

Issues, Challenges and Opportunities for Economic Evaluation in Rare Diseases

Tobias Sydendal Grand

BSc, MSc (Pharmaceutical Sciences), MSc (Health Economics & Health Policy)

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

The University of Sheffield

Faculty of Medicine, Dentistry and Health

School of Health and Related Research

Submitted March 2026

Acknowledgements

I am very grateful to many people for the opportunity to do this PhD. A special thanks to my wife, Simone, for her patience and unwavering support over the years. The last three years have been some of the worst and best years of our lives, and certainly the most eventful: selling an apartment, buying a house that is causing much hardship, birth of our second daughter. Sadly, we also experienced the tragic loss of my wife's father, Eleonora and Merle's grandpa, and my father-in-law. Kim Ziegler was prominent Danish footballer in the 1980's, who made his appearance in the Danish national team and was crowned Danish champion with Odense football club in 1989. On the pitch, he was a fierce and relentless fighter known for his strength and determination. His was renowned for bold tackles that inspired a generation of footballers to come. I would like to dedicate this PhD to Kim Ziegler. In recognition of him and to honour his legacy, I like to believe that I completed this PhD in his indomitable fighting spirit.

I would like to thank Agathe Le Lay for agreeing to this PhD project many years ago. The topic changed multiple times, but I am forever grateful for this opportunity, and for the chance to contribute with research in rare diseases. The unmet need for patients is enormous, which emphasises this research privilege. In addition, I would like thank Stephane Regnier (company supervisor) for always providing his expertise and honest opinions. Daniel Oudin Åström (company supervisor) deserves special recognition. I write directly to you now - you were always there – always supporting & cheering – always helping – and always paving the way for me. I cannot express in writing how important you have been for my sanity and for my ability to complete this project.

My supervisor team at University of Sheffield was second to none. Collectively, your expertise is unrivalled, and any PhD student who gets to study under your supervision should count themselves beyond privileged. Your ability to guide without enforcement was unsurpassed, and I barely noticed your interventions and guidance along the way. The openness to my inputs and ideas provided me with a feeling that my research was effortlessly guided by an invisible hand. Over time, I have come to realise that it only happened due to your extraordinary leadership. Kate, Praveen, and Sarah, I am forever grateful for your supervision.

Abstract

Background and objectives

Orphan drugs have surged in recent decades, but economic evaluation in rare diseases remains challenging. The main research aim set out to explore issues, challenges and opportunities for economic evaluation in rare disease.

Methods

A comprehensive umbrella review was undertaken to identify issues and challenges for economic evaluation of orphan drugs, though opportunities emerged as well. The findings were further evaluated in an early health technology assessment of multiple system atrophy (rare disease), which was disseminated as a user-friendly Shiny app. In addition, cost-effectiveness modelling using Markov and partitioned survival models was conducted to further explore opportunities and corroborate challenges.

Results

The most obvious finding to emerge from the umbrella review was the scarcity of health-economic parameters. It was, moreover, a challenge for cost-effectiveness modelling in multiple system atrophy, e.g., initially for disability scales and later for clinical progression milestones used to inform health states, where health-economic parameters became increasingly scarce. Not surprisingly, a limited quantity of economic evaluations was identified in the umbrella review. Similarly, the evidence synthesis for multiple system atrophy found no economic evaluations despite a wealth of economic evaluations for disease analogues.

Discussion

Missing health-economic parameters make model development challenging. However, the very first economic evaluation was developed in multiple system atrophy. Early modelling in an iterative modelling framework should guide data collection efforts, e.g., using comprehensive and flexible models. This is important, because timely preparation and data generation is key to address missing health-economic parameters, which cause much difficulty when constructing economic evaluations.

Conclusions

To conclude, demonstrating the value of orphan drugs in rare diseases is notoriously difficult, because of pertinent challenges such as dealing with data scarcities for health-economic parameters. Solutions are not plentiful and rarely forthcoming; however, opportunities were identified, but they need further development and refinement by future researchers.

Table of contents

Acknowledgements 1
Abstract 2
Table of contents 3
List of Figures 9
List of Tables 9
Code availability 12
Abbreviations 13
Chapter 1 Introduction 16
1.1 An introduction to nomenclature and foundational concepts 16
1.1.1 Orphan drugs and rare diseases 16
1.1.2 Heterogeneity of diseases 16
1.1.3 Utilitarianism 17
1.1.4 Health-economic evaluation 17
1.1.5 Cost-effectiveness, cost-utility and cost-benefit analysis 18
1.1.6 Decision-analytic modelling 18
1.1.7 Value of information analysis 18
1.1.8 Budget impact analyses 19
1.1.9 Advanced therapeutic medicinal products 20
1.1.10 Risk-sharing agreements 20
1.2 Rare disease policy: the past, the present and the future 21
1.2.1 The past: orphan drug act of 1983 21
1.2.2 The present: policy landscapes 22
1.2.3 The future: continued scepticism 25
1.3 Economic evaluation of rare diseases 26
1.3.1 The surge of orphan drugs 26
1.3.2 Health technology assessment 26
1.3.3 Challenges for economic evaluation 27
1.3.4 Ethical considerations 27
1.3.5 Health technology assessment in the United Kingdom 29
1.4 Aims of the thesis 29
1.5 Outline of the thesis 31

Chapter 2 Issues, challenges and opportunities for economic evaluations of orphan drugs in rare diseases: an umbrella review 35
2.1 Introduction 35
2.1.1 Rationale 35
2.1.2 Objective 37
2.2 Methods 37
2.2.1 Scoping 37
2.2.2 Eligibility Criteria 38
2.2.3 Search 39
2.2.4 Study selection 39
2.2.5 Data collection process 39
2.2.5.1 Summary of characteristics 40
2.2.5.2 Critical appraisal 40
2.2.5.3 Data extraction 40
2.2.6 Reporting 40
2.3 Results 40
2.3.1 Scoping searches 41
2.3.2 Search and study selection 41
2.3.3 Study characteristics 41
2.3.4 Critical appraisal 44
2.3.5 Data extraction 44
2.3.6 Issues with health-economic parameters 47
2.3.6.1 Natural history of disease 47
2.3.6.2 Clinical effectiveness 47
2.3.6.3 Costs 48
2.3.6.4 Quality of life 48
2.3.7 Issues with health-economic evaluations 49
2.3.7.1 Cost effectiveness 49
2.3.7.2 Budget impact 51
2.3.8 Issues with estimating value and reimbursement 51
2.3.8.1 Value frameworks and thresholds 51
2.3.8.2 Multiple criteria decision analysis 52
2.4 Discussion 52
2.4.1 Issues with health-economic parameters 52
2.4.2 Issues with health-economic evaluations 54
2.4.3 Issues with estimating value and reimbursement 55

2.4.4 Recommendations for further research 55
2.4.4.1 The potential of comprehensive and flexible cost-effectiveness models 55
2.4.4.2 Effect of publication bias on the ability to meet cost-effectiveness thresholds 56
2.4.4.3 The importance of estimating prevalence 57
2.4.5 Limitations 58
2.5 Conclusion 58
Chapter 3 Cost-effectiveness modelling of multiple system atrophy for early health technology assessment 60
3.1 Introduction 60
3.1.1 Rationale 60
3.1.2 Objectives 61
3.1.3 Multiple system atrophy 61
3.2 Methods 63
3.2.1 Clinical advice 63
3.2.2 Selection of disease analogues 63
3.2.3 Literature searches 64
3.2.4 Model development 64
3.2.4.1 Target population 65
3.2.4.2 Comparators 65
3.2.4.3 Model perspective 65
3.2.4.4 Model structure 66
3.2.5 Estimated health-economic parameters 67
3.2.5.1 Estimating transition probabilities 67
3.2.5.2 Estimating utilities 71
3.2.6 Other parameters 71
3.2.6.1 Costs 71
3.2.6.2 Clinical events and disutilities 73
3.2.7 Scenario analyses 75
3.2.8 Probabilistic sensitivity analysis 76
3.2.9 Shiny app 76
3.3 Results 79
3.3.1 Literature searches 79
3.3.1.1 Cost-effectiveness studies for multiple system atrophy and disease analogues 79
3.3.1.2 Quality of life studies for multiple system atrophy 80
3.3.1.3 Natural history of disease studies for multiple system atrophy 81
3.3.1.4 Cost-of-illness studies for multiple system atrophy 81

3.3.2 Cost-effectiveness modelling 81
3.3.2.1 Estimated transition probabilities 81
3.3.2.2 Estimated utilities 83
3.3.2.3 Deterministic cost-effectiveness results 84
3.3.2.4 Scenario analyses 84
3.3.2.5 Probabilistic cost-effectiveness results 85
3.3.3 Shiny app 86
3.4 Discussion 86
3.4.1 How can researchers better capture the health economic value of multiple system atrophy? 86
3.4.2 Can data gaps be minimised, or cost-effectiveness modelling facilitated by exploration of rare disease analogues? 87
3.4.3 Was early health technology assessment feasible for multiple system atrophy with currently available evidence? 90
3.4.3.1 Model structure 90
3.4.3.2 Early cost-effectiveness modelling 91
3.4.3.3 Future work for early health technology assessment 93
3.4.4 Is it worthwhile developing user-friendly interfaces for early health technology assessment in the context of rare diseases? 93
3.4.5 Limitations 94
3.5 Conclusions 95
Chapter 4 Conceptual modelling of functional milestones in multiple system atrophy: comparing Markov and partitioned survival models using digitised time-to-event data 96
4.1 Introduction 96
4.1.1 Rationale 96
4.1.2 Objectives 97
4.2 Methods 97
4.2.1 Study population and perspective 98
4.2.2 Model structure and development 98
4.2.3 Estimated health-economic parameters 99
4.2.3.1 Digitising Kaplan-Meier curves for functional milestones 99
4.2.3.2 Efficacy modification for the hypothetical drug 100
4.2.3.3 Reconstructed individual-patient level data 100
4.2.3.4 The assumption of sequential progression 100
4.2.4 Other health economic parameters 101
4.2.4.1 Costs 101
4.2.4.2 Utilities and disutilities 101

4.2.5 Comparison of Markov partitioned survival model approaches 103
4.2.5.1 Markov model 103
4.2.5.2 Partitioned survival model 104
4.2.6 Reporting of analyses 104
4.3 Results 105
4.3.1 Digitised Kaplan-Meier Curves and fitting of distributions 105
4.3.2 Cost-effectiveness results 106
4.3.3 State occupancy 106
4.3.4 Cost-effectiveness plan and acceptability curves 107
4.3.5 Expected value of perfect information 109
4.3.6 Quality-adjusted life year shortfall calculator (sensitivity analysis) 110
4.4 Discussion 111
4.4.1 What are the implications of using alternative model structures and strategies? 111
4.4.2 How do model structures and data sources influence cost-effectiveness? 113
4.4.2.1 Markov and partitioned survival models 113
4.4.2.2 What are the notable differences in comparison to Chapter 3 113
4.4.3 The importance of quality-adjusted life year shortfall multipliers 114
4.4.4 Is it worth conducting expected value of perfect information analyses for rare diseases? 114
4.4.5 Limitations 116
4.5 Conclusions 117
Chapter 5 Discussion 118
5.1 Summary of the thesis 118
5.1.1 Aim 118
5.1.2 Chapters and research questions 118
5.1.3 Main findings 122
5.1.4 Integrated summary 123
5.2 Novel contributions 124
5.2.1 Cost-effectiveness modelling in rare diseases is challenged by missing data 124
5.2.2 Missing health-economic evaluations in rare diseases 125
5.2.3 Disease-specific contributions: novelty through exploration of multiple system atrophy 127
5.2.3.1 The very first cost-effectiveness analysis in multiple system atrophy 127
5.2.3.2 A widespread issue in multiple system atrophy was missing cost data 127
5.2.3.3 Considering data and structural uncertainty with different model approaches (Markov and partitioned survival models) 128
5.3 Impact of this work 130

5.4 Recommendations for economic evaluations in rare diseases 130
5.4.1 Recommendations for nomenclature 130
5.4.1.1 The definition of “orphan drug” and “rare disease” 130
5.4.2 Recommendations for dealing with data scarcities in rare diseases 131
5.4.2.1 Is it an opportunity to use disease analogues to facilitate model development? 131
5.4.2.2 Recommendations for early planning of evidence generation 132
5.4.2.3 Is value of information analysis an overlooked opportunity in rare diseases? 133
5.4.3 Recommendations for economic evaluation of multiple system atrophy 134
5.5 Limitations 135
5.5.1 Drawbacks of the work presented 135
5.5.2 What is missing from this PhD 136
5.5.3 What is uncertain about this work 137
5.6 Future work 137
Chapter 6 Conclusions 139
References 140
Appendices 164
I Appendix (Umbrella Review) 164
I.I Appendix: Tables 164
I.II Appendix: Figures 220
II Appendix (Cost-effectiveness modelling early health technology assessment) 222
II.I Appendix: Tables 222
II.II Appendix: Figure 250
III Appendix: (Conceptual modelling of functional milestones) 251
III.I Modelling report 251
IV Appendix: Publications 282
IV.I Issues, Challenges and Opportunities for Economic Evaluation of Orphan Drugs in Rare Diseases: An Umbrella Review 282
IV.II Cost-Effectiveness Modelling of Multiple System Atrophy for Early Health Technology Assessment 296
IV.III Conceptual modelling of functional milestones in multiple system atrophy: comparing Markov and partitioned survival models using digitised time-to-event data 315
V Appendix (Use of GenAI declaration) 338

List of Figures

Figure 1: Design of thesis by research aim, chapters, research questions, and integrated results.....32
Figure 2: LinkedIn post made by Chris Carswell, editor-in-chief at PharmacoEconomics.....35
Figure 3: Flow diagram. 43
Figure 4: The distribution of JBI (Joanna Briggs Institute) checklist scores across studies. 44
Figure 5: Decision tree (mind map) illustrating extraction themes, sub-topics and issues.....46
Figure 6: Model structure for multiple system atrophy.....66
Figure 7: Output from msm package for prevalence by time in each state. State 1 (Ci), State 2 (Md), State 3 (Vd), State 4 (Td), and state 5 (Death) 83
Figure 8: State occupation by model cycle.....84
Figure 9: Cost-effectiveness plane for PSA 86
Figure 10: Shiny app user interface.....86
Figure 11: Model structure with functional milestones in multiple system atrophy 99
Figure 12: The digitised Kaplan-Meier Curves using IPDfromKM.....105
Figure 13: Comparison of health state occupancy 107
Figure 14: Cost-effectiveness scatterplot 108
Figure 15: The cost-effectiveness acceptability curves.....109
Figure 16: EVPI at the individual level 110
Figure 17: Venn diagram that illustrates overlap between topics across research projects (chapters) 120
Figure 18: Force-directed graph of references created through https://www.connectedpapers.com.....220
Figure 19: Screenshot of the JBI critical appraisal checklist for systematic reviews and research syntheses. Included with permission from the authors. 221
Figure 20: JBI checklist scores per question. The dotted red line represents the number of studies included. 222

List of Tables

Table 1: Orphan drug designation criteria and potential incentives for the European Union, US, and Japan.22
Table 2: Inclusion and exclusion criteria 38
Table 3: Naming conventions for transition probabilities 67
Table 4: Lowest and highest transition probabilities over at six-monthly intervals.....68
Table 5: Transition matrix for BSC 69
Table 6: Transition matrix for hypothetical intervention plus BSC 70
Table 7: Six-monthly cost data estimated from figure 1 in Winter et al. 2011.....72

Table 8: Rates for conversion and inflation adjustments 72
Table 9: Estimated cost data for use in analysis (monthly) 72
Table 10: Overview of clinical events in multiple system atrophy. 73
Table 11: Scenario analyses for exploration 75
Table 12: Specifications for Shiny app (R project) 77
Table 13: Monthly transition probabilities from multistate modelling 81
Table 14: EQ-5D-3L (utilities) for UMSARS Part IV health states 83
Table 15: Deterministic cost-effectiveness results 84
Table 16: Results from scenario analyses 84
Table 17: Cost-effectiveness results from probabilistic sensitivity analysis 85
Table 18: Cost data for functional milestones 101
Table 19: Mean index scores for functional milestones 101
Table 20: Disutilities and percentages for clinical events 102
Table 21: Mean utilities after adjustment for disutility increments, and distributions 102
Table 22: ICER table with cost-effectiveness results across models and analyses 106
Table 23: Results with QALY weights from the shortfall calculator 110
Table 24: Comparison of UMSARS Part IV items and functional milestones 112
Table 25: Scoping searches in Centre for Reviews and Dissemination (CRD) database conducted on the 10th of January 2023. 164
Table 26: Scoping searches in PROSPERO conducted on 10th of January 2023. 164
Table 27: Scoping searches in MEDLINE (1946 to January 9, 2023) by OVID conducted on the 10th of January 2023. 164
Table 28: MEDLINE (1946 to January 9, 2023) searches on Ovid platform performed on the 10th of January 2023. 165
Table 29: EMBASE (1974 to Week 1, 2023) searches performed on the 11th of January 2023. 168
Table 30: Cochrane Database of Systematic Reviews (Issue 1 of 12, January 23) searches performed on the 11th of January 2023. 171
Table 31: Papers included for full screening. 172
Table 32: Summary of study characteristics. 184
Table 33: Answers to each quality assessment question for all studies. 205
Table 34: Issues for economic evaluation of rare diseases. 209
Table 35:MEDLINE (1946 to May 06, 2024) systematic searches on Ovid platform for economic models and multiple system atrophy performed on the 7th of May 2024. 222
Table 36: MEDLINE (1946 to May 6, 2024) systematic searches on Ovid platform for health utilities and multiple system atrophy performed on the 7th of May 2024. 224

Table 37: EMBASE (1974 to 2024 Week 19) systematic searches on Ovid platform for economic models and multiple system atrophy performed on the 14th of May 2024. 226
Table 38: EMBASE (1974 to 2024 Week 19) systematic searches on Ovid platform for health utilities and multiple system atrophy performed on the 14th of May 2024. 228
Table 39: MEDLINE (1946 to May 12, 2024) targeted searches on Ovid platform for economic evaluations and rare-disease analogues performed on the 13th of May 2024. 230
Table 40: EMBASE (1974 to 2024 Week 19) targeted searches on Ovid platform for economic evaluations and rare-disease analogues performed on the 14th of May 2024. 231
Table 41: Searches performed in the NHS EED database on the 7th of May 2024 for multiple system atrophy. Disease analogues only for the period 2019 – 2024 (last 5 years). 232
Table 42: Cost-effectiveness analyses identified from literature and hand searches (only disease analogues) 233

Code availability

Chapter 2

Issues, challenges and opportunities for economic evaluations of orphan drugs in rare diseases: an umbrella review

Flow diagram.

- R code: https://github.com/TobiasGrand/PRISMA2020-Flow-Diagram

Decision tree.

- R code: https://github.com/TobiasGrand/-data-extraction-themes.github.io
- Mind map (html): https://tobiasgrand.github.io/-data-extraction-themes.github.io/

Chapter 3

Cost-effectiveness modelling of multiple system atrophy for early Cost-effectiveness modelling of multiple system atrophy for early health technology assessment

Data-analysis report.

- Markdown (html): https://tobiasgrand.github.io/CEA_MSA_DataAnalysis/

Shiny app (cost-effectiveness model).

- R code: https://github.com/TobiasGrand/CEA_MSA_ShinyApp
- Shiny app: https://sheffield-university.shinyapps.io/Markov-Model-MSA/

Chapter 4

Conceptual modelling of functional milestones in multiple system atrophy: comparing Markov and partitioned survival models using digitised time-to-event data

Cost-effectiveness modelling report.

- R code: https://github.com/TobiasGrand/PSM-and-Markov-for-milestones-in-MSA
- Markdown (html): https://tobiasgrand.github.io/Report_PSM_Markov_MSA/

Abbreviations

Acronym	Meaning
AIC	Akaike Information Criterion
ALS	Amyotrophic Lateral Sclerosis
AMSTAR	Assessment of Multiple Systematic Reviews
ATMP	Advanced Therapeutic Medicinal Product
BIC	Bayesian Information Criterion
BSC	Best Supportive Care
CADTH	Canada's Drug and Health Technology Agency
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
Ci	Completely or not completely independent
CINAHL	Cumulative Index to Nursing & Allied Health
COMPASS	Composite Autonomic Symptom Score
CRD	Centre for Reviews and Dissemination
CSV	Comma-Separated Values
DARTH	Decision Analysis in R for Technologies in Health
DSM IV	Diagnostic and Statistical Manual of Mental Disorders – fourth edition
EDSS	Expanded Disability Status Scale
EED	Economic Evaluation Database
EMA	European Medicines Agency
EMBASE	Excerpta Medica Database
EMSA	European Multiple System Atrophy
EMSA-SG	European Multiple System Atrophy – Study Group

ENBS	Expected Net Benefit of Sampling
EQ-5D or EQ-5D-3L	EuroQol-5Dimensions-3Levels
EU	European Union
EVPI	Expected Value of Perfect Information
EVPPI	Expected Value of Partial Perfect Information
EVSI	Expected Value of Sample Information
FDA	Food and Drug Administration
HIV	Human Immunodeficiency Virus
HST	Highly Specialised Technology
HTA	Health Technology Assessment
ICER	Incremental Cost-Effectiveness Ratio
IPD	Individual Patient-level Data
ISPOR	Society for Pharmacoeconomics and Outcomes Research
JBI	Joanna Briggs Institute
MAP	Modelling Analysis Plan
MCDA	Multiple Criteria Decision Analysis
Md	More dependent
MEDLINE	Medical Literature Analysis and Retrieval System Online
MeSH	Medical Subject Headings
MS	Multiple Sclerosis
MSA	Multiple System Atrophy
MSA-C	Multiple System Atrophy – Cerebellar
MSA-P	Multiple System Atrophy – Parkinsonian
NHS	National Health Service

NICE	National Institute for Health and Care Excellence
PICO	Population, Intervention, Comparison and Outcome
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO	Prospective Register of Systematic Review
PSA	Probabilistic Sensitivity Analysis
QALY	Quality-Adjusted Life-Year
RIS	Research Information Systems
SF-36	36-Item Short Form Survey
SIGN	Scottish Intercollegiate Guidelines Network
SPICE	(Setting, Perspective, Interest, Comparison, Evaluation)
Td	Totally dependent
UK	United Kingdom
UMSARS	Unified Multiple System Atrophy Rating Scale
US	United States
Vd	Very dependent

Chapter 1 Introduction

In this thesis, I study the issues, challenges and opportunities for economic evaluation in rare diseases. This chapter introduces nomenclature and foundational concepts in Section 1.1, which is important for understanding the wording and thesis scope. Section 1.2 provides an overview of rare-disease policies throughout time (the past, the present and the future), which is critical to understand issues for economic evaluation in rare diseases. Section 1.3 explains the effect of policies and the resulting issues for economic evaluation. In addition, a brief overview is provided on the ethics associated with resource allocation for economic evaluations in rare diseases. Section 1.4 outlines the aims of the thesis whereas Section 1.5 outlines the thesis structure, e.g., the content included in each of the chapters.

