBound(arr1)
            newseq(n) = sequence(n)
        Next n

```

```

loopcount = 1

Do
    r_num2 = Rnd
    insert_selection_index = Int(r_num2 * 14) + 1
    insert_selection_tx = txlist(insert_selection_index)
    Range("Z35").Offset(i - 1, 0).Value = insert_selection_tx
    If IsInArray(insert_selection_tx, sequence) Then

        Else
            Exit Do
        End If
    If loopcount = 1000 Then
        Exit Do
    End If
    loopcount = loopcount + 1
Loop
newseq(selection_index + 1) = insert_selection_tx

For n = 1 To UBound(arr2)
    newseq(n + selection_index + 1) = sequence(n + selection_index)
Next n
Else
    ReDim newseq(1 To prevseqlength + 1) As String
    loopcount = 1
    Do
        r_num2 = Rnd
        insert_selection_index = Int(r_num2 * 14) + 1
        insert_selection_tx = txlist(insert_selection_index)
        Range("Z35").Offset(i - 1, 0).Value = insert_selection_tx
        If IsInArray(insert_selection_tx, sequence) Then

            Else
                Exit Do
            End If
        If loopcount = 1000 Then
            Exit Do
        End If
        loopcount = loopcount + 1
    Loop
    newseq(selection_index) = insert_selection_tx

```

```

For n = 1 To UBound(newseq) - 1
    newseq(n + 1) = sequence(n)
Next n
End If

a = 1
Do Until a > prevseqlength + 1
    Range("B35").Offset(i - 1, a - 1).Value = newseq(a)
    a = a + 1
Loop

currentseqlength = prevseqlength + 1
Call eligible(newseq, currentseqlength, seqfault)
Range("AC35").Offset(i - 1, 0).Value = seqfault
currentseqlength = prevseqlength - 1
Range("AD35").Offset(i - 1, 0).Value = throwaway

throwaway = throwaway + 1
If throwaway > 200 Then
    throwaway = throwaway
    currentseq = prevseq
    seqfault = 0
End If
Loop

'If selection_index > 1 Then
' Range("B35").Offset(i - 1, selection_index).Interior.ColorIndex = 4
'Else
' Range("B35").Offset(i - 1, selection_index - 1).Interior.ColorIndex = 4
'End If

currentseq = newseq
If throwaway > 200 Then
    currentseqlength = prevseqlength
    currentseq = prevseq
Else
    currentseqlength = prevseqlength + 1
End If

Range("X35").Offset(i - 1, 0).Value = "ADDITION"

End Sub

```