1.1 An introduction to nomenclature and foundational concepts

The varying nomenclature in health economics can be challenging but it may be even more complex for rare diseases. The following paragraphs clarify nomenclature and introduce foundational concepts in this PhD thesis.

1.1.1 Orphan drugs and rare diseases

Institutions such as regulatory or health technology assessment (HTA) bodies may use different or varying prevalence thresholds to define rare diseases. This document uses the term orphan drug as recommended by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) when a drug is indicated for the treatment of a rare disease. The term rare disease is used for diseases with a minimum prevalence threshold of 40 to 50 patients per 100,000 people (Richter *et al.*, 2015).

1.1.2 Heterogeneity of diseases

Rare diseases are heterogeneous because of diverse and overlapping symptoms, which has implications, e.g., for diagnosis where specialised diagnostic tests and genetic screening may be needed. Clinical trials are difficult to conduct because of varying disease manifestations and small sample sizes. For example, heterogeneity in included sub-groups. The clinical trials may be further complicated by ethical concerns, e.g., blinding or placebo arms in diseases where treatments are missing (Chaudhary and Kumar, 2025). The lack of robust clinical trial data in rare diseases was noted by others (Nestler-Parr *et al.*, 2018).

It is critical for me to pre-emptively state that because of the heterogeneity in rare diseases, all challenges and opportunities identified in my research may apply to a varying degree in rare diseases. The relevance of the findings in this PhD exists on a spectrum – the relevance will vary depending on the rare disease in question.

Similarly, the findings may be applicable or non-applicable to prevalent diseases, but the degree to which they do apply was not possible to determine in my research. Please see Section 2.4 for further explanation of these difficulties. Interestingly, research into genomics is a place where prevalent and rare disease have been dubbed "a false dichotomy", because shared traits in genomic pathways or features may be a better way to identify critical drug targets that can potentially work for both rare and prevalent diseases (Chung, Chau and Wong, 2021; Reinholdt *et al.*, 2025). One could hypothesise that if genomics can be similar, so can challenges between rare and prevalent diseases in economic evaluation, but this discussion will likely continue, because it is difficult to generalise across all diseases. However, it seems that we have moved beyond the general perception that economic evaluation of rare diseases can meet traditional cost-effectiveness thresholds without facing challenges attributed to the rarity of the disease. For example, the perception that it works –*" you just do not like the answers*" (Rittenhouse and Nicod, 2024).

1.1.3 Utilitarianism

The ethical framework that is utilitarianism has largely been attributed to Jeremy Bentham and Stuart Mills who were British philosophers (Marseille and Kahn, 2019; Bentham, 1823). It was referred to as the principle of utility where utility was thought of as a property in any object that produces a benefit, e.g., happiness, or prevents the opposite, e.g., unhappiness for the party (community or individual) (Bentham, 1823). The most famous quote is likely "the greatest good for the greatest number" (Marseille and Kahn, 2019).

A key point in relation to this, is that because this ethical framework considers populations over groups, it may leave certain individuals worse of, if resources are used to the greatest benefit of the entire population. In addition, a criticism is that it does not take the dignity of human life into account, e.g., end-of-life decision making (Felzmann, 2017). Generally, in healthcare, resources are allocated under budget constraints in accordance with these ethical principles, with the aim of maximising overall health benefits, e.g., improved overall survival.

1.1.4 Health-economic evaluation

Health-economic evaluation is defined as an analysis of costs and effects for interventions including downstream consequences for a population over a certain time horizon. The purpose of economic evaluation is mainly to support decision-making on reimbursement decisions, e.g., for HTA processes. The output of these analyses is evaluated against cost-effectiveness thresholds which may differ across value frameworks. For example, country specific value frameworks include varying cost-effectiveness thresholds, which may even differ within countries depending on disease pathways, e.g., rare as opposed to prevalent diseases. For this reason, economic evaluation was considered as a critical component for rare diseases (York Health Economics Consortium, 2016b).

1.1.5 Cost-effectiveness, cost-utility and cost-benefit analysis

The methods used in economic evaluation are cost-benefit or cost-effectiveness analyses, which are used to assist in efficient allocation of resources under budget constraints (Marseille and Kahn, 2019). Cost-benefit analyses evaluate costs and benefits in monetary terms and are mostly used in non-healthcare sectors (York Health Economics Consortium, 2025a). Cost-effectiveness analyses evaluate effectiveness relative to the costs for one or more interventions, e.g., costs per unit of outcome. For example, costs per stroke avoided. Cost-effectiveness analyses that specifically focus on cost per quality-adjusted life year (QALY) are also referred to as cost-utility analyses (York Health Economics Consortium, 2025b). Collectively, these analyses are broadly referred to as being health-economic evaluations. Thus, it has been stated that one must read publications of economic evaluations to know which method, e.g., cost utility, that was used as there is often inconsistent labelling (Brent, 2023).

1.1.6 Decision-analytic modelling

Decision-analytic modelling is often part of economic evaluations, e.g., cost-utility analyses. Decision analysis has been widely used in a range of disciplines, e.g., business analysis or engineering (Briggs, Sculpher and Claxton, 2006). The purpose of decision or mathematical modelling in healthcare is to structure economic and clinical outcomes to inform decision-making for resource allocation, as previously mentioned. A model is a mathematical framework, e.g., Markov models or decision trees, that are informed by data sources. For example, cost-of-illness studies or clinical trials (Weinstein *et al.*, 2003; Gray *et al.*, 2011).

A Markov model is a decision-analytic model type, which assumes that subjects are always in one or more discrete states called Markov states, e.g., alive or dead. The subject can then transition from alive to death, which is determined by a transition probability. The model time horizon of the analysis is divided into increments, often called the cycle length. Each health state can be assigned a weight, utility or costs, for each model cycle, which subjects accumulate depending on the time spent in each state (Sonnenberg and Beck, 1993).

1.1.7 Value of information analysis

I have explained how decision-analytic models outline consequences of various decisions, e.g., interventions and their costs and benefits. However, these models can also evaluate the uncertainty associated with decision problems. It is important to assess uncertainty to avoid suboptimal decisions that could be mitigated by further evidence collection. The concept of value of information originates from statistical decision theory, e.g., making optimal decisions under uncertainty (Heath, Kunst and Jackson, 2024).

A method to assess uncertainty is expected value of perfect information (EVPI). The EVPI indicates if it is worthwhile to fund and conduct further research, e.g., the value of eliminating all uncertainty in all model parameters (Briggs *et al.*, 2012; Barton, Briggs and Fenwick, 2008; Kunst *et al.*, 2023a; Heath, Kunst and Jackson, 2024). In situations where research is worthwhile, it is important to acknowledge that additional research is unlikely to provide perfect information but rather reduce uncertainty. However, it is nonetheless important for decision-makers to consider, because they need to have confidence in their decisions, e.g., would the conclusion change if we had more information? Additional evidence will not have value in value of information analysis if additional evidence will not change the decision.

Multiple types of analyses exist within the realm of value of information analysis. For example, EVPI can be calculated at the individual and population level. Moreover, expected value of partial perfect information (EVPPI) can inform decision-makers about where the uncertainty is coming from. For example, from a parameter or from sets of parameters, which can help determine which parameters contribute most to uncertainty (Briggs, Sculpher and Claxton, 2006; Heath, Kunst and Jackson, 2024).

Expected value of sample information (EVSI) may be useful to obtain information on prioritisation and study design by highlighting where most potential value is provided (York Health Economics Consortium, 2025c). However, it is a complex process which requires that the potential data to arise from a study are estimated, e.g., a treatment effect from a clinical trial. The probabilistic distributions (priors) from the original model are subsequently updated using a Bayesian approach where the model is updated using posterior distributions from the potential study under investigation. It means that the value of a given decision is based on the data from the estimated study (Heath, Kunst and Jackson, 2024).

Lastly, I will briefly touch upon expected net benefit of sampling (ENBS), which in combination with EVSI may help inform study designs. It has been defined as a measure of the net total benefit of undertaking a specific research study, accounting for the number of individuals who would be impacted by any change in decision and the cost of undertaking research (Heath, Kunst and Jackson, 2024). Kunst and colleagues have published definitions for EVPI, EVPPI, EVSI and ENBS together with a checklist for reporting of value of information analysis alongside economic evaluations (Kunst *et al.*, 2023b).

1.1.8 Budget impact analyses

Budget-impact analyses are typically used by decision-makers to assess affordability, but usually not explicitly considered as part of an economic evaluation (York Health Economics Consortium, 2016a). However, they were considered relevant in parts of this thesis, e.g., the published umbrella review (Grand *et al.*, 2024b). The main reason for inclusion was that those interested in economic evaluations, might be equally interested in budget-impact challenges for rare diseases. Grouping of cost-effectiveness and

budget-impact analyses under the term economic evaluation has precedence, but it is not commonplace (Wynne *et al.*, 2023).

1.1.9 Advanced therapeutic medicinal products

Advanced therapeutic medicinal products (ATMPs) were excluded from the thesis for the following reasons: Firstly, orphan drugs can classify as ATMPs, if they meet pre-defined criteria for orphan drugs. 15 percent of orphan drug designations in Europe met ATMP criteria during the period 2009 – 2016, which is an increase from 4.5% during 2000 – 2008, but they are clearly not the majority yet (Farkas *et al.*, 2017).

Secondly, it is debated if ATMPs should classify as drugs (Angelis, Naci and Hackshaw, 2020). ATMPs can be somatic-cell-therapy or tissue-engineered products, which involve cells or tissue being substantially manipulated to provide them with an intended function, or cell therapies where the therapeutic effect comes from genetic manipulation, e.g., manipulating a recombinant-nucleic-acid sequence, which may occur inside or outside the human body (European Medicines Agency, 2015).

Thirdly, the attributes of ATMPs, e.g., curative potential, result in certain challenges being more pronounced as compared to orphan drugs that are not classified as ATMPs. Moreover, it has been suggested to consider medical, ethical, and economic aspects separately for curative and non-curative drugs (Zelei *et al.*, 2016). Challenges for health-economic parameters are common in cost-effectiveness analyses of orphan drugs (Pearson *et al.*, 2018; Nestler-Parr *et al.*, 2018). For ATMPs, non-randomised trials and their immature survival data caused by a limited study duration, were highlighted as critical parameters that were subject to extreme sensitivity analysis in cost-effectiveness analysis (Fiorenza *et al.*, 2020). They are typically one-time or short-term treatments, which explain their higher one-off costs. Importantly, therein lies an idiosyncratic problem of affordability for ATMPs, which require payers to risk an upfront payment for uncertain outcomes, e.g., a cure that materialises over a patients' lifetime (Angelis, Naci and Hackshaw, 2020; Coyle *et al.*, 2020). By contrast, most orphan drugs tend to be developed for chronic diseases (Pearson *et al.*, 2018; Schuller, Hollak and Biegstraaten, 2015; Drummond *et al.*, 2007).

1.1.10 Risk-sharing agreements

Decision-makers have implemented alternative ways of financing in response to high uncertainty in economic evaluations, e.g., for orphan drugs where future outcomes are of critical importance but they are associated with high uncertainty (Drummond, 2015; Facey *et al.*, 2021). These arrangements are in place to facilitate risk-sharing between those supplying (manufacturers) and paying (healthcare providers) for health interventions, which is why they have broadly been referred to as risk-sharing, pay-for-performance or managed-entry agreements. Although, nomenclature is not consistent, they can generally be divided into two categories: health-outcome-based or non-outcome-based agreements (Ferrario and Kanavos, 2015; Carlson *et al.*, 2010).

For example, health-outcome-based agreements may include a reward when predefined milestones are successfully met. For example, confirming a prespecified survival rate through further data collection resulting in a monetary reward or a change from conditional to permanent reimbursement. Whereas non-outcome-based agreements may include financial obligations such as discounts, rebates, and price-volume agreements. In the present thesis, the term managed-entry agreement is used to collectively refer to both health-outcome-based and non-outcome-based agreements.

1.2 Rare disease policy: the past, the present and the future

Incentivising polices have played a significant role for the development of orphan drugs. The history of these policies provides important learnings for those trying to understand the issues of today. The coming paragraphs provide examples of rare-disease policies.

1.2.1 The past: orphan drug act of 1983

The first serious discussions on orphan drugs appeared in the decades leading up to the 1980s, which was a period shaped by the introduction of rigorous safety and efficacy testing required for drug approval. The requirements for human clinical trials increased development costs, which made investments in rare diseases less attractive for manufacturers. The fundamental issue was that selling drugs to few people was a poor business case, especially when the alternative for companies was selling to many people (Pulsinelli, 1999; Grossman, 1984; Arno, Bonuck and Davis, 1995).

Eventually, the conclusion became that too many disincentives existed for orphan drug development, and that this had to change. Undoubtedly, a conclusion reached under concerted pressure from physicians and patients. Although, this patient group was heterogeneous and suffered from variety of rare diseases, all together, it consisted of between 10 to 20 million people (Grossman, 1984; Arno, Bonuck and Davis, 1995).

The United States (US) orphan drug act of 1983 is one of the most important and debated policies, which deserves to be regarded as nothing less than a pivotal moment in the history of rare diseases. A manufacturer could apply for orphan drug designation after the introduction of the orphan drug act. It is currently defined as a beneficial status reserved for interventions that show promise (European Medicines Agency, 2018; National Cancer Institute, 2022). In the beginning, orphan drugs received this status if there was no reasonable expectation of recovering research and development costs through in-country sales.

The 1984 amendment adopted a prevalence threshold as part of the definition. A disease was considered rare if it affected less than 200,000 people in the US (The Senate and House of Representatives, 1984). Only 22 designations were granted over the course of the first year since the implementation of the 1983 orphan drug act, but this number almost tripled less than 5 months after the 1984 amendment (Lawton, 1991). The subsequent evaluation of the 1984 amendment showed that not a single sponsor had applied for orphan

designation based on expected research costs and future earnings (Arno, Bonuck and Davis, 1995). To this day, the prevalence criterion is still in use.

1.2.2 The present: policy landscapes

The US orphan drug act is used as the main example, because it is the oldest of the policies incentivising research for rare diseases. Nevertheless, similar policies emerged across the world. For instance, the Orphan Regulation (Regulation (EC) No 141/2000) and Orphan Product Development Support Programme were introduced in the European Union (EU) 1999 and Japan 1993, respectively (The European Parliament, 2000; National Institutes of Biomedical Innovation, 2022).

Looking across the legislation in Table 1, there are apparent similarities when comparing both criteria for orphan designation and incentives across the three regions: United States, Europe, and Japan. These regions were chosen as examples because of their market size, thus, increasing relevance to readers without unnecessary repetition. Similar legislations exist in other markets, which are not included, for example Australia (Australian Government, 2018; Australian Government, 2021). The influence of the orphan drug act is not difficult to recognise. The readers should note that the legislation is accurately reported according to original sources for the purpose of comparison (European Medicines Agency, 2022a; European Medicines Agency, 2022b; The European Parliament, 2000; European Comission, 2021; European Comission, 2022; Legal Information Institute, 2022a; Legal Information Institute, 2022b; Legal Information Institute, 2022c; Legal Information Institute, 2022d; Legal Information Institute, 2022e; Legal Information Institute, 2022f; Legal Information Institute, 2022g; Legal Information Institute, 2022h; Legal Information Institute, 2022i; National Institutes of Biomedical Innovation, 2022).

Table 1: Orphan drug designation criteria and potential incentives for the European Union, US, and Japan.

Institution (Region)	Designation criteria	Incentives
European Medicines Agency (Europe)	The drug is intended for the diagnosis, prevention, or treatment of a life-threatening or chronically debilitating condition. • affecting not more than five in 10,000 persons in the Community when the application is made **OR** without incentives it is unlikely that the marketing of the medicinal product in the Community would generate	**Economic** • Ten years of market exclusivity. This period can be extended by two years if a paediatric investigation plan was agreed upon. • Fee reductions for regulatory activities: protocol assistance, marketing-authorisation application, inspections before authorisation, applications for changes to

	sufficient return to justify the necessary investment, • that there exists no satisfactory method of diagnosis, prevention or treatment of the condition in question that has been authorised in the Community or, if such method exists, that the medicinal product will be of significant benefit to those affected by that condition.	authorisation after approval, and reduced annual fees. • Grants are not available but can be obtained through other sources. For example, through the Horizon Europe programme (2021-2027) funded by the European Commission. • Other incentives depending on company size and its member state location. **Research** • Protocol assistance: advice on tests and trials to demonstrate quality, safety, and efficacy, but also regulatory support to define the content of the authorisation application. • Compulsory access to the centralised procedure.
Food and Drug Administration (United States)	The term "rare disease or condition'' means any disease or condition which • affects less than 200,000 persons in the United States **OR** affects more than 200,000 in the United States and for which there is no reasonable expectation that the cost of developing and making available in the United States a drug for such disease or condition will be recovered from sales in the United States of such drug.	**Economic** • Tax credits: determined as 25 percent of the human clinical testing expenses, defined as the in-house and contract research expenses, for the taxable year. • Grants and contracts: for development of drugs, observational studies, and pre-clinical and human clinical testing. o Grants for rare neurogenerative diseases for research, e.g., to explore natural history of disease.

		• Potential for seven years of exclusivity after approval. **Research** • Opportunity for written recommendations for non-clinical and clinical investigations. • Priority review for paediatric indications. • Open protocol for investigations.
Ministry of Health, Labour and Welfare of Japan* (Japan)	The drugs, medical devices or regenerative medicine should be indicated for the treatment of serious diseases, including difficult-to-treat diseases. In addition, they must be drugs, medical devices or regenerative medicine for which there are high medical needs. • The number of patients who may use the drugs, medical device or regenerative medicine should be less than 50,000 in Japan. • There is no appropriate alternative drug/medical device/regenerative medicine or treatment **OR** high efficacy or safety is expected compared with existing products. • There should be a theoretical rationale for the use of the product for the target disease, and the development plan should be appropriate.	**Economic** • A fixed subsidy period, from the time of orphan drug designation to the end of the fiscal year in which an application for marketing authorization is submitted, but with a maximum period of three years. • Covered expenses: honoraria, travels, equipment, consumables, printing and binding, communication and transportation, rents and leases, meetings, labour services, and subcontracting fee. Maximum ceiling of half the research and development costs. • Tax-deductible expenses 20% x [R&D cost – subsidy amount]. • Orphan developers who receive grants must refund part of the profits accruing from sales of the commercialised orphan products. Sales above 100 JPY million require refunds. No refund required for sales less than 100 JPY million. **Research**

	• Guidance and advice on interpreting orphan drug designation criteria and clinical development.

*Orphan drugs in Japan are collectively managed by Ministry of Health, Labour and Welfare, Pharmaceuticals and Devices Agency, and National Institute of Biomedical Innovation, Health and Nutrition. R&D (Research and Development) and JPY (Japanese Yen).

It is clear from Table 1 that each of the three regions include a measure of prevalence as a prerequisite for orphan drug designation. The orphan-drug status provides various incentives if granted. For example, market exclusivity among others, but not patents. Patents are well-known from other industries as compared to market exclusivity, which calls for explanation. In the context of diseases, patents may have broad claims and can expire at any time, while market exclusivity may delay or prevent competitors from launching within a certain disease indication. It is possible to protect an orphan drug through the use of both patents and market exclusivity (Food and Drug Administration, 2020).

1.2.3 The future: continued scepticism

Rare-disease policies are considered widely successful because of the enormous increase in orphan drugs that they have stimulated. Despite this, there seem to be a constant temptation to change them (Pulsinelli, 1999; Arno, Bonuck and Davis, 1995; Miller and Lanthier, 2018). For example, the 1994 proposal suggested to shorten the period of market-exclusivity to four years but depending on sales with an opportunity to extend to seven years, and to withdraw orphan drug designations if they exceeded prevalence thresholds throughout the sales period. This amendment was never implemented (Congress, 1994).

Most recently, it was proposed that companies should demonstrate that they do not expect to recoup research and development costs within 12 years of launch to maintain market exclusivity (Congress, 2021). As mentioned, not a single sponsor used this option when the orphan drug act was first introduced in 1983 (Arno, Bonuck and Davis, 1995).

Implementing such a proposal could have stark consequences. It was already noted in 1984 by Senator Hatch (Lawton, 1991, p. 330).

"*If drug manufacturers must in every case not only make difficult development cost and marketing projections with sufficient detail and confidence to satisfy federal auditors, but also open their books and disclose backup financial data, many will find participation in the orphan drug program is simply not worth the bother.*"

Others have argued that it would stimulate insufficient research for rare diseases if incentives were provided only for the non-profitable drugs (Lawton, 1991).

1.3 Economic evaluation of rare diseases

Orphan drugs typically face many challenges when appraised through economic evaluation, which means that they de facto become inaccessible to patients if they fail to meet cost-effectiveness thresholds (Hughes-Wilson *et al.*, 2012; Nestler-Parr *et al.*, 2018; Gammie, Lu and Babar, 2015). In terms of appraisal outcomes, there is a major disconnect between regulatory and economic-evaluation policies for rare diseases. Thus, authors have rightfully asked: If we know that no one will be able to use orphan drugs if appraised by economic evaluation, why do we have incentives to develop them (Zelei *et al.*, 2016)?

1.3.1 The surge of orphan drugs

Orphan drugs have seen an unprecedented increase over the last decades according to data from the European Medicines Agency (EMA) and Food and Drug Administration (FDA). For example, in EU where the number of orphan-drug submissions rose from 548 to 1293 across the periods 2000-2005 to 2015-2020. Similarly, a 10-fold increase was observed in the US for drugs which received an orphan designation from the 1980s to 2010s and recently by a factor of almost three, from 1,153 to 2,979 between the 2000s and 2010s (Miller, Fermaglich and Maynard, 2021; European Medicines Agency, 2022c).

Analysts believe that the proliferation of orphan drugs is due to the success of incentivising policies, which is further supported by observation of similar trends in Japan, which adopted the Orphan Product Development Support Programme in 1993 (Giannuzzi *et al.*, 2017; Miller and Lanthier, 2018; Sakushima, Takeda and Aoi, 2021; Murakami and Narukawa, 2016; Wellman-Labadie and Zhou, 2010)

1.3.2 Health technology assessment

In the past, there was a widely held view that drug prices would naturally be higher for orphan drugs, but this was before the proliferation of orphan drugs. There was an expectation for companies to make a profit from even small patient populations in exchange for their research and development investments. It can best be described as an informal and unspoken “gentleman’s agreement” (Ollendorf, Chapman and Pearson, 2018). The conduct was only possible because of the negligible impact it had on budgets but when the collection of orphan drugs grew too large. This informal and unwritten agreement became most difficult for health service providers to honour. Ignoring the affordability issue was no longer possible.

Nowadays, many countries have HTA bodies who conduct economic evaluation of orphan drugs, which are systematic multi-disciplinary evaluations of cost and effectiveness with the purpose of aiding decision-makers when making reimbursement decisions (York Health Economics Consortium, 2016c). HTA bodies can recommend drugs for reimbursement if they are cost-effective and affordable. This is different from regulatory bodies such as EMA who makes recommendations based on a benefit-risk evaluation, if the

benefits exceed the risks for a drug, it could be recommended for approval (European Medicines Agency, 2020).

Health service providers need to reward innovators of medicine because they direct their efforts toward profitable areas. Fail to do so – and they will stop providing the innovations, e.g., orphan drugs (Towse, 2010). By way of explanation, it has been argued that there should be a mutual interest for both innovators and policymakers to succeed in bringing orphan drugs to the market since policymakers are incentivising research for rare diseases. Stimulating research for the purpose of regulatory approval is not sufficient, because economic evaluation of rare diseases is an equally important hurdle to pass.

1.3.3 Challenges for economic evaluation

The nature of rare diseases means that there are clinical challenges such as low prevalence, use of surrogate endpoints or biomarkers as compared to "hard" clinical outcomes (clear and verifiable outcomes such as a stroke), single-arm clinical trials, and patient heterogeneity. Moreover, these clinical challenges influence or translate directly into issues for economic evaluation. As an example, there might be little data on natural history of disease, effectiveness, and cost, which are important health-economic parameters for cost-effectiveness modelling. These challenges are problematic for economic evaluation because they lead to high uncertainty, which negatively affects the cost-effectiveness results (Nestler-Parr *et al.*, 2018; Pearson *et al.*, 2018). This means that HTA bodies recommend few orphan drugs for reimbursement as they do not meet required evidence standards, or cost-effectiveness thresholds.

In response, countries have loosened their evaluation criteria for orphan drugs to consider different cost-effectiveness and budget-impact thresholds together with other elements. For example, acceptance of higher uncertainty or impact of informal care (Ollendorf, Chapman and Pearson, 2018). One of the issues that has emerged, likely as a cause, is differing evaluation outcomes. For instance, Nicod and colleagues found that recommendations were contrasting in three out of five assessments across England, Scotland, Sweden, and France (Nicod, 2017).

1.3.4 Ethical considerations

As mentioned, healthcare systems have budget constraints and use economic evaluation to efficiently allocate finite resources. The theoretical foundation for this is utilitarianism, which was also previously outlined, and it looks to maximise health. Although, the degree to which cost-utility analysis, a sub-type of cost-effectiveness analysis, adheres to this ethical framework has been discussed at length (Dolan, 2001; Coast, 2009).

Ethical debates are common in rare diseases because orphan drugs rarely have favourable outcomes when appraised through economic evaluation that use traditional cost-effectiveness thresholds, e.g., those used

for standard technologies. For example, if orphan drugs are reimbursed, this results in displacement of other treatments, e.g., more cost-effective treatments for more prevalent and equally worthy diseases (Hughes, Tunnage and Yeo, 2005). Others have highlighted the problem of small populations and effect sizes together with high prices for orphan drugs as a cause for budget-allocation inefficiencies (Eichler *et al.*, 2023).

Those who favour the utilitarian framework argue for maximising health or most value for money, but what does that mean in practice? It means that healthcare providers need to consider many trade-offs if they are to maximise health, because their budgets are finite. They cannot finance every single intervention. This trade-off is often referred to as the opportunity costs, e.g., in a healthcare system with fixed budgets, the opportunity costs are defined as the health lost as a result of the displacement of activities to fund the selected intervention (York Health Economics Consortium, 2025d).

In response to these challenges, other ethicists, in favour of rare diseases, have proposed fairness principles which cover multiple principles that promote fair resource allocation. It is self-evident that utilitarianists consider their own ethical perspective equally fair, if not fairer. Many of these fairness principles are unlikely to converge with utilitarianism. For example, the rule of rescue (Marseille and Kahn, 2019).

The rule of rescue, which is a fairness principle, reflects human propensity to rescue individuals while disregarding costs. For example, outside a healthcare setting, rescue helicopters aiding tourists who drift offshore on paddleboards. In a healthcare setting, one treatment of prophylaxis for haemophilia amounts to $300,000 which may provide more value if allocated elsewhere (Largent and Pearson, 2012). The excessive cost of rescue may be difficult to justify in both examples, if not for an ethical principle like the rule of rescue principle.

Others recommend a fair innings principle: treating younger patients because they have a need for more life years as compared to elderly who already lived through theirs. It should be thought of as the potential life years that can be gained for a patient rather than how young patients are (Sinclair, 2022; Largent and Pearson, 2012).

Furthermore, Zelei *et al.* (2016) provide an excellent overview of ethical principles that favour orphan drugs: non-abandonment (rare-disease patients should not be abandoned), rights-based approach (all have the right to high-quality healthcare), and equality of opportunity (all need a fair chance of improving their health). Ollendorf, Chapman and Pearson (2018) argue for a balanced approach between utilitarianism and fairness principles.

1.3.5 Health technology assessment in the United Kingdom

One may argue that the United Kingdom (UK) is a country that seeks a balanced approach between utilitarianism and fairness principles or at least is accommodating fairness principles despite having utilitarianism as its main foundation.

For illustration, the principles of National Institute for Health and Care Excellence (NICE) outlines *"when making decisions, NICE and our committees strive to balance the need to achieve the most overall benefit for the greatest number of people, with the need to ensure fairness and respect for individual choice"* (National Institute for Health and Care Excellence, 2026). Whilst these principles mainly are utilitarian, indicated by *"most overall benefit for the greatest number of people"*, there is also mentioning of fairness principles, "*ensure fairness and respect for individual choice"*, which indicates that there are exemptions to the utilitarian foundation.

Practically, this is reflected in their guidance documents by having separate processes: a single technology appraisal and highly specialised technology evaluation. A single technology appraisal is for the common diseases, which currently uses a cost-effectiveness threshold of £20,000 – £30,000 per QALY, but this threshold will be raised to £25,000 – £35,000 per QALY as of April 2026, which will likely result in an additional recommendation of 3 to 5 medicines or indications per year (National Institute for Health and Care Excellence, 2025a).

Orphan drugs can enter the highly specialised technology evaluation under exceptional circumstances if they fulfil certain routing criteria. For example, the disease is rare (prevalence is less than 1 per 50,000 people), no more than 300 people are eligible for the licensed indication or no more than 500 people across all indications, it is a rare disease that significantly shortens life or severely impairs quality of life, there are no significant treatment options and orphan drug under evaluation that provides significant benefit over existing treatments (National Institute for Health and Care Excellence, 2022c). In the process NICE applies a much higher cost-effectiveness thresholds of £100,000 per QALY, but decision modifiers (QALY weights) could technically produce thresholds of up to £300,000 per unweighted QALY (National Institute for Health and Care Excellence, 2022d) .

1.4 Aims of the thesis

This PhD set out to explore issues, challenges and opportunities for economic evaluation in rare disease. The research questions were designed to contribute with information to help achieve the main aim, and they were:

- In health economic research settings are there any issues and challenges for the economic evaluation of orphan drugs in rare diseases which apply less to other drugs?

Economic evaluation is explained in Sections 1.1.4 - 1.1.6 and challenges for economic evaluation are introduced in Section 1.3.3.

- How can a conceptual model be developed for Multiple System Atrophy (MSA) to support cost-effectiveness analysis in early HTA, and what evidence gaps currently limit the reliability of such modelling?

Section 3.1.3 includes a description of MSA while the evidence shortcomings are highlighted in Sections 3.3.1.1 - 3.3.1.4.

- Can data gaps in early cost-effectiveness modelling be reduced, or model development facilitated, by conducting literature searches for disease analogues?

The concept of diseases analogues is explained in Section 3.2.2, and limitations further discussed in Section 3.4.2.

- Can patient-level data be reliably digitised and reconstructed from existing published sources?

The need for digitisation of patient-level data and the process is discussed in Section 4.1.1 and 4.2.3.1, respectively.

- Can alternative model structures based on functional milestones be developed to compare the advantages and disadvantages of different modelling strategies, such as Markov and partitioned survival models?

The development of Markov and partitioned survival models are outlined in Sections 4.2.5.1 - 4.2.5.2, but the advantages and disadvantages of model types are discussed in Section 4.4.2.1.

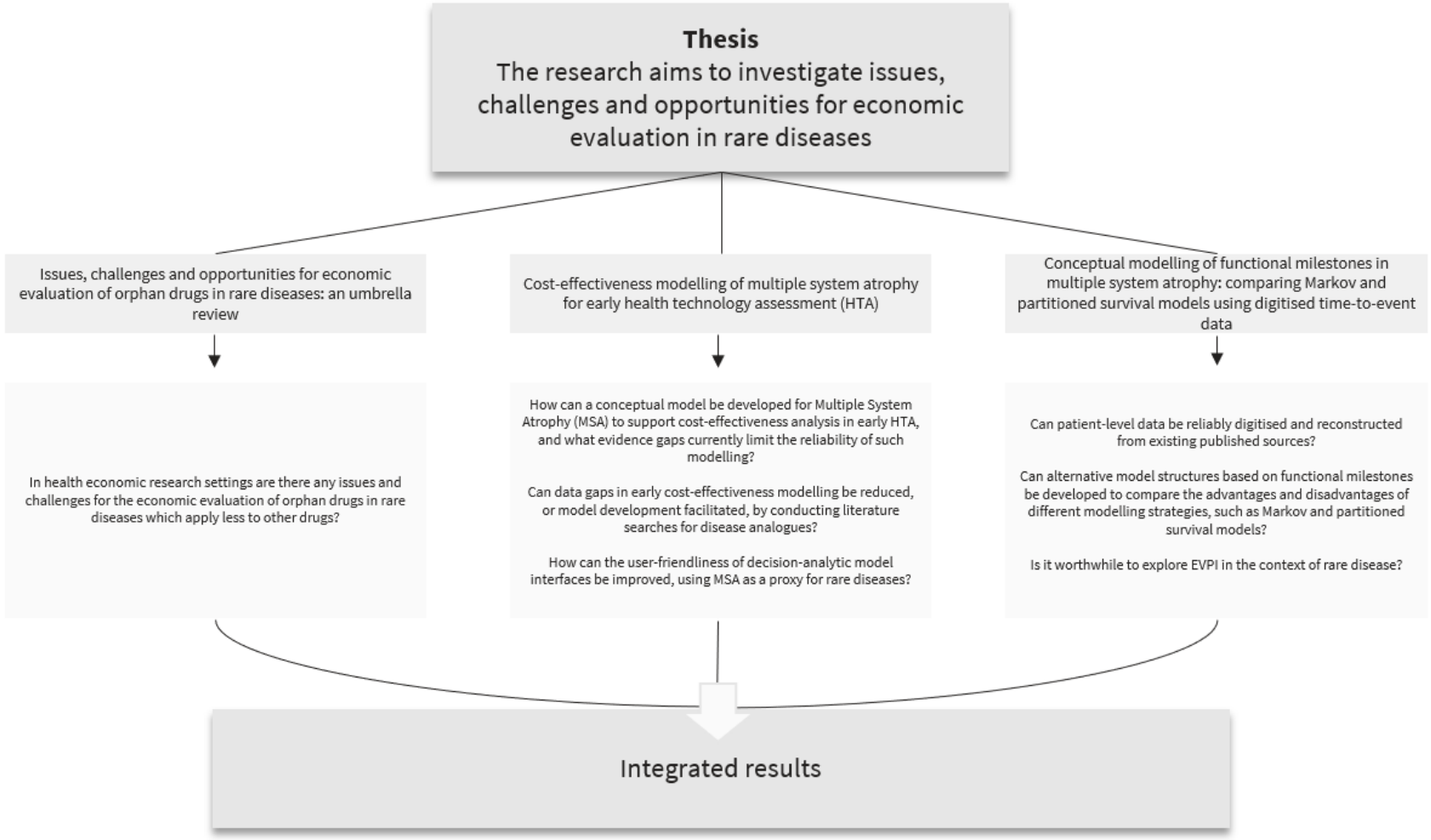
- Is it worthwhile to explore EVPI in the context of rare disease?

The concept of value of information analysis is outlined in Section 1.1.7, but discussed in relation to rare diseases in Section 4.4.4.

1.5 Outline of the thesis

This work takes the form of several research projects or studies, which are reported as independent chapters in the dissertation. Figure 1 provides an overview of the research framework. The main research aim of the thesis is at the top of the framework. Three independent chapters, with their titles in the three boxes, seek to contribute to the overarching research aim by addressing different aspects related to the main research aim. Beneath the chapter title boxes, there are three boxes with the research questions that each of the chapters seek to answer. The information was then collected and discussed collectively, which is represented by the integrated results box that binds everything together.

Figure 1: Design of thesis by research aim, chapters, research questions, and integrated results



The methodological approach taken in this PhD makes use of different methods, e.g., literature reviews and cost-effectiveness modelling, which are applied in multiple studies. The benefit was that each of these research projects were able to clarify and investigate aspects related to economic evaluation of rare disease using different methodology. It provided an exciting opportunity to identify both issues, challenges and opportunities and further advance the current understanding of these, e.g., to identify and corroborate challenges through application of multiple methods in a multi-study framework. In other words – whatever research findings that were made could potentially be examined in more than one project and by various methods, e.g., data scarcities for health-economic parameters through both literature searching and case-studies using decision-analytic modelling, which was an example of a recurrent issue across all research projects. The chapters in the dissertation are described in following paragraphs:

Chapter 2 presents an umbrella review, a systematic review of systematic reviews, which sought to provide an overview of the current state of knowledge and opinion of scholars in the public debate. It was instrumental in identifying challenges for economic evaluation of orphan drugs in rare diseases, and they were systematically categorised by topics. Although, opportunities were not part of the research question as such, this study identified several opportunities that could be explored to alleviate challenges for economic evaluation of rare diseases.

Chapter 3 presents cost-effectiveness modelling of MSA for early HTA as no previous models were identified during literature searching. However, this is mainly a model demonstration of how a cost-effectiveness analysis can be used in MSA, because the intervention is hypothetical. Moreover, data scarcities were highlighted as a major issue in the umbrella review, which is why this was crucial to explore in a case study of MSA. It turned out that this was also an issue for MSA as very little data existed. Lastly, a recommendation for comprehensive and flexible modelling was investigated by coding a user-friendly interface as a Shiny app. Another recommendation was sourcing of surrogate data from disease analogues, e.g., diseases with overlapping characteristics, but with higher prevalence than MSA. It meant that extensive literature searching was conducted to evaluate this opportunity.

Chapter 4 investigated functional milestones to inform health states by different model approaches, e.g., Markov and partitioned survival models, but this required digitising and reconstruction of patient-level data. Again – data scarcities proved to be an issue. The study provided insights into the pros and cons of model strategies and the importance of choosing health states, but it was similarly, a model demonstration based on digitised and surrogate data. Lastly, I conducted and discussed EVPI in the context of rare diseases, which fills an important research gap.

Chapter 5 contains a summary and discussion of the thesis contribution. For example, discussing each of the chapters and resulting research questions. In addition, the contributions to research were outlined, together with the recommendations for future research. Lastly, the limitations were highlighted.

Chapter 6 contains the summary of key findings and recommendations, and lastly the conclusions.

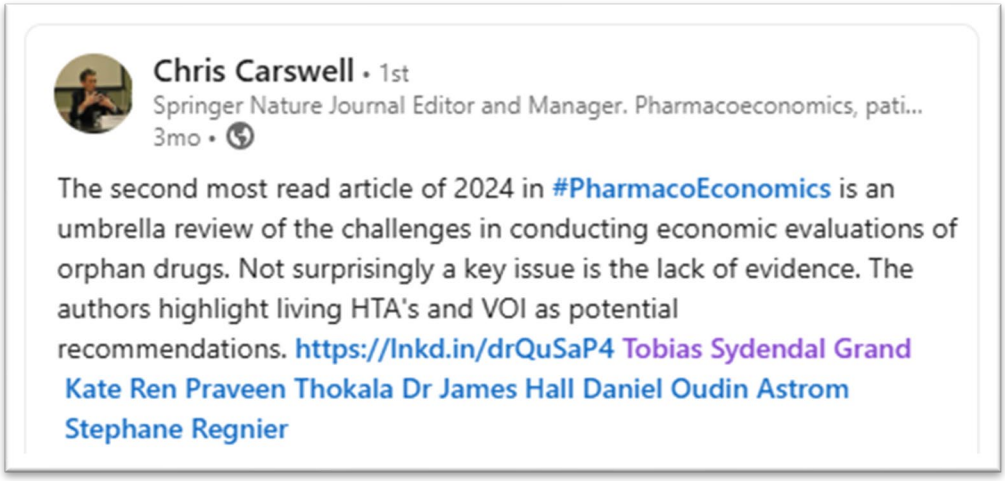
Chapter 2

Issues, challenges and opportunities for economic evaluations of orphan drugs in rare diseases: an umbrella review

This chapter contains my 1st study, which was an umbrella review to identify issues, challenges and opportunities for economic evaluations of orphan drugs in rare diseases. Umbrella reviews aim to synthesise systematic literature reviews, with or without meta-analyses, and the benefit of them is that they can provide meaning of broader topic areas, even if the body of literature is growing rapidly (Aromataris *et al.*, 2015).

The umbrella review was published in PharmacoEconomics (Grand *et al.*, 2024b). A review protocol was made publicly available prior to the commencement of data extraction (Grand *et al.*, 2023). The publication was the 2nd most read paper in PharmacoEconomics 2024, which was announced on LinkedIn by the journal editor (Figure 2).

Figure 2: LinkedIn post made by Chris Carswell, editor-in-chief at PharmacoEconomics



2.1 Introduction

The introduction outlines the rationale for the review and research objective.

2.1.1 Rationale

New methods constantly emerge to address clinical and economic challenges for economic evaluation of rare diseases, while policymakers simultaneously are trying to find new ways of implementing innovative treatments. Scoping searches revealed that the three most influential review papers were all published during 2018. Their influence was judged by their number of citations. A force-directed graph was created

through the website, Connected Papers, and suggested that these papers were indeed some of the most recent and cited papers on the topic of interest (Connected Papers, 2022). See Figure 18 in Appendix .

Two of them focussed on the broader challenges for economic evaluation of rare diseases (Ollendorf, Chapman and Pearson, 2018; Nestler-Parr *et al.*, 2018). The remaining paper was concerned with the economic modelling component of economic evaluations (Pearson *et al.*, 2018). As a result of their year of publication, these three review papers, together with much of the literature identified through scoping searches, did not by themselves sufficiently account for recent developments. For example, living HTA, which is the concept of continuous or even automation of updates to health-economic evaluations (Smith, Schneider and Mohammed, 2022).

It means that regular or scheduled updates, e.g., manual or semi-manual updates, are made to avoid that analyses become outdated if new data on comparators or other data types emerge. This is important, because usually HTA's are one-off analyses. Living HTA has been defined as *"a full HTA that is planned from the outset to be updated at regular intervals or at specific trigger point" (Thokala et al., 2023)*. The process is known from other disciplines, e.g., living systematic reviews (Simmonds *et al.*, 2022). However, few cases of living HTA exist, e.g., an open-source model developed for rheumatoid arthritis by iteratively using feedback and inputs from the public. The iteration process had four phases, which were to be repeated: release initial model, receive feedback, review of feedback and prioritise revisions, revise model and release updated model (Jansen, Incerti and Linthicum, 2019). The model was developed in R software because of the benefits that version control systems provided in managing open-source code (R Core Team, 2021; Jansen, Incerti and Linthicum, 2019)

Another important attribute of living HTA is the ability to increase data protection: firstly, protection of patient data is crucial, and not only due to its value. The management of data may have legal and ethical implications, an issue captured by Ollendorf, Chapman and Pearson (2018), but initially described by Rodriguez-Monguio, Spargo and Seoane-Vazquez (2017). Secondly, data management play a significant role for managed-entry agreements, which as mentioned, are risk-sharing agreements for orphan drugs with high uncertainty where further data are collected, typically for a clinical endpoint (Carlson, Chen and Garrison, 2017; Facey *et al.*, 2021; Drummond, 2015). Nevertheless, with living HTA, the process of data collection might become better controlled and more efficient, e.g., sending scripts to a protected and centralised database that returns results without direct data access. It could also allow for broader focus in managed-entry agreements than simply the collection of pre-defined clinical endpoints, e.g., fully updated cost-effectiveness models across multiple health-economic parameters.

Turning now to recent policy developments in England, the introduction of The Innovative Medicines Fund by the NICE. This fund supports the pipeline of promising medicines for rare conditions and allows for a temporary recommendation while data are collected to alleviate uncertainties (National Institute for Health

and Care Excellence, 2022b). Previously a dedicated fund to finance managed-entry agreements only existed for cancer medicines (National Health Service England, 2016). There are, however, examples of managed-entry agreements for innovative medicines that pre-date The Innovative Medicines Fund, but a dedicated fund will vastly reduce the time it takes for medicines to reach patients (National Health Service England, 2021; National Institute for Health and Care Excellence, 2022a; National Institute for Health and Care Excellence, 2023).

To conclude, issues, challenges and opportunities associated with economic evaluations of orphan drugs rapidly transpire. Consequently, existing studies did not include all methodological breakthroughs, simply because recent events had superseded them. These findings warrant a systematic and rigorous effort to synthesise the literature.

2.1.2 Objective

It was decided to use a modified version of SPICE (Setting, Perspective, Interest, Comparison, Evaluation) framework to scope the research question, which can be more appropriate for social sciences due to a broadening of the evaluation framework in comparison to PICO (Population, Intervention, Comparison and Outcome) (Booth, 2006). I decided to omit the perspective component because all perspectives were relevant.

When applying the framework with its parameters in brackets, e.g., [Setting]. The research question became: **In health economic research settings** [Setting] **are there any issues and challenges** [Evaluation] **for the economic evaluation of orphan drugs in rare diseases** [Interest], **which apply less to other drugs** [Comparison]**?**

2.2 Methods

The following sections describe the methods applied.

2.2.1 Scoping

On the 10th of January 2023, scoping searches were executed in the following databases: Prospective Register of Systematic Review (PROSPERO), the Centre for Reviews and Dissemination (CRD) and Medical Literature Analysis and Retrieval System Online (MEDLINE) by Ovid. The purpose was to establish whether existing reviews were ongoing covering the exact topic of this review and to assess the body of literature. The search applied quotation marks to retrieve exact phrases, asterisks to truncate words, Boolean operators and a search filter for economic studies. The search filter was developed by CRD at the University of York but modified by the Scottish Intercollegiate Guidelines Network (SIGN) (SIGN, 2021a).

The scooping searches are listed in Appendix : Table 25, Table 26 and Table 27.

2.2.2 Eligibility Criteria

The research question informed inclusion and exclusion criteria for the literature search, which are listed in Table 2.

Table 2: Inclusion and exclusion criteria

	Inclusion	Exclusion
Setting	• Health economic research settings	

Interest	• Issues and challenges for economic evaluation in rare diseases. Including, but not limited to:
	  ○ Cost data
	  ○ Effectiveness data
	  ○ Natural history of disease
	  ○ Burden of disease
		• Studies not describing issues and challenges for economic evaluation in rare diseases.

Comparison	• Orphan drugs
	  ○ Papers describing orphan and ultra-orphan drugs
		• Papers exclusively on
		  ○ ATMPs* are excluded, even if they have orphan drug designations. For example, gene and cell therapies
		  ○ Personalised medicines
		  ○ Other drugs that are not orphan or ultra-orphan
		• Interventions not classified as drugs

Evaluation	• Systematic reviews of ○ Economic studies	• Comments • Editorials • Letters • Animal studies • Case studies • Observational studies • Clinical trials • Abstracts • Empirical economic studies
Language	• English	• Other

Date of publication	• Year 2010 – January 2023	• Studies published before 2010

* Advanced-Therapeutic Medicinal Product (ATMP)

2.2.3 Search

In collaboration with a librarian from the University of Sheffield it was decided that the most relevant databases for the umbrella review were MEDLINE, Cochrane, and Excerpta Medica Database (EMBASE). Although, not all research on database combinations reach identical conclusions, this combination of databases have shown satisfactory recall (Bramer *et al.*, 2017). Other databases were considered, for example Scopus, and Cumulative Index to Nursing & Allied Health (CINAHL). The former was excluded because of the overlap with EMBASE and the latter due to the narrow focus on nursing and allied health.

Thus, MEDLINE and EMBASE were accessed through the Ovid platform and Cochrane independently through its website. For both MEDLINE and EMBASE appropriate search filters for economic evaluation and models, and systematic reviews were sourced from Canada's Drug and Health Technology Agency (CADTH) and SIGN, respectively (CADTH, 2022b; CADTH, 2022a; SIGN, 2021b). These filters were combined with search terms for orphan drugs and rare diseases. As part of the search, studies published before 2010, comments, editorials, letters, and animal studies were excluded. For the Cochrane library, an exploded Medical Subject Heading (MeSH) descriptor for rare diseases was used. The literature searches are listed in Appendix : Table 28, Table 29 and Table 30.

As recommend by Booth, Sutton and Papaioannou (2016), a hand search of references and bibliographies of a selection of the most relevant papers from the review was completed. This was followed by a verification process where it was checked if any known and relevant papers were missing from the review.

2.2.4 Study selection

The citations were exported from the databases as Research Information Systems (RIS) files and subsequently imported to Endnote™ 20 (Version 20.4.1). Titles and abstracts were screened by two independent researchers against the inclusion and exclusion criteria.

2.2.5 Data collection process

The papers that met the inclusion criteria after screening title and abstract, were subjugated to full screening by only one researcher. At this stage, abstracts were excluded, because they contained insufficient information for meaningful data analysis. Those papers that met the inclusion criteria were included in the data collection process. This process was divided into three steps: summary of characteristics, critical appraisal, and data extraction.

2.2.5.1 Summary of characteristics

For the study characteristics, a modified extraction list was created based on recommendations for umbrella reviews (Aromataris *et al.*, 2015). The extraction of study characteristics included citations details, type of review, objectives, date range of database searches, number of studies, Joanna Briggs Institute (JBI) checklist rating and reviewer comments.

2.2.5.2 Critical appraisal

For critical appraisal, the JBI critical appraisal checklist for systematic reviews and research syntheses was applied. For illustration, the JBI critical appraisal checklist was included with permission from authors in Appendix : Figure 19 (Aromataris *et al.*, 2015).

For this tool, there is high degree of freedom for deciding on a scoring system. It was decided not to exclude any papers based on quality appraisal to avoid missing any information. The papers were scored according to an approach previously used for the measurement tool, Assessment of Multiple Systematic Reviews (AMSTAR), which to a considerable extent, resembles the JBI checklist. The systematic reviews were divided into three levels according to scores: 8 – 11 (high quality), 4 – 7 (moderate quality) and 0 – 3 (low quality) (Sharif *et al.*, 2013; Shea *et al.*, 2007).

2.2.5.3 Data extraction

The included reviews were carefully assessed to identify broader themes that were considered issues for the economic evaluation of orphan drugs in rare diseases. The issues were extracted and tabulated according to their themes.

2.2.6 Reporting

Although the Preferred Reporting Items for Systematic reviews and Meta-analyses (PRISMA) statement is mainly intended for health interventions, the checklist of reporting items was used where applicable to structure and report the findings of this umbrella review (Page *et al.*, 2021). A compliant PRISMA flow diagram was constructed using the approach referred to as: the PRISMA2020 R package – for users with coding experience, which was developed by Haddaway *et al.* (2022). The process involved editing a Comma Separated Values (CSV) file with the R Commander package followed by creation of a flow diagram using the PRISMA2020 R package (Fox and Bouchet-Valat, 2022). The R code for the PRISMA diagram is available from: https://github.com/TobiasGrand/PRISMA2020-Flow-Diagram.

2.3 Results

This section describes the results of the umbrella review.

2.3.1 Scoping searches

No comparable systematic literature reviews were identified in PROSPERO. However, in the CRD database one review was concerned with policies for orphan drugs and rare diseases, albeit not with the challenges for economic evaluation (Denis *et al.*, 2009).

The search in MEDLINE by Ovid resulted in 1,565 hits and a sizeable proportion of the papers did not come across as relevant to the research question. For example, they were purely clinical in their nature, and noteworthy for this review, did not consider clinical implications in the context of economic evaluations. Nevertheless, the scoping searches helped improve the literature review. For example, it was decided to change the search filter for economic studies to a filter for economic evaluation and models. Furthermore, the following phrases were added to the search strategy: rare condition and orphan condition. More importantly, a search filter for systematic reviews was introduced, because of the difficulties associated with assessing the full body of literature.

Ultimately, the decision to conduct an umbrella review was made, which in this case, was deemed as an appropriate solution. Previously, this approach proved useful in similar situations, where fields of research expanded rapidly, and consequently, resulted in a diffuse body of literature (Alsulamy *et al.*, 2020; De Freitas *et al.*, 2018; Salleh *et al.*, 2017). See Appendix : Table 25, Table 26 and Table 27 for scooping searches.

2.3.2 Search and study selection

The study selection was illustrated by a PRISMA flow diagram in Figure 3. The number of identified records was 282. They were retrieved from the following databases: EMBASE (n = 211), MEDLINE (n = 67) and Cochrane Library (n = 4). A total of 64 references were duplicate records, and they were removed from the study selection.

Moreover, 172 records were excluded during screening of abstract and title, which left 46 studies for full screening. Of those, 11 were excluded due to them not containing components for economic evaluations (n = 5) or systematic reviews (n = 4), or because they were abstracts (n = 2). Overall, 35 studies from the database searches were deemed eligible for inclusion.

The hand search yielded 5 papers, of which four were excluded for the following reasons: no component of economic evaluations (n = 1) or systematic reviews (n = 2), and for not being concerned with orphan drugs (n = 1). It meant that one paper was carried forward from the hand search, which brought the total number of eligible studies to 36. See Appendix : Table 31, for an overview of the papers included for full screening.

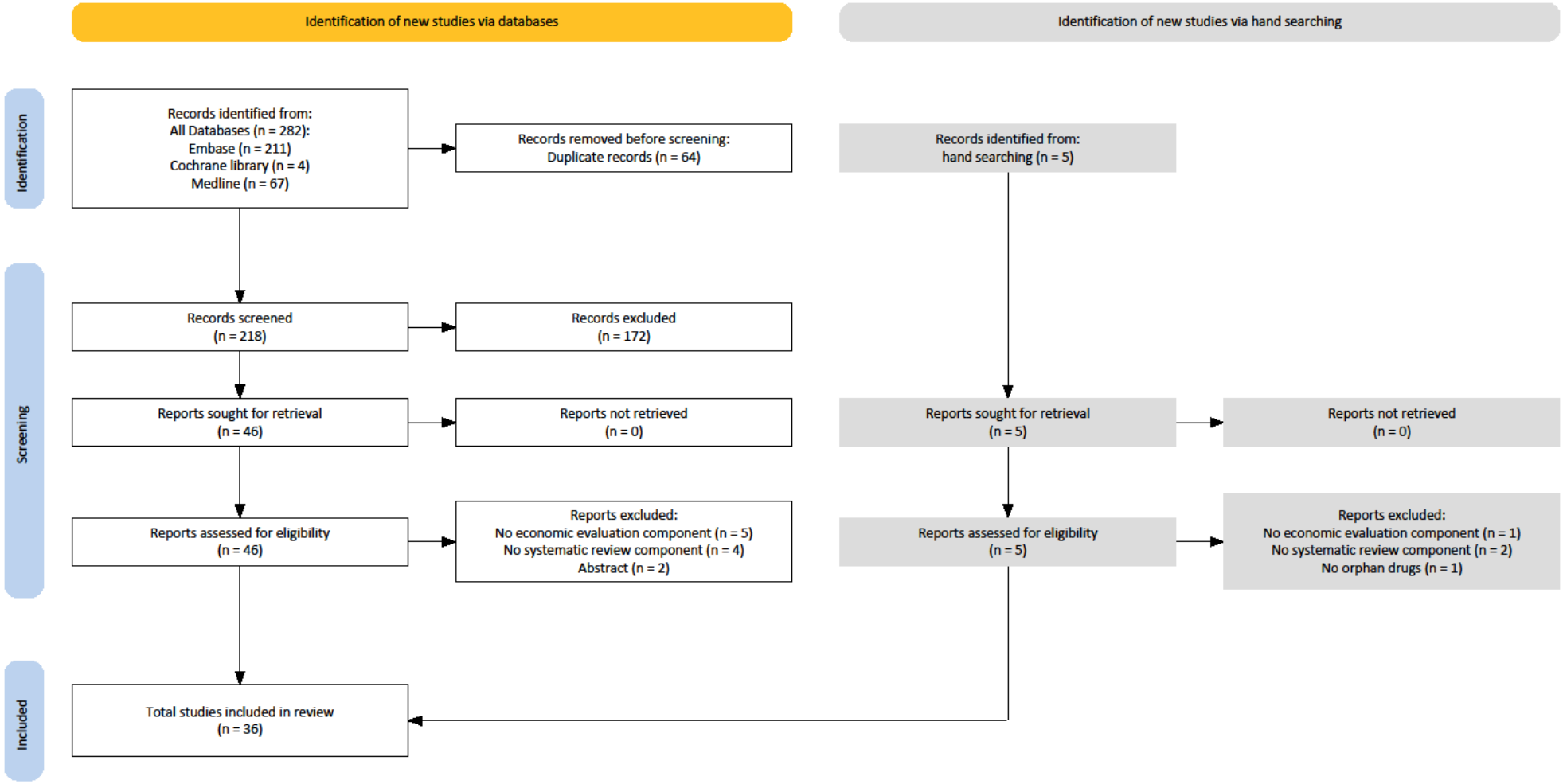
2.3.3 Study characteristics

Two of the identified papers stated that they were scoping reviews but were included because the methods were deemed sufficiently systematic (Lasalvia *et al.*, 2019; Mohammadshahi *et al.*, 2022). Similarly, a study

described their approach as a series of targeted literature reviews, which was also deemed sufficiently systematic for inclusion (Pearson *et al.*, 2018).

Less than half of the studies reported meaningful date ranges for the period where data was collected. In this case, 17 studies provided two time points, as minimum start and end year, which was considered a meaningful range. It was common for studies only to report one timepoint, often for when the database searches were conducted. For example, by month and year, but without specifying a full period for data collection. Other authors stated that their search had no time limit or no restrictions, while providing only one timepoint for data collection. This was considered insufficient as it was not possible to backtrack their respective inclusion periods. The number of records included in the systematic reviews varied between 2 and 338. See Appendix : Table 32, for all extracted study characteristics.

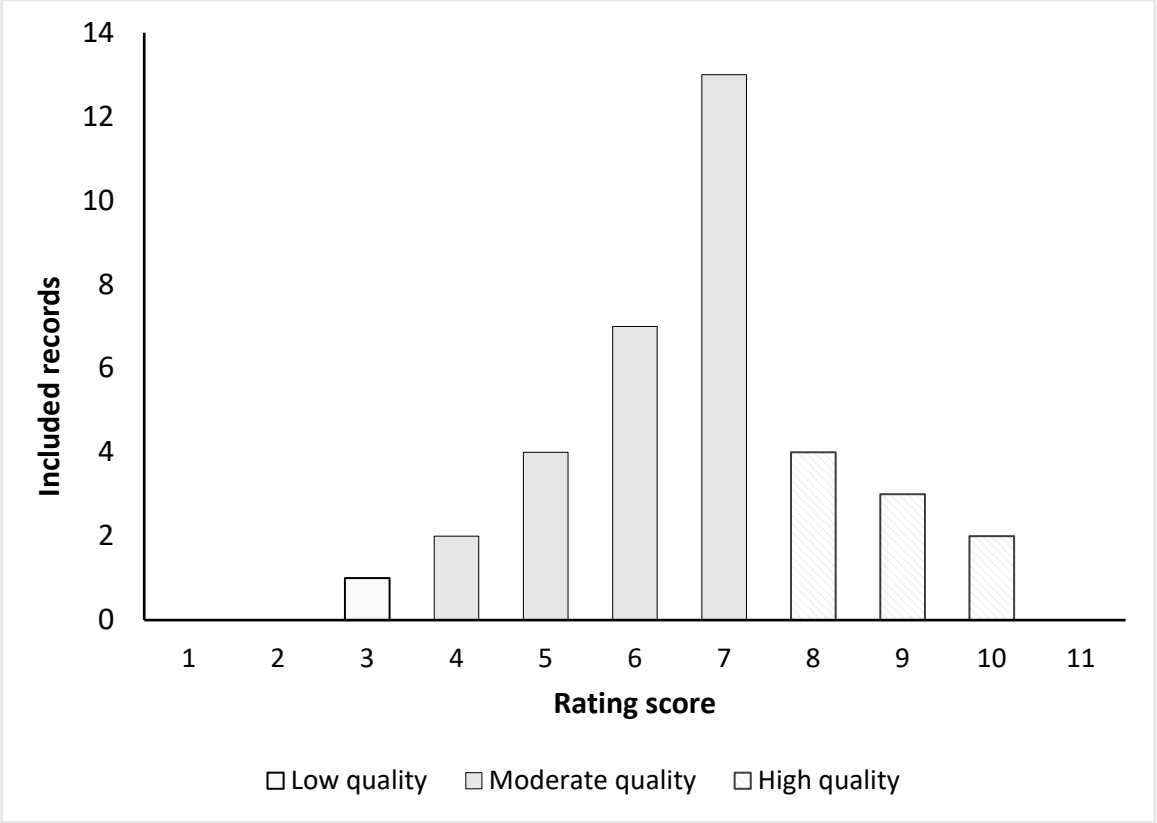
Figure 3: Flow diagram.



2.3.4 Critical appraisal

Figure 4 shows the distribution of quality-assessment scores using the JBI checklist. Only one study had low quality (Lasalvia *et al.*, 2019). Eight studies were rated as having high quality, but the highest frequency was found in the category of moderate quality, which comprised 27 studies. See a detailed list of JBI checklist responses for each study in Appendix : Table 33.

Figure 4: The distribution of JBI (Joanna Briggs Institute) checklist scores across studies.



Most studies obtained points from question 4, which was: were the sources and resources used to search for studies adequate? Question 8 was not widely applicable and was fulfilled by the fewest number of studies. Question 8 was: were the methods used to combine studies appropriate? See Appendix : Figure 20 for accumulated scores across each of the eleven JBI checklist questions.

2.3.5 Data extraction

The identified systematic reviews were carefully evaluated for data-extraction purposes. The systematic reviews were divided into two categories: namely, those who considered a specific rare disease and those who considered multiple rare diseases, 13 and 23 studies, respectively.

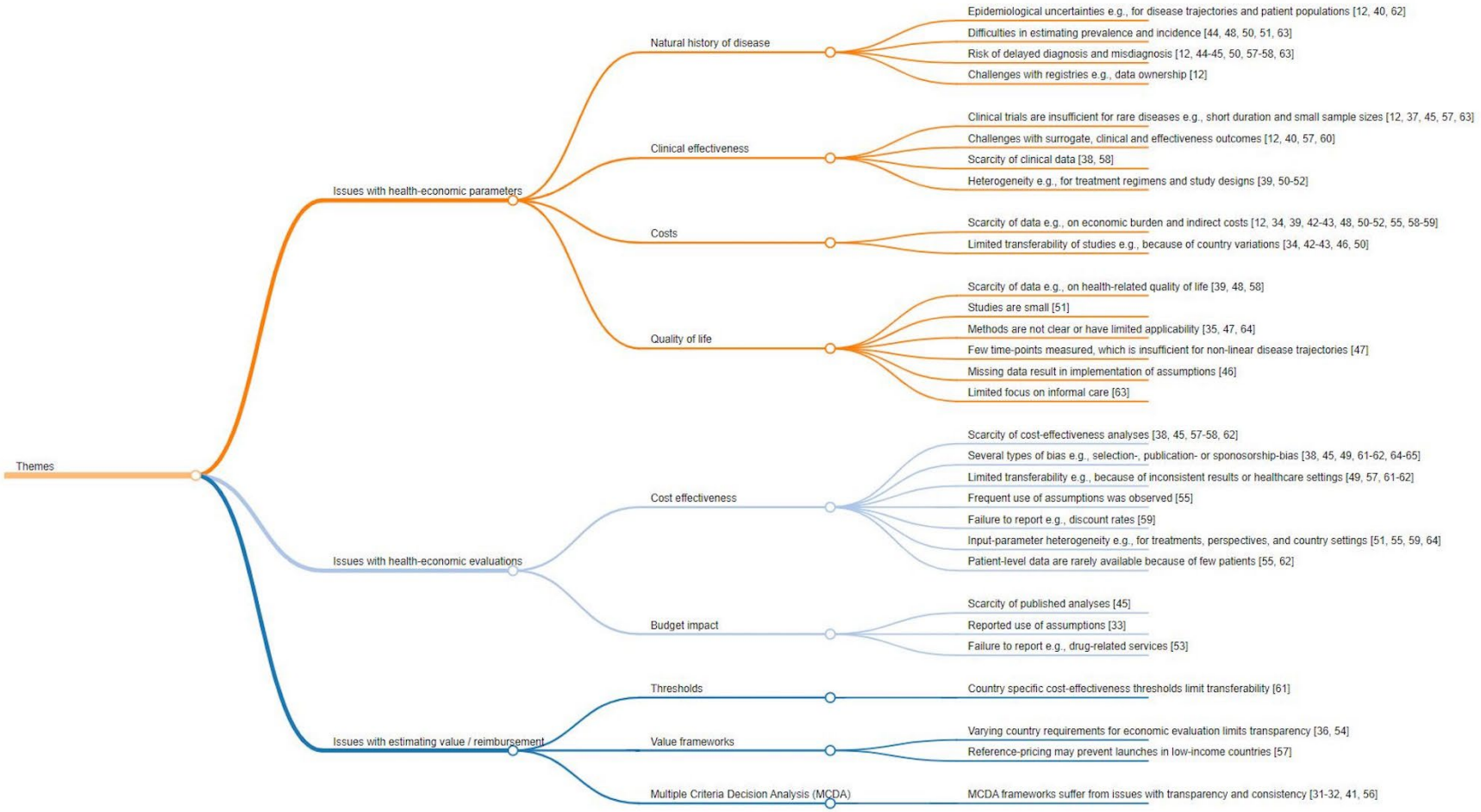
Furthermore, three types of themes were identified: issues with health-economic parameters, issues with health-economic evaluations, and issues with estimating value / reimbursement. For issues with health-

economic parameters, the following sub-topics were identified: natural history of disease, clinical effectiveness, costs, quality of life. For issues with health-economic evaluation the topics were: cost-effectiveness and budget impact. Lastly, for issues with estimating value or reimbursement these were: thresholds, value frameworks and multiple criteria decision analysis (MCDA).

In the category that considered a specific disease, evidence was missing for the following topics: budget impact, thresholds, value frameworks and MCDA. Multiple issues for the economic evaluation of rare diseases were identified during the data extraction. The full data extraction is tabulated in Appendix A: Table 34, to illustrate separate references which provides an indication of commonality for the identified issues.

For illustration, an interactive decision tree (mind map) showing data extraction themes, sub-topics and issues is available from: tobiasgrand.github.io/-data-extraction-themes.github.io/ or from the decision tree which is illustrated in Figure 5. Note that the references in brackets, e.g., [12] reflect the reference list from our publication in Appendix IV.I or (Grand *et al.*, 2024b), but the decision tree was built from Table 34 from Appendix: Tables wherein references are matching this document. The code for the interactive decision tree is available from: https://github.com/TobiasGrand/-data-extraction-themes.github.io.

Figure 5: Decision tree (mind map) illustrating extraction themes, sub-topics and issues.



2.3.6 Issues with health-economic parameters

An economic evaluation requires distinct types of evidence sources. Input parameters on natural history of disease, clinical effectiveness, costs, and quality of life are common types of evidence. The following paragraphs present challenges for each of them in the context of rare diseases.

2.3.6.1 Natural history of disease

Clinical trials are often insufficient to provide information on natural history of disease, because of their short duration. Rare diseases often progress slowly or are chronic by nature (Schuller, Hollak and Biegstraaten, 2015; Pearson *et al.*, 2018). This is an issue for economic evaluation, which often requires long-term modelling of rare-disease trajectories, e.g., using surrogate markers (Faulkner *et al.*, 2019).

The non-existence or limited number of studies which include data on prevalence and incidence further magnify issues (Kwon *et al.*, 2021; Raut *et al.*, 2022). Although, registries can alleviate data issues, they may suffer from challenges such as diverging disease and diagnostic codes, data ownership, missing comparator data, and biases (Querol *et al.*, 2021; Kwon *et al.*, 2021; Pearson *et al.*, 2018).

Furthermore, delayed diagnosis and misdiagnosis make it difficult to define treatment eligible cohorts (Raut *et al.*, 2022; Lee *et al.*, 2022; Querol *et al.*, 2021; Kanters *et al.*, 2013). Clinical experts are few and private practitioners may only encounter few rare-disease cases, which make them difficult to diagnose, and expert advice on rare diseases might not be easy to find (Kanters *et al.*, 2013; Pearson *et al.*, 2018; Zelei *et al.*, 2016).

2.3.6.2 Clinical effectiveness

Clinical trials are common sources for effectiveness data in economic evaluation of rare diseases. Unfortunately, they may not be available, or if they are available, they may suffer from small sample sizes, premature termination, inadequate power, missing data or missing placebo arms, e.g., for ethical reasons (Kanters *et al.*, 2013; Pearson *et al.*, 2018; Zelei *et al.*, 2016; Knoble *et al.*, 2018; Lee *et al.*, 2022).

Differences across study designs, e.g., for clinical outcomes, time-points measured, treatment definitions and reporting, hinder the use of clinical evidence. Other challenges are missing treatment guidelines, data to predict treatment responses, and concerns on the patient-relevance and use of surrogate endpoints (Querol *et al.*, 2021; Onakpoya *et al.*, 2015; Rice *et al.*, 2018; Sequeira *et al.*, 2021; Faulkner *et al.*, 2019). What is more, published long-term studies providing post-marketing data on safety and efficacy are rarely available (Cheng *et al.*, 2012). Onakpoya *et al.* (2015) found that studies reporting clinical evidence had low to moderate quality, and none of them had high quality.

Studies naturally suffer from common evidence challenges: limited number of trials, small study samples and short duration of trials. Obtaining paediatric data might pose additional challenges. For example,

aggregated data for adults and children further limit the potential use of data, e.g., if there is a need to model a paediatric sub-group individually (Chen *et al.*, 2021).

Lastly, information on comparators or standard of care is essential as economic evaluation is a comparative analysis. Data for rare diseases might be missing, and if they are available there might not be consensus on the use of treatment regimens or treatment eligibility of patients, which result in differences across studies (Dhawan *et al.*, 2019; Kwon *et al.*, 2021; Querol *et al.*, 2021; Raut *et al.*, 2022).

2.3.6.3 Costs

For cost-of-illness or burden-of-disease studies there is a scarcity of evidence, studies might not exist or be limited in numbers for a particular disease (Dhawan *et al.*, 2019; Gruhn *et al.*, 2022; Jakes *et al.*, 2021; Kwon *et al.*, 2021; Rice *et al.*, 2018; Weidlich, Kefalas and Guest, 2016; Angelis, Tordrup and Kanavos, 2015; Garcia-Perez *et al.*, 2021; Pearson *et al.*, 2018). Angelis, Tordrup and Kanavos (2015) found that costs were available if treatments were also available for rare diseases, but availability depended less on disease rarity and severity.

Of those studies available, most are retrospective and only a small proportion of studies report indirect, non-medical, or informal-care costs. (Lee *et al.*, 2022; Leonart *et al.*, 2018; Raut *et al.*, 2022; Angelis, Tordrup and Kanavos, 2015; Pearson *et al.*, 2018). Aggregated primary data are rarely available, hence, studies tend to report patient-reported, claims, or registry data (Garcia-Perez *et al.*, 2021). In addition, estimation of lost productivity does not follow recommendations for handling uncertainty (Garcia-Perez *et al.*, 2021).

It is complicated to transfer results between different rare-disease settings. In short, this is due to differences in study design, methods, and results (Garcia-Perez *et al.*, 2021). As can be expected, a multitude of factors influence generalisability such as data sources, geographical perspective, nomenclature, assumptions, discount rates, unit costs, treatment guidelines and value frameworks (Gruhn *et al.*, 2022; Querol *et al.*, 2021; Angelis, Tordrup and Kanavos, 2015; Katsigianni and Petrou, 2022).

2.3.6.4 Quality of life

Quality of life studies are either non-existent or present in limited quantities, but availability depends on the setting of interest (Dhawan *et al.*, 2019; Knoble *et al.*, 2018; Babac, Damm and Graf von der Schulenburg, 2019). There are data limitations for the quality of life of caregivers (Sequeira *et al.*, 2021). Katsigianni and Petrou (2022) showed that scarcity of evidence may lead to the use of assumptions. They highlighted one example for infant onset Pompe disease, which adopted an assumption of equal utility values across treatment arms.

A probable explanation for evidence scarcity is the limited applicability of quantitative methods for rare diseases such as choice experiments or conjoint analysis (Babac, Damm and Graf von der Schulenburg,

2019). Furthermore, studies tend to be small, not randomised or controlled, which decrease the reliability of conclusions (Raut *et al.*, 2022).

Other analysts highlight shortcomings for reporting. For example, the failure to include utility values, mapping algorithms, or insufficiently describing the elicitation of utility weights. These shortcomings have not escaped criticism, because an attempt to address them could have been made (Lee *et al.*, 2022; Thorat, Neumann and Chambers, 2018).

Another major drawback of studies is that they frequently measure only two time-points, which is not suitable for statistical mediation analysis. Furthermore, data are only sufficient for a linear disease trajectory, and a considerable proportion of rare diseases have non-linear disease trajectories (Knoble *et al.*, 2018).

Moving briefly to the issue of outcomes, trade-offs may occur when having to choose between a generic-preference-based or condition-specific measures for rare diseases. It is easier to administer the former, but it may be insensitive to certain disease characteristics, whereas the latter may suffer from lack of validation (Knoble *et al.*, 2018).

2.3.7 Issues with health-economic evaluations

The following two sections present issues for health-economic evaluations.

2.3.7.1 Cost effectiveness

Economic evaluations for rare diseases are scarce. For example, Lee *et al.* (2022) failed to identify any studies, whereas Kanters *et al.* (2013) noted a remarkable absence of pharmacoeconomic evidence, which include cost-effectiveness studies. A notable opinion on the cause of scarcity is that limited input parameters simply deter people from attempting to construct cost-effectiveness analyses, because it is presumed unachievable (Schuller, Hollak and Biegstraaten, 2015). In brief, the causes are missing patient-level data, high drug costs and inability to measure effects for clinical or quality-of-life outcomes (Weidlich, Kefalas and Guest, 2016; Schuller, Hollak and Biegstraaten, 2015; Zelei *et al.*, 2016).

The difficulties for economic evaluation are driving factors for the use of assumptions to overcome challenges for cost-effectiveness modelling. For example, assumptions on mortality, efficacy, treatment, and complications (Weidlich, Kefalas and Guest, 2016). Similarly, it is commonplace to use modelling techniques such as mapping algorithms or long-term extrapolation for outcomes, because of data limitations (Knoble *et al.*, 2018; Cheng *et al.*, 2012). On the other hand, limited patient numbers coupled with unreliable estimates of effects, symptoms, and complications, suggest that methods such as patient-level-simulation modelling may have limited applicability in rare diseases (Schuller, Hollak and Biegstraaten, 2015).

Additionally, challenges for cost-effectiveness modelling may give rise to biases. The literature searches identified two types of biases: publication bias and selection bias. Publication bias was every so often described in relation to positive-results or industry-sponsorship bias. It may occur when manufacturers decide to publish only if they have favourable cost-effectiveness results, a post-marketing obligation, or an opportunity to adopt favourable input parameters and an advantageous interpretation of results (Plüddemann, Banerjee and O'Sullivan, 2017; Holman, Bero and Mintzes, 2019; Woersching, Borrego and Raisch, 2016; Kanters *et al.*, 2013; Schuller, Hollak and Biegstraaten, 2015).

Firstly, numerous sources suggest issues of publication bias (Cheng *et al.*, 2012; Kanters *et al.*, 2013; Park, Griggs and Suh, 2015; Picavet, Cassiman and Simoens, 2015; Woersching, Borrego and Raisch, 2016). For example, Schuller, Hollak and Biegstraaten (2015) indicated a higher frequency of analyses in countries with post-marketing obligations. Secondly, only Picavet, Cassiman and Simoens (2015) highlighted selection bias to explain conflicting cost-effectiveness results for a particular drug. Failure to address bias should be considered a methodological shortcoming, but this was also highlighted as an issue. For example, Cheng *et al.* (2012) and others found that studies failed to discuss the direction and magnitude of bias, despite using data from potentially biased sources (Park, Griggs and Suh, 2015; Woersching, Borrego and Raisch, 2016).

In support of these findings, Thorat, Neumann and Chambers (2018) found that most studies were industry funded in a systematic review of cost-utility analyses for haemophilia. Furthermore, it was shown that incremental cost-utility ratios were significantly lower when published by industry compared to foundations and academia (Picavet, Cassiman and Simoens, 2015).

Economic evaluations have moderate quality (Leonart *et al.*, 2018). The failure to reach high quality may be partly attributed to utility values that do not account for patient characteristics and disease severity or because economic evaluations omit lifetime horizons for chronic rare diseases (Weidlich, Kefalas and Guest, 2016; Park, Griggs and Suh, 2015). Moreover, problems with reporting are frequently highlighted as another factor which may contribute to insufficient quality. For example, not adequately reporting discount rates, sensitivity analyses, utility weights, patient characteristics, funding sources, and time horizons (Thorat, Neumann and Chambers, 2018; Leonart *et al.*, 2018; Cheng *et al.*, 2012).

For cost-effectiveness results, the most serious disadvantage is transferability (Zelei *et al.*, 2016). Cost-effectiveness analyses are heterogenous, because of modelling variations for treatments, patient populations, time horizons, countries, cost-effectiveness thresholds, settings, year of analysis, comparators, and assumptions (Schuller, Hollak and Biegstraaten, 2015; Picavet, Cassiman and Simoens, 2015; Leonart *et al.*, 2018; Raut *et al.*, 2022; Thorat, Neumann and Chambers, 2018; Weidlich, Kefalas and Guest, 2016; Park, Griggs and Suh, 2015). For these reasons, results are rarely generalisable, which is why attention to modelling variations and a high degree of caution is advised when interpreting them (Park, Griggs and Suh, 2015).

Finally, issues that are not related to transferability may affect results. Naturally, the comparator's price may decide cost-effectiveness outcomes because the analysis compares two or more interventions. In one example, this seemed to be the case with drugs indicated for pulmonary arterial hypertension compared to those for metabolic diseases (Park, Griggs and Suh, 2015; Schuller, Hollak and Biegstraaten, 2015). Furthermore, HTA bodies often report lack of clarity on treatment-eligible patient populations (Faulkner *et al.*, 2019). For low-income countries, uncertainty may further add to the existing financial controversy related to reimbursement of high-priced orphan drugs. As a result, this may prevent or delay reimbursement (Zelei *et al.*, 2016).

2.3.7.2 Budget impact

Studies on budget impact are few, mostly from high-income or native English-speaking countries. If Kanters *et al.* (2013) suggestion is accurate, it is not possible to rule out publication bias as a cause for the scarcity of studies on budget impact.

Budget-impact analyses are low quality and show poor adherence to guidelines (Kanters *et al.*, 2013; Abdallah *et al.*, 2021). Furthermore, it has been highlighted that a proportion of budget-impact studies fail to report side effects, drug-related services, life-extension costs, savings from mortality reductions and validation methods (Schlander, Dintsios and Gandjour, 2018; Abdallah *et al.*, 2021)

A challenge for budget-impact modelling is to account for events that occur within a specified time horizon, because it requires knowledge of changing market dynamics. For example, change of drug status, e.g., change of market exclusivity during the modelled period (Schlander, Dintsios and Gandjour, 2018). Another challenge is the frequent use of assumptions on target population, population size, intervention, comparators, costs, and market uptake.

Schlander, Dintsios and Gandjour (2018) make a critical point on the interpretation of budget-impact-modelling results, which is that the reliability of results may be time-dependent because of changing market dynamics. In relation to this, they may have overlooked the importance of assumptions, but these factors together may explain within country variation across different timepoints (Abdallah *et al.*, 2021; Schlander, Dintsios and Gandjour, 2018).

2.3.8 Issues with estimating value and reimbursement

Cost-effectiveness thresholds, value frameworks, and MCDA are important for economic evaluation of rare diseases. The following paragraphs outline issues for them.

2.3.8.1 Value frameworks and thresholds

As explained earlier, most countries require budget-impact and cost-effectiveness models as part of HTAs, but the appraisal process may vary across countries. In Europe, reference pricing further adds to the

complexity and may prevent launches of orphan drugs in low-income countries (Zelei *et al.*, 2016). What is more, input parameters on prevalence, incidence, number of treatment-eligible patients, and clinical benefits are often required when estimating budget impact and cost effectiveness (Short, Stafinski and Menon, 2015).

Value frameworks may suffer from issues with transparency and consistency. For example, cost-effectiveness thresholds may differ, which means that cost-effectiveness outcomes for orphan drugs are far from universal. In fact, these differences largely make budget-impact and cost-effectiveness analyses country specific (Park, Griggs and Suh, 2015; Baran-Kooiker, Czech and Kooiker, 2018).

2.3.8.2 Multiple criteria decision analysis

The use of MCDA is an emerging value framework for orphan drugs. It offers an opportunity to include a broad range of value criteria. For instance, societal, economic, disease or treatment criteria (Friedmann *et al.*, 2018; Lasalvia *et al.*, 2019). Critics highlight variations in scoring functions for value criteria as a significant limitation. In addition, for decision making, it is difficult to observe consistent recommendations (Friedmann *et al.*, 2018; Zelei *et al.*, 2021).

Friedmann *et al.* (2018) showed that studies typically consider economic criteria unimportant in traditional value frameworks. Instead, the most reported value criterion is disease severity. By contrast, Mohammadshahi *et al.* (2022) found that, in the category of economics, the most important criteria were cost effectiveness and budget impact.

2.4 Discussion

The results indicated multiple issues for the economic evaluation of orphan drugs. However, it was not possible to assert whether the issues and challenges identified for orphan drugs applied less to other drugs, which was part of the research objective. The identified studies did not make this comparison, which is why this part of the research question could not be sufficiently answered. The identified challenges for rare diseases were nonetheless grouped by the following themes: health-economic parameters, health-economic evaluations, value frameworks and reimbursement.

2.4.1 Issues with health-economic parameters

For issues with evidence, the results showed a substantial scarcity of evidence across multiple data sources. Scarcity of data was reported for natural history of disease, clinical effectiveness, costs, and quality of life (Kwon *et al.*, 2021; Raut *et al.*, 2022; Knoble *et al.*, 2018; Lee *et al.*, 2022; Kanters *et al.*, 2013; Zelei *et al.*, 2016; Pearson *et al.*, 2018; Dhawan *et al.*, 2019; Gruhn *et al.*, 2022; Jakes *et al.*, 2021; Rice *et al.*, 2018; Weidlich, Kefalas and Guest, 2016; Angelis, Tordrup and Kanavos, 2015; Garcia-Perez *et al.*, 2021).

Recent papers indicate growing concern among analysts due to the scarcity of evidence. It was previously pointed out that there was simply no easy answers to the problem of assessing evidence for orphan drugs (Ollendorf, Chapman and Pearson, 2018). In this review, this was exemplified by Chen *et al.* (2021) who expressed a hope, rather than an actionable plan, for better availability of clinical trials with longer time-horizons to conduct a thorough analysis of cost-effectiveness for paediatric pulmonary arterial hypertension. In support of this statement, it was further claimed that high drug prices and inability to measure effects would discourage people from even attempting to construct cost-effectiveness analyses (Schuller, Hollak and Biegstraaten, 2015). Nonetheless, it is important to highlight that guidelines recommend designing the analysis first, and then afterwards consider data availability. It is an option to use of expert opinion if little data is available (Caro *et al.*, 2012; Briggs *et al.*, 2012).

Authors suggested to practice timely planning of registries to alleviate issues for clinical evidence. Namely, because registries have the potential to inform modelling on natural history of disease or to help construct a replacement for standard of care which may be relevant for trials without a control arm (Schuller, Hollak and Biegstraaten, 2015; Pearson *et al.*, 2018; Sequeira *et al.*, 2021).

Although, surrogate markers can play a vital role when clinical trials have short durations, they may be difficult to validate without long-term data (Zelei *et al.*, 2016). Readers may consider this a validation paradox. A critical solution to this problem is to use research from elsewhere. Pearson *et al.* (2018) have drawn attention to this matter and highlighted the importance of consulting experts and explore more prevalent diseases to source data.

Other authors further add suggestions for improving the status quo for clinical evidence. For example, investigating the impact of misdiagnosis, geographical variation on treatment patterns, reporting of side effects, long-term trials in disease areas with little evidence, and a Cochrane collaboration to reduce evidence gaps (Querol *et al.*, 2021; Kwon *et al.*, 2021; Onakpoya *et al.*, 2015; Chen *et al.*, 2021).

Authors experienced overlapping issues with costs-of-illness or burden-of-disease studies, which led them to recommend the following: firstly, studies should be clear on their perspective; secondly, report costs separately for indirect costs, e.g., lost productivity, and direct costs; thirdly, report costs associated with prevented comorbidities; fourthly, provide clarity on applied discount rates (Sequeira *et al.*, 2021; Leonart *et al.*, 2018; Garcia-Perez *et al.*, 2021; Angelis, Tordrup and Kanavos, 2015).

The importance of future research for informal care, in terms of costs and quality of life, was highlighted by multiple authors, because rare diseases may have severe implications for the closest providers of care, e.g., family and friends (Sequeira *et al.*, 2021; Angelis, Tordrup and Kanavos, 2015; Weidlich, Kefalas and Guest, 2016).

For quality-of-life data, few papers did not provide much direction and made a somewhat classic statement: more research is needed for rare diseases (Kwon *et al.*, 2021; Leonart *et al.*, 2018). Other studies supported the notion of including three or more time-points for measurement of quality of life. They further emphasised the importance of thorough reporting for qualitative methods, like for quantitative methods, although the latter is less applicable to rare diseases (Knoble *et al.*, 2018; Babac, Damm and Graf von der Schulenburg, 2019).

2.4.2 Issues with health-economic evaluations

Multiple studies reported a scarcity of cost-effectiveness modelling studies (Lee *et al.*, 2022; Kanters *et al.*, 2013). As alluded to earlier, it could suggest a strong link between evidence issues and paucity of cost-effectiveness studies (Schuller, Hollak and Biegstraaten, 2015).

As explained earlier, heterogeneity limits the comparability of results (Zelei *et al.*, 2016). Thus, Leonart *et al.* (2018) proposed the principle that if analyses for a specific intervention report consistent results across multiple settings, researchers may show more confidence in the superiority of this intervention. This principle could prove helpful for those who are not yet familiar with challenges for economic evaluation of rare diseases.

Park, Griggs and Suh (2015) want economic evaluations with higher quality and extended time-horizons. To achieve this aim, without conducting a clinical trial, researchers could evaluate: managed-entry agreements, registries for data collection, patient surveys to assess burden of disease, Delphi techniques for validation, expert opinion for estimation, match-adjusted indirect comparison to account for patient characteristics, and rare events with high costs (Pearson *et al.*, 2018; Thorat, Neumann and Chambers, 2018).

Previous paragraphs demonstrated issues for cost-effectiveness modelling, it is now necessary to discuss explanations for the paucity of budget impact models. In terms of input parameters, less are needed, but issues with estimating prevalence or incidence may contribute to their paucity (Kwon *et al.*, 2021; Raut *et al.*, 2022). Budget-impacts models were low quality and rarely validated. Summarising recommendations for improvement, they simply were that researchers should adhere to guidelines (Mauskopf *et al.*, 2007; Abdallah *et al.*, 2021).

The introduction of an orphan drug requires displacement of other treatments to avoid exceeding budgets. HTA bodies often require budget-impact models, but for manufacturers, being the cause of increased healthcare spending might not be a message worth communicating, thus providing an explanation for potential publication bias (Short, Stafinski and Menon, 2015; Kanters *et al.*, 2013).

2.4.3 Issues with estimating value and reimbursement

There has been much debate on the appropriateness of value frameworks in the context of rare diseases. For traditional value frameworks, examples of proposed solutions are: weighting of QALYs according to disease severity and prevalence, categorising QALYs based on disease states, implementing higher cost-effectiveness thresholds, special rules for those that exceed thresholds, e.g., managed-entry agreements and stopping rules for cost containment (Zelei *et al.*, 2016; Pearson *et al.*, 2018).

The UK is an example where some of these measures have been incorporated through the Innovative Medicines Fund or Highly Specialised Technology (HST) evaluation (National Institute for Health and Care Excellence, 2022b; National Institute for Health and Care Excellence, 2022c).

As highlighted throughout this review, criticism of traditional value frameworks has partly been related to their limited transparency and transferability of results. Critics have suggested policymakers to explore other frameworks, e.g., MCDA. So far, this method has only seen sporadic implementation, but it is clearly emerging (Baran-Kooiker, Czech and Kooiker, 2018; Friedmann *et al.*, 2018). The benefit of MCDA is the ability to include a range of value criteria, e.g., burden on caregivers (Friedmann *et al.*, 2018). However, like traditional frameworks, transferability and transparency for MCDA are areas that warrant further research (Friedmann *et al.*, 2018; Zelei *et al.*, 2021).

Recent publications indicate that MCDA is not yet, a strong alternative to traditional value frameworks if the aim is to market more drugs for rare diseases. A study showed that better evaluation outcomes are not guaranteed, because their comparison found slightly more favourable outcomes for orphan drugs evaluated through HTA as compared to MCDA (Lasalvia *et al.*, 2019). Zelei *et al.* (2021) found that twelve out of sixty-two papers did not include the rationale for scoring functions. Furthermore, scoring functions were overlapping, which can lead to double counting (Baran-Kooiker, Czech and Kooiker, 2018). Finally, one point worth mentioning, using a different value framework will not solve issues of evidence scarcity.

2.4.4 Recommendations for further research

From the identified papers, it was clear that challenges were abundant and that solutions were not plentiful and rarely forthcoming. Stakeholders, however, must recognise that certain types of research are costly and requiring these could further eliminate company incentives to research rare diseases (Zelei *et al.*, 2016). For example, clinical trials with extended time horizons. Thus, there is a need for recommendations that are more sustainable. The following paragraphs report recommendations to improve economic evaluations.

2.4.4.1 The potential of comprehensive and flexible cost-effectiveness models

Data availability is critical at the time of economic evaluation for rare diseases, which is why economic models should be transparent and set up for continuous updating as data become available over time

(Leonart *et al.*, 2018). Continuous updating of cost-effectiveness models with new data is an unexplored opportunity, especially, considering the necessity of post-launch-monitoring or real-world data (Onakpoya *et al.*, 2015; Pearson *et al.*, 2018).

As for continuous updating, the identified papers failed to fully define such a framework. Fortunately, advances in data sciences have provided researchers, manufacturers, and payers with the tools to automate updating of cost-effectiveness models. Such a framework has been referred to as living HTA, and apart from automatic updating, it may better protect data privacy or intellectual property (Smith, Schneider and Mohammed, 2022; Thokala *et al.*, 2023). Furthermore, transparency may increase for other stakeholders that are not trained researchers because user-friendly interfaces allow them to "safely" explore model scenarios without having to face backend code (Smith and Schneider, 2020).

Consequently, for economic evaluation of rare diseases, there is untapped potential for using living HTA. In other words, continuous updating of flexible models that contain coded functions that read input parameters from data sheets, rather than, e.g., models in software which require error prone copy pasting between cells. The degree of automation can vary depending on the data sources, but the process of updating with new data as they become available will be more efficient and likely reduce uncertainty within the models. What is more, it has been recommended to use cost-effectiveness models in rare diseases to facilitate EVPI analysis using inputs from, e.g., phase II or registry data (Pearson *et al.*, 2018). In a best-case scenario, it provides researchers with an opportunity to address root causes of uncertainty by reprioritising or initiating data-collection efforts. For example, before initiation of a phase III trial or HTA.

It was discouraging that none of the identified studies used this approach. The recommendation is for cost-effective models to be comprehensive and flexible, and in their analysis report value of information as suggested by Pearson and colleagues (Pearson *et al.*, 2018).

2.4.4.2 Effect of publication bias on the ability to meet cost-effectiveness thresholds

Publication bias was the type of bias most frequently cited by authors (Schuller, Hollak and Biegstraaten, 2015; Woersching, Borrego and Raisch, 2016; Park, Griggs and Suh, 2015; Kanters *et al.*, 2013; Cheng *et al.*, 2012). Not everyone accounts sufficiently for this, for example Picavet, Cassiman and Simoens (2015) would appear overambitious in their claim that orphan drugs can achieve cost-effectiveness when evaluated by common cost-effectiveness thresholds, which could imply that there is no disadvantage for orphan drugs in economic evaluation. The problem is that their study sample is not fully representative for economic evaluations of rare diseases, because of suspected publication bias within their sample.

The sample is not representative because studies were identified mainly from the literature, 23 from Tuft registry (including MEDLINE) and 167 from EMBASE of which 106 analyses met their inclusion criteria (sources not further specified), and if the hypothesis of publication bias is correct, there must be a higher

likelihood that these studies were cost-effective, simply because they were published. Not surprisingly, the majority of identified studies were industry funded. Their conclusion would have higher credibility if publication bias was reduced in their sample of published cost-effectiveness analyses.

In other words, if the analogy of an iceberg is used to illustrate these points: their conclusion is based on the tip of the iceberg, which is made up, only of publications with favourable outcomes. They fail to consider everything beneath the surface, which is where one will find unpublished studies, previously submitted to HTA bodies, who failed to meet cost-effectiveness thresholds. This results in a flawed conclusion.

Ongoing efforts to improve reimbursement conditions for orphan drugs could lose momentum if policy makers take their conclusion at face value. The overall conclusion that cost-effectiveness analyses can meet common cost-effectiveness thresholds seems strongly contested by the findings of this review. The research would have been more convincing if the authors had considered cost-effectiveness analyses submitted to HTA bodies as compared to those available in literature. It requires further research to determine the effect of publication bias on the ability to meet cost-effectiveness thresholds.

2.4.4.3 The importance of estimating prevalence

Turning now to the issue of missing prevalence and incidence data. I argue in coming paragraphs that prevalence is important on more than one level throughout the life cycle of orphan drugs. Firstly, on a regulatory level because meeting prevalence thresholds may result in various research incentives. Secondly, it is important for budget-impact modelling. Thirdly, prevalence thresholds may provide access to attractive economic-evaluation processes.

The US and EU reward orphan drugs with tax exemptions and market exclusivity if they meet the definition of a rare disease. Both countries include a measure of prevalence. For example, the disease affects less than 200,000 persons in the United States or not more than five in 10,000 persons in the EU.

The importance of budget impact models has soared, from being inconsequential to a primary concern for policymakers (Schlander, Dintsios and Gandjour, 2018; Zelei *et al.*, 2016). This might be due to the surge of orphan drugs (Miller, Fermaglich and Maynard, 2021; European Medicines Agency, 2022c). Prevalence is a primary input parameter in these models used to answer: how many patients are eligible for treatment?

In the UK, there is an opportunity for evaluation through the HST evaluation, which includes special conditions for orphan drugs. Generally, depending on the HTA pathway, examples are weighting of QALYs depending on size of benefit or a higher cost-effectiveness threshold (National Institute for Health and Care Excellence, 2022d; National Institute for Health and Care Excellence, 2022c; National Institute for Health and Care Excellence, 2017a). Although, guidelines for HST evaluation define, e.g., routing criteria as small target groups, concentrated in few specialised centres. Short, Stafinski and Menon (2015) interpreted that as a prevalence threshold of less than five hundred patients.

To summarize, reliable prevalence estimates are important for orphan drugs. It is problematic for stakeholders in rare diseases, e.g., manufacturers, that researchers indicate issues for estimating prevalence. It is hard to determine prevalence without quality data (Sequeira *et al.*, 2021; Garcia-Perez *et al.*, 2021). Thus, it is important to find ways of estimating prevalence across multiple settings. For example, an orphan drug in Europe would need prevalence estimates for: the EU, countries, and even within-country estimates, e.g., for regional budget-impact analyses.

2.4.5 Limitations

Important limitations to this review, which need consideration when interpreting the research are as follows:

- Two researchers conducted screening of titles and abstracts, but only one reviewer conducted the full screening and quality assessment. For this reason, the reliability could have been higher. To make up for this, I transparently report full screening and quality assessment in Appendix .
- Exclusion of studies that did not qualify as systematic reviews meant that there was a chance of missing valuable information. Such an example was a narrative review of orphan drugs, which could have supported my findings (Ollendorf, Chapman and Pearson, 2018). Moreover, the search only included studies from 2010. However, the literature searches were partly based on search filters, which balanced sensitivity and specificity. In addition, the identified papers showed a high degree of cross-referencing, which indicate capture of relevant studies.
- All studies were included, no matter their quality rating. The purpose was to maximise inputs into the study. This resulted in the inclusion of one study with a low-quality rating (Lasalvia *et al.*, 2019).
- ATMPs were excluded from this analysis, even if they were considered orphan drugs. It has been much debated whether they should qualify as drugs because the production process typically involves modifying cells or genes. There are challenges for economic evaluation of ATMPs. Thus, the identified opportunities for orphan drugs could apply equally to them. However, not everything is the same, they are frequently curative with a one-off cost, with major challenges being and long-term uncertainty (Angelis, Naci and Hackshaw, 2020; Coyle *et al.*, 2020; Fiorenza *et al.*, 2020). Furthermore, it was previously suggested to consider economic aspects for curative and non-curative treatments separately (Zelei *et al.*, 2016).

2.5 Conclusion

This umbrella review set out to determine issues for the economic evaluation of orphan drugs. The most obvious finding to emerge from this study was scarcity of evidence for clinical effectiveness, costs, quality of life, and natural history of disease. Scarcity of evidence and publication bias emerged as possible causes for

the limited quantity of economic evaluations from literature. The results support the notion that economic evaluation of rare diseases is challenging.

I recommend that researchers focus on sustainable initiatives and explore flexible cost-effectiveness models, e.g., using living HTA. It is highlighted that further research is required to determine the effect of publication bias on the ability to meet cost-effectiveness thresholds.

Chapter 3

Cost-effectiveness modelling of multiple system atrophy for early health technology assessment

This chapter describes my 2nd study, which was cost-effectiveness modelling of MSA for early HTA. We make use of a hypothetical intervention to demonstrate important aspects of cost-effectiveness modelling for MSA, which is why this is a model demonstration. The paper is published in PharmacoEconomics Open (Grand *et al.*, 2025). In addition, an abstract and poster on the concept of disease-analogues were presented at ISPOR before publication of the manuscript (Grand *et al.*, 2024a).

3.1 Introduction

The background is provided in the coming paragraphs. The study design was chosen to facilitate exploration of recommendations made in the umbrella review while fulfilling important research gaps for MSA.

3.1.1 Rationale

The umbrella review found that the main challenges for economic evaluation of orphan drugs in rare diseases were lack of evidence for clinical effectiveness, costs, quality of life and natural history of disease data, and because health-economic input parameters are needed for economic evaluation, they may partly explain the scarcity of economic evaluations for rare diseases (Grand *et al.*, 2024b).

The review provided multiple recommendations that demanded further research. For example, comprehensive and flexible cost-effectiveness models, effect of publication bias on the ability to meet cost-effectiveness thresholds and the importance of estimating prevalence (Grand *et al.*, 2024b). It was concluded that exploring flexible and comprehensive cost-effectiveness modelling, e.g., timely or early conceptual modelling and sourcing of data from disease analogues, had the highest potential in terms of novelty and academic contribution (Grand *et al.*, 2024a; Grand *et al.*, 2024b; Pearson *et al.*, 2018; Ramsey *et al.*, 2005; Smith and Schneider, 2020; Hart *et al.*, 2020). In addition, these recommendations were considered attainable with modest resources and capabilities, which may further improve adaptation by peers.

Consequently, it was decided not to explore the two other recommendations. Estimating prevalence may be considered more of an epidemiological challenge, which cannot be appropriately solved within the realm of health economics. Exploring publication bias in the context of cost-effectiveness thresholds may result in less of an academic contribution since it is difficult as a single individual to assert an influence on both cost-effectiveness thresholds and publication bias.

The best research opportunity at my disposal to critically assess the selected recommendation was cost-effectiveness modelling for early HTA of MSA. The PhD studentship was partly sponsored by Lundbeck A/S (a

pharmaceutical company) which is actively researching MSA, and they were able to facilitate access to individual patient-level data (IPD) and clinical experts. However, this also meant that the choice of disease was partly opportunistic, e.g., in that other rare diseases were not considered. I do not believe that this decision compromises the study integrity, because MSA is a rare-disease archetype with properties that are common for other rare diseases. For example, delayed diagnosis, misdiagnosis, more than one phenotype, inadequate treatment opportunities and low prevalence (Krismer and Wenning, 2017; Mészáros *et al.*, 2020; Gilman *et al.*, 2008; Chrysostome *et al.*, 2004; Bower *et al.*, 1997; Osaki *et al.*, 2011; Sakushima *et al.*, 2015; Bjornsdottir *et al.*, 2013; Chiò, Magnani and Schiffer, 1998; Mazzucato *et al.*, 2014; Grand *et al.*, 2024b). This was important for the validity of the research. Moreover, to my knowledge, no economic evaluation exists for MSA as none were identified, e.g., during literature searching, and little is known about health-economic parameters, and it is not clear if current available data are sufficient to construct an early conceptual cost-effectiveness model for early HTA (Ijzerman *et al.*, 2017).

Lastly, the potential importance of sourcing surrogate data for rare diseases in cost-effectiveness analyses was previously highlighted, which is why it was decided also to explore this opportunity (Pearson *et al.*, 2018; Grand *et al.*, 2024b). However, formal guidance for appropriate surrogate data sourcing for rare diseases are currently missing. In this study, a disease analogue was defined by a disease with overlapping symptoms with MSA to increase applicability of the surrogate data despite not having the same underlying biology. The disease analogues should have higher prevalence than MSA, because of an expected positive correlation between disease prevalence and evidence availability, which ought to increase time-efficiency of the process (Grand *et al.*, 2024a; Shourick, Wack and Jannot, 2021).

3.1.2 Objectives

The study aimed to address the following research objectives:

- How can a conceptual model be developed for MSA to support cost-effectiveness analysis in early HTA, and what evidence gaps currently limit the reliability of such modelling?
- Can data gaps in early cost-effectiveness modelling be reduced, or model development facilitated, by conducting literature searches for disease analogues?

3.1.3 Multiple system atrophy

MSA is a rapidly progressing neurodegenerative disorder with a survival of between 6.2 to 9.8 years from the onset of symptoms (Ben-Shlomo *et al.*, 1997; Wenning *et al.*, 1994; Wenning *et al.*, 2013; Low *et al.*, 2015). Furthermore, Watanabe and colleagues group disease deterioration by three disease states: aid-requiring walking; confinement to wheelchair; bedridden, and the median time to reach these states are 3, 5 and 8 years, respectively (Watanabe *et al.*, 2002). The most common cause of death is sudden nocturnal death

(Shimohata *et al.*, 2008). Accordingly, it is considered a debilitating, devastating, and fatal disease (Krismer and Wenning, 2017).

The mean age at onset of disease is between 50 to 59 years. Age is furthermore, the only widely accepted risk factor. Diagnosis is difficult and frequently mistaken for other diseases, for example Parkinson's disease, which results in improper treatment or diminishing opportunities for clinical trial participation (Krismer and Wenning, 2017; Mészáros *et al.*, 2020).

There is no disease modifying treatment available. Thus, treatment aims at managing symptoms, which include orthostatic hypotension, constipation, genitourinary and breathing disorders. The multifaceted symptoms mean that several medical specialities are involved. For example, neurology, cardiology, gastroenterology, and urology. Furthermore, social services are crucial for end-of-life planning and for supporting caregivers (Burns and McFarland, 2020; Flabeau, Meissner and Tison, 2010; Grossauer *et al.*, 2022).

Guidelines define clinical symptoms by two phenotypes, although they frequently overlap. Firstly, MSA-Parkinsonian (MSA-P) where symptoms of Parkinsonism are dominating. For example, bradykinesia, rigidity, tremor, or freezing gait. The similarity of symptoms between early-stage MSA-P and Parkinson's disease, even apply to a proportion of those suffering from MSA-Cerebellar (MSA-C). A fast-declining response to levodopa might be an early sign of MSA-P rather than Parkinson's disease. (Krismer and Wenning, 2017; Mészáros *et al.*, 2020). Secondly, MSA-C shows primarily cerebellar symptoms such as gait, limb ataxia, explosive speech, and cerebellar oculomotor disturbances (Krismer and Wenning, 2017; Mészáros *et al.*, 2020; Gilman *et al.*, 2008).

Although the disease is rare, MSA-P is more frequent in Europe and US, compared to Japan, where MSA-C is dominating (Sakushima *et al.*, 2015). The estimated prevalence in Europe is between 2 to 5 cases per 100,000 people (Chrysostome *et al.*, 2004; Bjornsdottir *et al.*, 2013; Chiò, Magnani and Schiffer, 1998). In the US, one study estimated an incidence rate of 3 cases per 100,000 person-years in the state of Minnesota (Bower *et al.*, 1997). In Japan, the estimated prevalence is between 13 – 20 cases per 100,000 people (Osaki *et al.*, 2011; Sakushima *et al.*, 2015). However, comparability of prevalence results is limited, because of differences in age distributions, methods, and diagnostic criteria (Kaplan, 2024). In addition, definitive diagnosis requires confirmation of the presence of alpha-synuclein, which is detected by biopsy and positron emission tomography (Watanabe *et al.*, 2018; Gilman *et al.*, 2008; Krismer and Wenning, 2017; Monzio Compagnoni and Di Fonzo, 2019; Burns and McFarland, 2020).

The complicated interplay of mechanisms leading to neurodegeneration is yet to be fully understood even though a large amount of research has been devoted to the disease (Krismer and Wenning, 2017; Monzio Compagnoni and Di Fonzo, 2019). As far as the aggregation of alpha-synuclein is concerned, this is likely the

most researched subject, resulting in an improved comprehension in recent years (Monzio Compagnoni and Di Fonzo, 2019). Research shows that a fundamental pathological property of MSA is an abundant aggregation of intracellular protein called glial-cytoplasmic inclusions primarily in the oligodendrocytes, which is mainly consisting of abnormally folded alpha-synuclein.

The key role of alpha-synuclein in the neurodegenerative pathway makes it a desirable treatment target. It might be possible to change the cause of disease by reducing or stopping alpha-synuclein accumulation, in terms of both severity and duration (Mészáros *et al.*, 2020). Compounds have already shown positive effects in pre-clinical models (Meissner *et al.*, 2019).

3.2 Methods

The coming paragraphs introduce the methods that were applied in the study.

3.2.1 Clinical advice

Clinical experts, e.g., movement disorder specialists with experience in treating MSA, e.g., in specialist centres were recruited as advisors. The contact was facilitated by Lundbeck A/S, a series of meetings and emails were exchanged from June 2025 to October 2025. First, to discuss the concept of using disease-analogues to fill the potential data gaps for health-economic parameters, which resulted in an abstract being published (Grand *et al.*, 2024a). Secondly, the clinicians provided inputs to the relevance of model structures and use of input parameters to improve face validity (Caro *et al.*, 2012). This arrangement was continued when work began on Chapter 4.

3.2.2 Selection of disease analogues

Clinical experts were consulted and identified several disease analogues for MSA. The shortlisted diseases were amyotrophic lateral sclerosis (ALS), multiple sclerosis (MS), and loss of ambulation through, e.g., late-stage Parkinson disease or spinal cord injury (Grand *et al.*, 2024a). I include all disease analogues here, but in the publication, the concept was further refined, which meant that spinal cord injury was excluded to sharpen the focus, as this is an injury of the central nervous system rather than a disease of the central nervous system (Grand *et al.*, 2025). They were chosen because of overlapping symptoms with MSA, e.g., with significant impact on functioning, quality of life and survival, and with higher prevalence because of known evidence scarcities for rare disease (Kanters *et al.*, 2013; Grand *et al.*, 2024b). Disease analogues were explored for the sourcing of health-economic parameters and health-economic evaluations to aid model development in MSA.

3.2.3 Literature searches

Separate literature searches were conducted for MSA and disease analogues. Literature searches for MSA were the most comprehensive and were conducted in EMBASE, MEDLINE, and National Health Service (NHS) Economic Evaluation Database (EED). Search filters were used to identify cost-effectiveness analyses, cost analyses (direct and indirect costs), and quality-of-life analyses (health-state utilities) (CADTH, 2022b; CADTH, 2022a; CADTH, 2023b; CADTH, 2023a). However, no economic evaluations, e.g., cost-effectiveness analyses were identified for MSA. Hand searching was used to expand literature searches when critical data for input parameters were missing.

However, the literature searches for disease analogues were made in the same databases as for MSA, but the searches were more focussed. For example, a narrow search filter was applied with a search limit of 5 years for economic evaluations (CADTH, 2024). Please find literature searches in Appendix: Tables.

3.2.4 Model development

The modelling analysis plan (MAP) was endorsed by study collaborators, e.g., clinical experts or supervisors. It is available from the Shiny app, but the majority of content is covered in the coming paragraphs (Chang *et al.*, 2023). The data to inform the model were sourced from published literature or derived from natural history of disease data collected by the European MSA - Study Group (EMSA-SG) (Wenning *et al.*, 2013).

A Markov model, or specifically, a time-dependent cohort state-transition model was developed in R software (R Core Team, 2021). The mean age at diagnosis for MSA is 56 years, which was the cohort age at model entry (Wenning *et al.*, 2013). The time horizon was 44 years and cycle length 1 month, which resulted in 528 model cycles. The time horizon reflects that survival distributions are highly skewed in natural history of disease studies, e.g., few subjects remain alive even at 12 or 25 years, depending on whether survival is estimated from probable diagnosis or symptom onset, respectively (Watanabe *et al.*, 2002; Wenning *et al.*, 2013). In addition, the time horizon needed to capture the extended survival caused by a hypothetical intervention. The choice of cycle length was based on two criteria: first, the methods for estimating transition probabilities were flexible and the use of a monthly cycle length produced state occupancy plots with improved aesthetics. Second, observed transition probabilities in transition matrices were more intuitively aligned with the permitted health-state transitions that were illustrated in the model structure if cycle lengths were shorter, which was due to multistate modelling.

The modelling was conducted in a freely available software called R using a Shiny package to assist with development of interfaces (R Core Team, 2023; Chang et al., 2023). The cost-effectiveness model was made available online as a Shiny app for user-friendliness and exploration by healthcare and non-healthcare professionals because these model interfaces can facilitate wider use of models without revealing sensitive

data. It means that non-technical users can interact with the interface and reflect on the model parameters and cost-effectiveness results without facing backend code (Smith and Schneider, 2020).

3.2.4.1 Target population

The targeted population for this analysis was people living with MSA, regardless of whether clinically possible or probable diagnostic criteria were applied (Gilman *et al.*, 2008). The inclusion and exclusion criteria were aligned with the European MSA collaboration work package III. The work package is data on file (not published); however, the criteria are reflected in published literature. For example, The natural history of multiple system atrophy: a prospective European cohort study (Wenning *et al.*, 2013).

Inclusion criteria

The target population encompasses the secondary data collected by EMSA-SG, wherein the inclusion criteria were clinically possible or probable MSA and provision of written informed consent (Wenning *et al.*, 2013).

Exclusion criteria

The exclusion criteria are onset under 30 years of age, family history of a similar disorder, secondary cause by history, secondary cause by investigation, hallucinations unrelated to drugs, dementia according to DSM IV (Diagnostic and Statistical Manual of Mental Disorders – fourth edition), prominent slowing of vertical saccades, vertical supranuclear palsy (down/upward gaze palsy), aphasia, alien limb syndrome, parietal dysfunction, and generalised areflexia.

3.2.4.2 Comparators

The comparator was a hypothetical intervention, which was a necessary but appropriate comparator in the absence of a disease-modifying treatment, and because such an intervention is likely to become an add-on to existing symptomatic treatment or best supportive care (BSC). The treatments for comparison in the economic evaluation were BSC vs. BSC plus a hypothetical intervention, the latter with a modelled ability to delay disease progression and reduce clinical events. Similarly, for NICE's assessment of Parkinson's disease, a critical aim was to halt or delay disease progression into health states that require full-time care, which can represent a large impact on quality of life and costs (National Institute for Health and Care Excellence, 2017b; Winter *et al.*, 2011b).

3.2.4.3 Model perspective

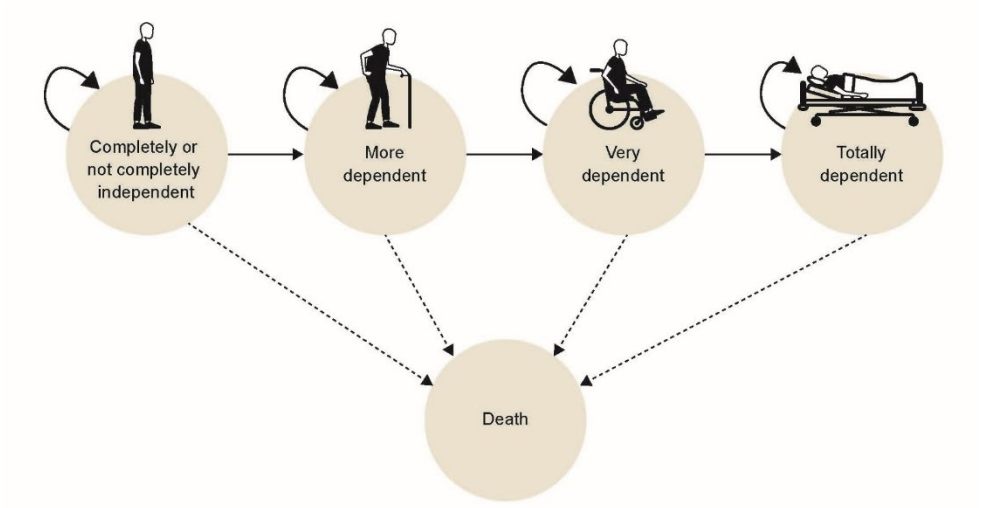
The analysis used the UK NHS perspective. The cost-effectiveness threshold for the analysis was £30,000 per QALY, because £20,000 – £30,000 is regarded as good value for money by the NICE (Bouvy, 2024). The HST process was excluded as the MSA prevalence ranged from 2 to 5 cases per 100,000 people as compared to the HST routing criterion of lower than 1 case per 50,000 people (York Health Economics Consortium, 2016d;

National Institute for Health and Care Excellence, 2024; Fowler, 2011; NHS Trust, 2023). Discount rates were 3.5% for costs and effects as per NICE guidelines (National Institute for Health and Care Excellence, 2022e).

3.2.4.4 Model structure

Clinicians' opinion and findings from the literature searches informed the model structure, which made use of Unified MSA Rating Scale (UMSARS) Part IV (Global Disability Scale), which is a disease-specific functional rating scale (Wenning *et al.*, 2004). Items 1 and 2 were merged as current available data indicated comparable utilities and a previous study reported aggregated costs (Winter *et al.*, 2011b). The model structure has the following health states: completely or not completely independent (walking), more dependent (aid-required walking), very dependent (wheelchair confinement), totally dependent (bedridden) and death. In addition, transition arrows indicated that reversal of disease was not possible, which is aligned with the current lack of a cure. In Figure 6 clinical milestones (images) were illustrated alongside health states. The clinical images were only included to align with a previous illustration that depicts UMSARS Part IV progression together with clinical milestones (Saulnier *et al.*, 2024), and secondly, for easier disease comprehension by non-disease experts as compared to a full item description and clinical explanations. The relationship between UMSARS Part IV and clinical progression milestones (functional milestones) is expanded upon in Chapter 4.

Figure 6: Model structure for multiple system atrophy



The probability for urinary, bowel and hypotension disorders are applied to health states. The first health state completely or not completely independent represents UMSARS part IV items: item 1 (completely independent) and item 2 (not completely

independent). The second health state represents item 3 (more dependent), the third health state represents item 4 (very dependent), and fourth health state represents item 5 (totally dependent). The fifth health state is death (absorbing state).

3.2.5 Estimated health-economic parameters

Estimated health-economic parameters were sourced from data analysis of IPD. The data were European natural history of disease, e.g., survival, UMSARS, and EuroQol-5Dimensions-3Levels (EQ-5D-3L) with a 2-year follow-up, which has been previously published by EMSA-SG (Wenning *et al.*, 2013). A data-analysis report with exploration of EQ-5D (utilities) and UMSARS Part IV (transition probabilities) are available from: https://tobiasgrand.github.io/CEA_MSA_DataAnalysis/. The decision-analytic modelling followed published tutorials and coding framework by Decision Analysis in R for Technologies in Health (DARTH) (Alarid-Escudero *et al.*, 2019; Alarid-Escudero *et al.*, 2023; Smith and Schneider, 2020).

3.2.5.1 Estimating transition probabilities

Table 3 describes the naming conventions for transition probabilities, e.g., probability (p) for moving from state (Ci) to state (Md) is written as p_CiMd.

Table 3: Naming conventions for transition probabilities

	Completely or not completely independent (Ci)	**More dependent (Md)**	**Very dependent (Vd)**	**Totally dependent (Td)**	**Death (D)**
Completely or not completely independent (Ci)	p_CiCi	p_CiMd	p_CiVd	p_CiTd	p_CiD
More dependent (Md)	N/A	p_MdMd	p_MdVd	p_MdTd	p_MdD
Very dependent (Vd)	N/A	N/A	p_VdVd	p_VdTd	p_VdD
Totally dependent (Td)	N/A	N/A	N/A	p_TdTd	p_TdD
Death (D)	N/A	N/A	N/A	N/A	1

European MSA (EMSA) data were captured at baseline, 6, 12, 18 and 24 months, which were used to calculate transition probabilities across UMSARS Part IV. The initial approach was to use a non-parametric count

method, but the results showed that transition probabilities were not constant over the 6 monthly intervals, which could be due to few subjects or insufficient follow-up. See Table 4 for an overview of lowest and highest probabilities captured at six-monthly intervals.

Table 4: Lowest and highest transition probabilities over at six-monthly intervals

	Completely or not completely independent (Ci)	**More dependent (Md)**	**Very dependent (Vd)**	**Totally dependent (Td)**	**Death (D)**
Completely or not completely independent (Ci)	0.55 - 0.8	0.2 - 0.29	0 - 0.16	0 – 0.02	N/A
More dependent (Md)	N/A	0.42 - 0.89	0.11 - 0.54	N/A	0 – 0.06
Very dependent (Vd)	N/A	N/A	0.63 - 0.7	0.2 - 0.26	0.05 - 0.17
Totally dependent (Td)	N/A	N/A	N/A	0.55 – 0.88	0.12 - 0.45
Death (D)	N/A	N/A	N/A	N/A	1

Consequently, multistate modelling was used, which rely on maximum likelihood estimation to endogenously estimate transition probabilities (Srivastava, Latimer and Tappenden, 2021). This method fits a multistate model to the EMSA data using state occupation and time. The calculations were made with the msm R package. The package can be used for discrete time assuming that there is a continuous process underlying the data, despite being mainly designed for models with continuous time measurements (Srivastava, Latimer and Tappenden, 2021; Jackson, 2011).

The transition intensity matrix Q shows instantaneous transitions which can also be thought of as a snapshot of the Markov process. The multistate model is fitted to relevant data, but controlled by the matrix Q (Jackson, 2011). The matrix Q is specified to best reflect clinical reality if patient transitions were captured at the exact time points, e.g., not followed at six-monthly intervals. For example, patients are likely to deteriorate sequentially through the health states. In other words, not moving from Ci to Td, but rather Ci to Md.

$$Q = \begin{pmatrix} -(q12+q15) & q12 & 0 & 0 & q15 \\ 0 & -(q23+q25) & q23 & 0 & q25 \\ 0 & 0 & -(q34+q35) & q34 & q35 \\ 0 & 0 & 0 & -q45 & q45 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Zeros in the matrix Q illustrates that transitions are not possible through these states if precise time was captured at each transition. An initial value for transition intensity is specified in the R code to inform where transitions were possible. However, for the matrix elements showing zeros in the Q matrix, if time intervals are long, then patients can still make these transitions within a cycle. Thus, a transition probability may be observed, but what is observed as the implausible transition from Ci to Vd in the transition matrix, is in theory, Ci to Md to Vd within the time interval t. A shorter cycle length, e.g., 1 month, will bring these clinically unexpected transitions, e.g., p_CiVd closer to zero.

The transition matrix for BSC in Table 5 accounts for both disease-specific and age-dependent all-cause mortality (v_p_Dage). All-cause mortality is informed by death rates by age bands in England and Wales (Our world in data, 2015).

Table 5: Transition matrix for BSC

	Completely or not completely independent (Ci)	**More dependent (Md)**	**Very dependent (Vd)**	**Totally dependent (Td)**	**Death (D)**
Completely or not completely independent (Ci)	(1 - v_p_Dage[i]) * (1 - (p_CiMd + p_CiVd + p_CiTd + p_CiD))	(1 - v_p_Dage[i]) * p_CiMd	(1 - v_p_Dage[i]) * p_CiVd	(1 - v_p_Dage[i]) * p_CiTd	(1 - v_p_Dage[i]) * p_CiD + v_p_Dage[i]
More dependent (Md)	N/A	(1 - v_p_Dage[i]) * (1 - (p_MdVd + p_MdTd + p_MdD))	(1 - v_p_Dage[i]) * p_MdVd	(1 - v_p_Dage[i]) * p_MdTd	(1 - v_p_Dage[i]) * p_MdD + v_p_Dage[i]
Very dependent (Vd)	N/A	N/A	(1 - v_p_Dage[i]) * (1 - (p_VdTd + p_VdD))	(1 - v_p_Dage[i]) * p_VdTd	(1 - v_p_Dage[i]) * p_VdD + v_p_Dage[i]

Totally dependent (Td)	N/A	N/A	N/A	(1 - v_p_Dage[i]) * (1 - p_TdD)	(1 - v_p_Dage[i]) * p_TdD + v_p_Dage[i] _p
Death (D)	N/A	N/A	N/A	N/A	p_DD = 1

* Explanatory notes: v_p_Dage (all-cause mortality vector for probabilities of death by age bands, and transition probabilities constructed by p = probability, from state = Ci, to state = Md resulting in p_CiMd. N/A is not applicable.

Table 6 illustrates how a user-specified treatment modifier, rr_trt_mod_p was multiplied with transition probabilities for the hypothetical intervention to reflect a disease mechanism that slows progression of disease. For example, monthly transition of p_CiMd * rr_trt_mod_p. The use of a treatment modifier was inspired by a cost-effectiveness analysis in human immunodeficiency virus (HIV) by Chancellor and colleagues, which is further highlighted in the book Decision Modelling for Health Economic Evaluation, wherein transition probabilities are adjusted by a relative risk reduction from a meta-analysis for combination therapy as compared to monotherapy (Chancellor *et al.*, 1997; Briggs, Sculpher and Claxton, 2006).

However, the treatment modification was not applied to the health state Td as prolonging a state with a negative utility that by definition is perceived worse than death could be considered unethical (York Health Economics Consortium, 2016f). The clinical experts challenged this definition and argued that they did not regard this health state as one being worse than death despite a negative utility value. However, they would consider it as a health state that could prompt a stopping rule, e.g., stopping treatment.

Table 6: Transition matrix for hypothetical intervention plus BSC

	Completely or not completely independent (Ci)	**More dependent (Md)**	**Very dependent (Vd)**	**Totally dependent (Td)**	**Death (D)**
Completely or not completely independent (Ci)	(1 - v_p_Dage[i]) * (1 - (p_CiMd + p_CiVd + p_CiTd + p_CiD) * rr_trt_mod_p)	(1 - v_p_Dage[i]) * p_CiMd * rr_trt_mod_p	(1 - v_p_Dage[i]) * p_CiVd * rr_trt_mod_p	(1 - v_p_Dage[i]) * p_CiTd * rr_trt_mod_p	(1 - v_p_Dage[i]) * p_CiD * rr_trt_mod_p + v_p_Dage[i]
More dependent (Md)	N/A	(1 - v_p_Dage[i]) * (1 - (p_MdVd + p_MdTd +	(1 - v_p_Dage[i]) *	(1 - v_p_Dage[i])	(1 - v_p_Dage[i]) * p_MdD *

		p_MdD) * rr_trt_mod_p)	p_MdVd * rr_trt_mod_p	* p_MdTd * rr_trt_mod_p	rr_trt_mod_p + v_p_Dage[i]
Very dependent (Vd)	N/A	N/A	(1 - v_p_Dage[i]) * (1 - (p_VdTd + p_VdD) * rr_trt_mod_p)	(1 - v_p_Dage[i]) * p_VdTd * rr_trt_mod_p	(1 - v_p_Dage[i]) * p_VdD * rr_trt_mod_p + v_p_Dage[i]
Totally dependent (Td)	N/A	N/A	N/A	(1 - v_p_Dage[i]) * (1 - p_TdD)	(1 - v_p_Dage[i]) * p_TdD + v_p_Dage[i]
Death (D)	N/A	N/A	N/A	N/A	p_DD = 1

* Explanatory notes: v_p_Dage (all-cause mortality vector for probabilities of death by age bands), rr_trt_mod_p (treatment modifier), and transition probabilities constructed by p = probability, from state = Ci, to state = Md resulting in p_CiMd. N/A is not applicable.

3.2.5.2 Estimating utilities

Three studies from literature searches reported mean EQ-5D index values with standard deviations in parenthesis, and they were between 0.2 (0.3) to 0.56 (0.28) (Schrag *et al.*, 2006; Schrag *et al.*, 2010; Xiao *et al.*, 2022). Others reported on the distribution of MSA patients across EQ-5D dimensions (Winter *et al.*, 2011a; Schrag *et al.*, 2007; Xiao *et al.*, 2022; Schrag *et al.*, 2010). The lack of stratification by model health states meant that data from literature were insufficient for estimation of utilities.

Instead, mean utility index scores were calculated from the EMSA IPD using the eq5d R package with UK EQ-5D-3L preference weights (Morton and Singh, 2024). For simplicity, the utility value was kept constant during sensitivity analysis for the health state Td, because the value was negative and to avoid sampling of implausible values, e.g., below the lower limit of -0.594 of EQ-5D-3L in the UK (Hvidberg and Alava, 2023). Moreover, disutilities were included in the analysis if they were considered to represent short-term (acute) clinical events which were expected to result in a temporary or rapid reduction in quality of life, which might not be captured by six-monthly EQ-5D-3L responses.

3.2.6 Other parameters

Other parameters are those sourced or modified from literature searches.

3.2.6.1 Costs

Only limited and antiquated cost data were found. The 2 studies claimed in their respective papers that no other cost-of-illness studies existed at their time of publication (2011) (Winter *et al.*, 2011b; McCrone *et al.*,

2011). Winter and colleagues investigated German patients and were the only ones to stratify costs by UMSARS Part IV (Wenning *et al.*, 2004). Thus, a simple and necessary way forward was to populate the analysis using this data. Table 7 shows the extracted data from the paper. An implicit and necessary assumption was that Germany costs were like costs in the UK. Subsequently, data required adjustment for inflation from 2011 to 2024, conversion to British pounds sterling, and conversion from six-monthly to monthly costs.

Table 7: Six-monthly cost data estimated from figure 1 in Winter et al. 2011.

Costs from paper (six-monthly)	Inpatient care	Drugs (BSC)	Special equipment	Home care	Other direct costs
Completely or not completely independent (Ci)	€ 1,200.00	€ 800.00	€ 800.00	€ -	€ 1,200.00
More dependent (Md)	€ 3,000.00	€ 1,200.00	€ 1,200.00	€ 200.00	€ 2,500.00
Very dependent (Vd)	€ 4,000.00	€ 1,300.00	€ 4,200.00	€ 1,600.00	€ 3,800.00
Totally dependent (Td)	€ 4,000.00	€ 1,300.00	€ 4,200.00	€ 1,600.00	€ 3,800.00

The cost data from Table 7 were converted from euros to pound using 2011 tariffs and adjusted for inflation using rates from Table 8 . Moreover, data were adjusted from six-monthly to monthly costs.

Table 8: Rates for conversion and inflation adjustments

Conversion rate (€ to £)	£	0.8872	Source: Exchange Rates, UK
Inflation adjustment	£	1.67	Source: Office for National Statistics, UK

Table 9 shows the monthly adjusted costs, values are rounded to 2 decimals, after inflation adjustments and conversion from euro (€) to British pounds sterling (£).

Table 9: Estimated cost data for use in analysis (monthly)

Converted costs for analysis (monthly)	Inpatient care	Drugs (BSC)	Special equipment	Home care	Other direct costs	Total
Completely or not completely independent	£296.32	£197.55	£197.55	-	£296.32	£987.75

(Ci)						
More dependent (Md)	£740.81	£296.32	£296.32	£49.387	£617.34	£2000.10
Very dependent (Vd)	£987.75	£321.02	£1037.14	£395.10	£938.36	£3679.37
Totally dependent (Td)	£987.75	£321.02	£1037.14	£395.10	£938.36	£3679.37

* Note that Drugs (BSC) is symptomatic treatment as part of best supportive care

However, costs were not modelled for short-term clinical events to avoid potential double-counting stemming from the use of aggregated costs. Despite a potential reduction in event driven costs for the hypothetical intervention. The intervention is hypothetical, which is why the intervention costs were set at a £1000 per month, a completely arbitrary value. Users can modify the treatment costs in the Shiny app.

3.2.6.2 Clinical events and disutilities

Short-term (acute) clinical events were included in the analysis if they were independent of model health states (motor function) and expected to result in a temporary reduction in quality of life, which may not be captured by the six-monthly EQ-5D-3L responses. However, this is an area that warrants further research for economic modelling in MSA. In Table 10 the clinical events were grouped by hypotension, urinary or bladder, and bowel disorders, and data further stratified by potential data sources found in literature searches. All identified percentages for clinical events are included in the table, but only those marked in bold font were used in the cost-effectiveness analysis. The data derived by Wenning and colleagues were percentages at baseline, but I assume that they remain constant over time, and source data for clinical events from this study. The clinical events were chosen for consistency as they were derived from EMSA-SG data where utilities and transition probabilities were also sourced from (Wenning *et al.*, 2013).

A user-specified treatment modifier, rr_trt_mod_e was multiplied with event probabilities, e.g., for orthostatic hypotension where 56.7% was converted to a probability of 0.567, for the hypothetical intervention, but similarly to the rr_trt_mod_p, it was not applied to the Td health state.

Table 10: Overview of clinical events in multiple system atrophy.

Clinical events	Percentages	References
Hypotension disorders	71 % orthostatic hypotension / 47.9 % supine hypotension (n = 261)	(Foubert-Samier *et al.*, 2020)
	60 % orthostatic hypotension (n = 27)	(Kollensperger *et al.*, 2007)

	67.4 % orthostatic hypotension (n = 536)	(Saulnier *et al.*, 2024)
	56.7 % orthostatic hypotension (n = 141)	**(Wenning *et al.*, 2013)**
Urinary or bladder disorders	88.8 % bladder symptoms (n = 80);	(Cao *et al.*, 2018)
	48 % urinary retention / 52 % urinary incontinence (n = 27)	(Kollensperger *et al.*, 2007)
	83 % incomplete bladder emptying / 87 % urinary incontinence (n = 175)	(Low *et al.*, 2015)
	68.1 % urinary disorder (n = 536)	(Saulnier *et al.*, 2024)
	73 % urinary incontinence **/ 51.1 % incomplete bladder emptying (n = 141)**	**(Wenning *et al.*, 2013)**
Bowel disorders	**58.2 % constipation (n = 141)**	**(Wenning *et al.*, 2013)**
	87 % constipation / 29 % faecal incontinence (n = 175)	(Low *et al.*, 2015)

* Note that percentages in bold were used in the analysis

An EQ-5D-5L value was identified from literature for MSA with and without orthostatic hypotension, which was 0.571 and 0.638, respectively. The disutility decrement was derived by subtraction to - 0.067 (Xiao *et al.*, 2022). The disutility is higher than in an economic evaluation of MS, where hypertension was - 0.037 but originally sourced from a time-trade off analysis by Stein and colleagues (Stein *et al.*, 2002; Lazzaro *et al.*, 2022). A disutility was not identified for supine hypotension, which is why the disutility value for orthostatic hypotension was used as a proxy for hypotension disorders (Foubert-Samier *et al.*, 2020; Kollensperger *et al.*, 2007; Saulnier *et al.*, 2024; Wenning *et al.*, 2013). This means that I may underestimate the disutility for hypotension disorders due to missing data.

For bowel disorders, cost-effectiveness analyses for MS included a disutility of - 0.240 for gastrointestinal disorders. The value was not possible to verify but one paper cited a non-referenced network meta-analysis and another a guidance document where the original value was not located (Pinheiro *et al.*, 2020; Michels *et al.*, 2019). Instead, I explored the utility for other functional intestinal disorders (ICD-10 Code K59.1 functional diarrhea and K59.09 other constipation) which included constipation (ICD10Data, 2024). The utility value was 0.635, which subtracted from the UK population average (0.829) yields a utility decrement of - 0.194 (Hvidberg and Alava, 2023). This estimate was deemed more conservative than those reported in cost-effectiveness analyses for MS.

The costs of urinary tract infections were commonly included in economic evaluations for disease analogues, but disutilities were not reported (Acosta *et al.*, 2021; Ayati *et al.*, 2021b; Hunter *et al.*, 2021; Martins *et al.*, 2023). For MSA, urinary disorders, e.g., incontinence and incomplete bladder emptying may lead to uncontrollable urinary tract infections (Papatsoris *et al.*, 2008). A MSA study (n = 21) reported that urinary tract infection was the secondary cause of death (23.8%) after cardiopulmonary arrest (33.3%). Similarly, recurrent lower urinary tract infections occurred in 61.3% patients (Papapetropoulos *et al.*, 2007). In the larger study by Wenning and colleagues, urinary disorders ranged between 73% (urinary incontinence) and 51.1% (incomplete bladder emptying). It was decided to use 51.1% for the probability of having urinary disorders that could lead to urinary tract infections to maintain a conservative modelling approach (Wenning *et al.*, 2013).

The utility values for recurrent urinary tract infections in English-speaking women varied but were in the range 0.63 to 0.85 (Le Neveu *et al.*, 2023; Thompson *et al.*, 2023). Similarly, a systematic review of EQ-5D scores showed large variation for diseases of the genitourinary system, e.g., disutilities were between -0.49 (urinary incontinence) to + 0.01 (early kidney disease) (Van Wilder *et al.*, 2019). For use in this analysis, the disutility value (-0.0054) for urinary tract disorder (ICD-9 599) was sourced from an UK EQ-5D catalogue, but by way of explanation, may not fully reflect the severity of recurrent lower urinary tract infections in MSA (Sullivan *et al.*, 2011).

3.2.7 Scenario analyses

Table 11 includes a selection of scenario analyses for exploration, which allows for a practical explanation of parameter interactions and reiteration of data limitations, e.g., for costs. However, readers are encouraged to explore the Shiny app by varying single or multiple parameters and construct their own scenarios. For example, in scenario 6, rr_trt_mod_p was set at the value 0.25 to modify transition probabilities for a hypothetical intervention as illustrated in Table 6. Likewise, rr_trt_mod_e with value 0.25 was multiplied with probabilities for clinical events, thus, reducing the overall disutility associated with these events. Lastly, costs were differentiated for the health states, very dependent (Vd) at £3,000 and totally dependent (Td) at £25,000, to circumvent one of the major data limitations. Namely, that published costs are aggregated, which is moreover an issue for the completely or not completely independent state where costs were reduced to £0.

Table 11: Scenario analyses for exploration

Scenario	Slowing of disease	Reduce clinical events	Monthly cost of Ci* (£)	Monthly cost of Vd (£)	Monthly cost of Td (£)
Base case	0.5	0.5	988	3,679	3,679

1	1	0.5	988	3,679	3,679
2	0.5	1	988	3,679	3,679
3	0.7	0.7	988	3,679	3,679
4	0.5	0.5	988	3,679	8,000
5	0.45	0.45	100	2,500	10,000
6	0.25	0.25	0	3,000	25,000

* Note that slowing of disease is the value for rr_trt_mod_p applied to transition probabilities, and reduce clinical events is the value for rr_trt_mod_e applied to event probabilities. The base-case settings correspond to those in the Shiny app. Abbreviations: Ci (completely or not completely independent, Vd (very dependent) and Td (totally dependent).

3.2.8 Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was conducted to assess uncertainty by random sampling of values from distributions that were assigned to relevant input parameters (York Health Economics Consortium, 2016g). The distributions were either partly informed by literature or derived from IPD. For example, from literature where alpha was estimated by n (subjects) multiped with the mean, and beta by n (subjects) minus alpha. Common distributions were chosen, e.g., beta distributions for utilities, transition, or event probabilities. A gamma distribution was assigned to cost data in the absence of relevant information to quantify uncertainty, but assigning a gamma distribution to cost data is commonplace (Briggs, Sculpher and Claxton, 2006). The distributions are listed in the MAP contained within the Shiny app or directly in model code, which also contains in-code explanations. The model code including Shiny-app code is hosted here: https://github.com/TobiasGrand/CEA_MSA_ShinyApp. The PSA was used to construct a cost-effectiveness plane for readers where they can visually inspect simulations. The default is set at 1000 simulations. It was decided not to conduct one-way sensitivity analyses on top of scenario and PSAs, because the purpose of a Shiny app was to allow readers to conduct their own exploration of parameters and results.

3.2.9 Shiny app

A Shiny app was developed for the early HTA. The app consisted of three main coding components: code for a Markov model, time-dependency (three-dimensional array), and Shiny interface. The user interface has a left side panel and upper navigation tabs; the panel allows the user to change model parameters and update the sensitivity analysis by pressing “RUN MODEL”. The navigation tabs contain model structure, deterministic analysis (fixed), sensitivity analyses (changeable), MAP, introduction for non-technical users, other resources. The command set.seed() in the model code, together with a download button for simulated PSA data, were implemented to facilitate reproduceable scenarios and, ultimately, to compare and scrutinise them. The Shiny app’s parameter specifications are included in Table 12. The R code was reviewed by a second modeller with vast experience and the deterministic analysis double-coded in Excel to see if there were any bugs in the code and if results could be reproduced.

Table 12: Specifications for Shiny app (R project)

Parameter type	Parameter name	Value	PSA	Shiny user specification
Model specifications				
Number of cycles	n_cycles	528	constant	N/A
Names of health states	v_names_states	Ci, Md, Vd, Td, D	constant	N/A
Background all-cause mortality	v_p_Dage	age-dependent	constant	N/A
Discount rate costs	d_c	0.035	constant	0 - 1
Discount rate QALYs	d_e	0.035	constant	0 - 1
Number of PSA samples	n_sim	1000	constant	100-10000
Transition probabilities (across all trial visits)				
Ci to Md	p_CiMd	0.086	rbeta (shape1 = 12.1, shape2 = 128.9)	N/A
Ci to Vd	p_CiVd	0.0047	constant	N/A
Ci to Td	p_CiTd	0.000051	constant	N/A
Ci to D	p_CiD	0.000033	constant	
Md to Vd	p_MdVd	0.010	rbeta (shape1 = 14.1, shape2 = 126.9)	N/A
Md to Td	p_MdTd	0.0016	constant	
Md to D	p_MdD	0.0010	constant	N/A
Vd to Td	p_VdTd	0.030	rbeta (shape1 = 4.2, shape2 = 136.8)	N/A
Vd to D	p_VdD	0.020	rbeta (shape1 = 2.8, shape2 = 138.2)	N/A
Td to D	p_TdD	0.031	rbeta (shape1 = 4.4, shape2 = 136.4)	N/A
D to D	p_DD	1	constant	N/A
Treatment modifier for transitions probabilities (slowing of disease progression)				

Relative risk treatment modifier for transition probabilities	rr_trt_mod_p	0.5	constant	0.01 - 1
Probabilities of events (monthly)				
Probability of urinary disorders	p_UD	0.058	rbeta (shape1 = 72.1; shape 2 = 68.9)	N/A
Probability of hypotension disorders	p_HD	0.067	rbeta (shape1 = 79.9; shape2 = 61.1)	N/A
Probability of bowel disorders	p_BD	0.582	rbeta (shape1 = 82.1; shape2 = 58.9)	N/A
Treatment modifier for probability of events (reduction in number of events)				
Relative risk treatment modifier for clinical events	rr_trt_mod_e	0.5	constant	0.01 - 1
Total costs (monthly)				
Costs for Ci	c_Ci	987.75	rgamma (shape = c_Ci/50, scale = 50)	£0 – 100,000
Costs for Md	c_Md	2000.10	rgamma (shape = c_Md/50 scale = 50)	£0 – 100,000
Costs for Vd	c_Vd	3679.37	rgamma (shape = c_Vd/50, scale = 50)	£0 – 100,000
Costs for Td	c_Td	3679.37	rgamma (shape = c_Td/50, scale = 50)	£0 – 100,000
Price of hypothetical intervention (monthly)	c_Trt	100.00	constant	£0 – 10,000
Utility values				
Utility for Ci	u_Ci	0.5787011	rbeta (125.68; 91.50)	N/A

Utility for Md	u_Md	0.4272593	rbeta (86.68; 116.20)	N/A
Utility for Vd	u_Vd	0.1435197	rbeta (27.44; 163.75)	N/A
Utility for Td	u_Td	-0.1255000	constant	N/A
Disutility values				
Disutility value for urinary disorders*	du_UD	0.0054	rbeta (0.7; 140.3)	N/A
Disutility value for hypotension disorders**	du_HD	0.013	constant	N/A
Disutility for bowel disorders***	du_BD	0.194	constant	N/A

List of terms used in literature: * Urinary disorder, bladder symptoms, urinary retention, incomplete bladder emptying, urinary incontinence ** Orthostatic and supine hypotension *** Constipation and faecal incontinence. Note that a selection of values is rounded for better illustration.

3.3 Results

3.3.1 Literature searches

Literature searches were conducted from the 7th to 14th of May 2024 in MEDLINE and EMBASE. The searches for MSA data relied on search filters for economic evaluation and models, and health-state utilities which yielded 519 papers, of which, 111 were duplicates. It meant that 408 papers were screened for inclusion. The disease analogue searches for economic evaluations provided 287 and 260 papers across MEDLINE and EMBASE. 211 of these were duplicates, which left 335 papers for potential exploration of supportive surrogate data. No studies were identified from the NHS EED databases for both MSA and disease analogues.

3.3.1.1 Cost-effectiveness studies for multiple system atrophy and disease analogues

Cost-effectiveness analyses were missing for MSA despite significant efforts, e.g., hand and systematic literature searches. However, 48 analyses were identified for diseases analogues. Namely, 42, 3, 2, 1 cost-effectiveness analyses from literature searches across MS, advanced Parkinson's disease, ALS, and spinal cord injury, respectively.

First, most models for MS were Markov models that used an Expanded Disability Status Scale (EDSS) to define health states (Auguste *et al.*, 2020; Baharnoori *et al.*, 2022; Cortesi *et al.*, 2022; Daroudi *et al.*, 2023; Ginestal *et al.*, 2023; Giovannoni *et al.*, 2019; Hua *et al.*, 2019; Lasalvia *et al.*, 2020; Rezaee *et al.*, 2019). A proportion of these models further accounted for on- and off-treatment periods or for disease progression,

e.g., remitting / relapsing or secondary progressive MS with and without EDSS scores to inform health states (Ayati *et al.*, 2021a; Ayati *et al.*, 2021b; Bolanos-Diaz, Sanabria-Montanez and Custodio-Capunay, 2019; Furneri *et al.*, 2019; Kremer *et al.*, 2020; Lazzaro *et al.*, 2022; Martins *et al.*, 2023; Michels *et al.*, 2019; Montgomery *et al.*, 2022; Nakhaipour *et al.*, 2020; Navarro and Betancur, 2023; Pinheiro *et al.*, 2020; Poveda *et al.*, 2020; Schur *et al.*, 2021; Spelman *et al.*, 2024; Spelman *et al.*, 2022; Stanisic *et al.*, 2019; Taheri, Sahraian and Yousefi, 2019; Visser *et al.*, 2022; Walter *et al.*, 2019; Xu *et al.*, 2019).

In addition, there were examples of other health states, e.g., responder / non-responder, exacerbation, severity states (mild, moderate and severe) (Acosta *et al.*, 2021; Hunter *et al.*, 2021; Oppe *et al.*, 2021; Zhao *et al.*, 2022; Palmer *et al.*, 2024). In a few cases, the model type was not explicitly reported, but a proportion of these presumably modelled incremental differences in costs and effects directly across comparator arms using either trial or registry data (Dimitrova *et al.*, 2023; Mantovani *et al.*, 2020; Michel *et al.*, 2022; Renfrew *et al.*, 2019; Rosato *et al.*, 2021). Lastly, two studies used a microsimulation model (Huygens and Versteegh, 2021; Versteegh *et al.*, 2022).

Second, for advanced Parkinson's disease all models used Hoehn and Yahr states, which is a measure of clinical disability, together with on- and off-states to reflect periods of controlled and uncontrollable motor symptoms, the latter being unresponsive to treatment (Chaudhuri *et al.*, 2022; Kalabina *et al.*, 2019; Pahwa *et al.*, 2023). Third, for ALS both models were Markov models. The first model had health states for hospitalisation, invasive and non-invasive ventilation (Gajdos *et al.*, 2021). The second model was based on a staging approach called Fine'til 9, which is an approach that is based on a functional rating scale (Thakore *et al.*, 2020). Fourth, for spinal cord injury the model structure was not clear, but the study compared training strategies for practice-based cohorts (Pinto *et al.*, 2023).

See Table 42 in Appendix: Tables for a tabulated overview of cost-effectiveness studies.

3.3.1.2 Quality of life studies for multiple system atrophy

A proportion of studies reported on progression of quality of life, and they were classified as either longitudinal or prospective studies that covered the geographical areas of China, Europe, Japan, and US. All of them included UMSARS, a rating scale for MSA (Cao *et al.*, 2018; Foubert-Samier *et al.*, 2020; Golden, McCreary and Vernino, 2022; Jecmenica-Lukic *et al.*, 2021; Jecmenica-Lukic *et al.*, 2018; Kollensperger *et al.*, 2007; Low *et al.*, 2015; May *et al.*, 2007; Wenning *et al.*, 2013; Watanabe *et al.*, 2002; Saulnier *et al.*, 2024). 36-Item Short Form Survey (SF-36) was included in four studies (Kollensperger *et al.*, 2007; Jecmenica-Lukic *et al.*, 2018; Low *et al.*, 2015; May *et al.*, 2007). The composite autonomic symptom score (COMPASS) was used in two studies (Golden, McCreary and Vernino, 2022; Low *et al.*, 2015). Lastly, two studies included the MSA quality of life scale (Golden, McCreary and Vernino, 2022; Saulnier *et al.*, 2024).

Lastly, three papers focussed on the burden of caregivers. The duration of care provided by caregivers ranged from 8 months to 8 years (Langford, Zhou and Miyasaki, 2023). 39.3 hours of practical care or 31.0 hours of emotional support was provided per week by caregivers if their association had neurogenic orthostatic hypotension, which is a condition caused by underlying neurodegenerative diseases such as MSA (Gallop *et al.*, 2023). Moreover, 63% of caregivers to MSA patients had depression (Miyashita *et al.*, 2009). Lastly, a caregiver quote encapsulates the burden of informal care *"I'm not going to ruin myself doing this, because if I ruin myself, I can't help him"* (Langford, Zhou and Miyasaki, 2023, p. 54).

3.3.1.3 Natural history of disease studies for multiple system atrophy

Survival was followed in five studies by either Kaplan-Meier or Cox-regression curves. The median survival from symptom onset to death ranged from 6.4 – 9.8 years (Cao *et al.*, 2018; Foubert-Samier *et al.*, 2020; Low *et al.*, 2015; Watanabe *et al.*, 2002; Wenning *et al.*, 2013). In addition, studies measured stridor, falls, orthostatic hypotension, depression, apathy, delusions, and urinary incontinence (Cao *et al.*, 2018; Foubert-Samier *et al.*, 2020; Jecmenica-Lukic *et al.*, 2021; Jecmenica-Lukic *et al.*, 2018; Kollensperger *et al.*, 2007; Saulnier *et al.*, 2024). Lastly, two studies monitored clinical manifestations or milestones of progression: falls per day, unintelligible speech, and time to aid-required walking, wheelchair confinement, and bedridden state (Watanabe *et al.*, 2002; Wenning *et al.*, 2013).

3.3.1.4 Cost-of-illness studies for multiple system atrophy

As previously mentioned, only two cost-of-illness studies existed, but cost data stratified by UMSARS Part IV were sourced for this analysis from Winter and colleagues (McCrone *et al.*, 2011; Winter *et al.*, 2011b). In addition, they showed a positive correlation for cost and the following variables: age groups in years, marital status, age at disease onset, depression, and UMSARS Part IV (Winter *et al.*, 2011b). McCrone and colleagues covered France, Germany, and the UK, but this study did not stratify costs by UMSARS Part IV. Patients required a range of services, but informal care was the main contributor of costs. Moreover, costs and severity of disease were correlated.

3.3.2 Cost-effectiveness modelling

For cost-effectiveness modelling several health-economic parameters had to be estimated before results could be generated. This section outlines the estimated model parameters and results.

3.3.2.1 Estimated transition probabilities

Table 13 contains estimated transition probabilities from EMSA IPD by multistate modelling in R.

Table 13: Monthly transition probabilities from multistate modelling

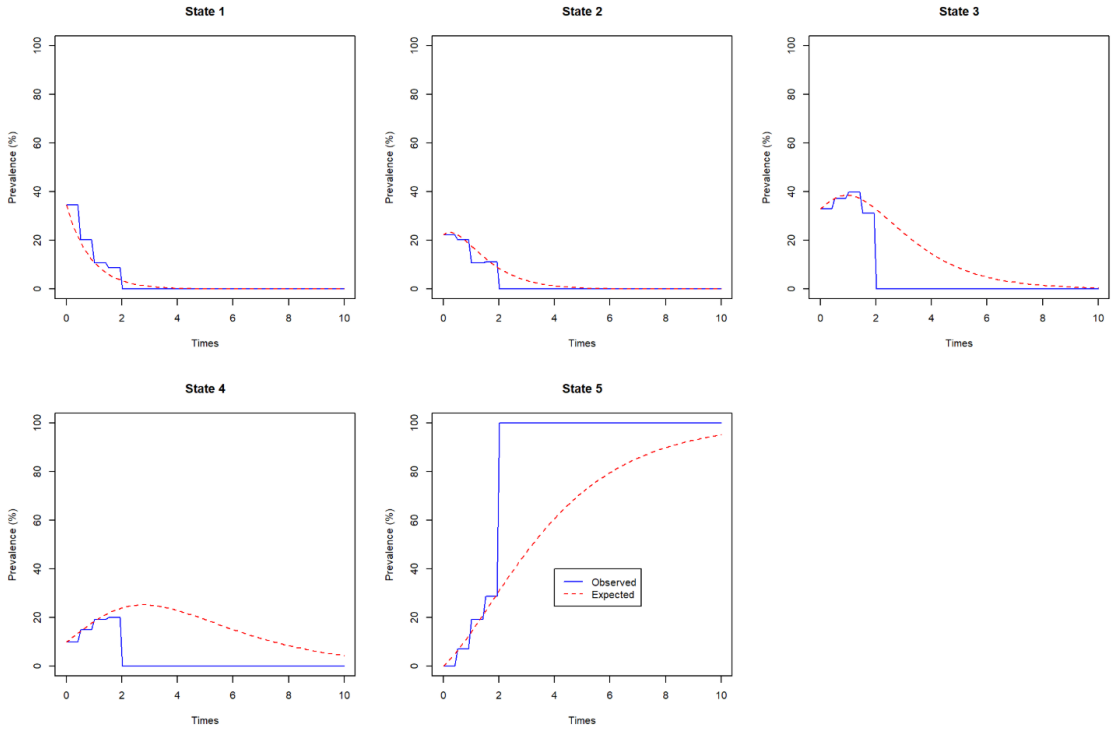
	Ci	Md	Vd	Td	Death

Completely or not completely independent (Ci)	0.91*	0.09	0.00	0.00	0.00
More dependent (Md)	N/A	0.90	0.10	0.00	0.00
Very dependent (Vd)	N/A	N/A	0.95	0.03	0.02
Totally dependent (Td)	N/A	N/A	N/A	0.97	0.03
Death	N/A	N/A	N/A	N/A	1

*Point estimate of transition probability. All values are rounded to 2 decimals.

Figure 7 is an output from the msm package that shows the predicted prevalence (dotted and red line) by time in each state, which was used to estimate transition probabilities by a maximum likelihood function. It is clear from the image that observed data (solid blue line) follow-up was 2 years.

Figure 7: Output from msm package for prevalence by time in each state. State 1 (Ci), State 2 (Md), State 3 (Vd), State 4 (Td), and state 5 (Death)



3.3.2.2 Estimated utilities

Table 14 shows estimated utilities from IPD using the eq5d R package using UK EQ-5D-3L preference weights (Morton and Singh, 2024). The health state utilities are correlated with disease severity.

Table 14: EQ-5D-3L (utilities) for UMSARS Part IV health states

UMSARS Part IV	Mean Index Score
Completely or not completely independent (Ci)	0.58*
More dependent (Md)	0.43
Very dependent (Vd)	0.14
Totally dependent (Td)	-0.13

*Point estimate for mean EQ-5D-3L index score. All values are rounded to 2 decimals

3.3.2.3 Deterministic cost-effectiveness results

The deterministic cost-effectiveness analysis provided an incremental cost-effectiveness ratio (ICER) of £142,036 per QALY in Table 15. A breakdown of costs and QALYs by health states over the model period is included in the Shiny app.

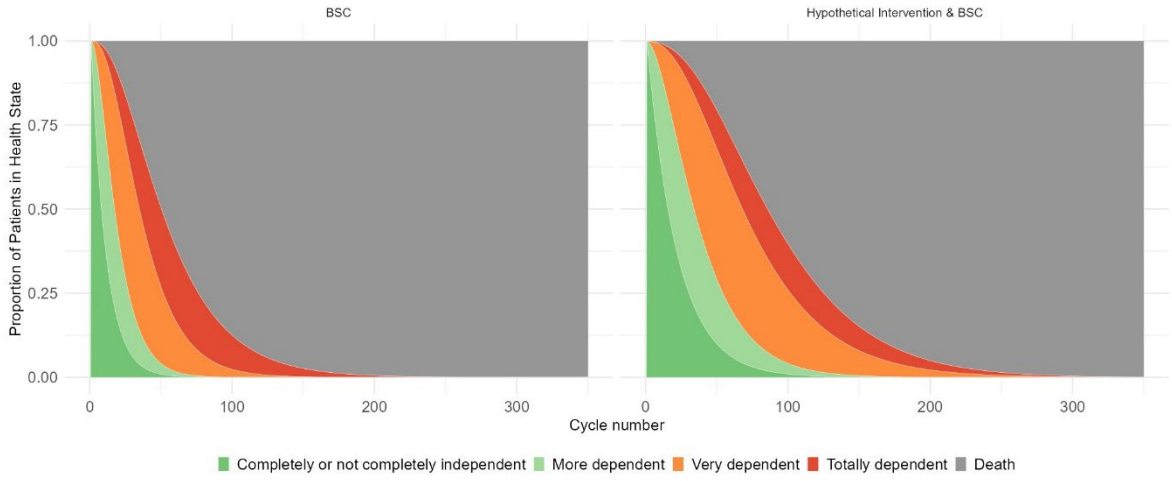
Table 15: Deterministic cost-effectiveness results

Deterministic analysis		Costs	QALYs	ICER
Comparators	Hypothetical Intervention & BSC	£279,782	1.72	
	BSC	£147,651	0.79	142,036

*Note that Drugs (BSC) is symptomatic treatment as part of best supportive care

The state occupation by model cycles in Figure 8 was calculated from the Markov traces. It illustrates the delay in disease progression from BSC plus a hypothetical intervention compared to BSC alone.

Figure 8: State occupation by model cycle



3.3.2.4 Scenario analyses

The scenario analyses were constructed using the Shiny app (sensitivity analysis tab) for replication purposes and the results are presented in Table 16. The ICERs varied, but as expected when considering each of the scenario specifications. Scenario 1 showed that disease modification by a reduction in clinical events alone, but without delayed progression, generated the least QALYs in the hypothetical intervention plus BSC arm. Scenario 6 was within the cost-effectiveness threshold of £30,000 per QALY.

Table 16: Results from scenario analyses

Scenario	Costs	QALYs	ICER

1	Hypothetical Intervention & BSC	£205,536	0.86	
	BSC	£165,846	0.83	1,296,894
2	Hypothetical Intervention & BSC	£304,234	1.75	
	BSC	£165,846	0.83	149,502
3	Hypothetical Intervention & BSC	£251,315	1.26	
	BSC	£165,846	0.83	194,865
4	Hypothetical Intervention & BSC	£369,847	1.81	
	BSC	£242,458	0.83	129,827
5	Hypothetical Intervention & BSC	£351,473	2.01	
	BSC	£244,508	0.83	90,683
6	Hypothetical Intervention & BSC	£589,718	3.38	
	BSC	£519,271	0.83	27,646

* Abbreviations: best supportive care (BSC), quality-adjusted life year (QALY), and incremental cost-effectiveness ratio (ICER). Scenarios from Table 11: #1 (Slowing of disease = 1. Remaining parameters are base case), #2 (Reduce clinical events = 1. Remaining parameters are base case), #3 (Slowing of disease and reduce clinical events = 0.7. Remaining parameters are base case), #4 (Slowing of disease and reduce clinical events = 0.5, and monthly costs of Td (£) = 8,000. Remaining parameters are base case), #5 (Slowing of disease and reduce clinical events = 0.45, and monthly costs of Ci (£) = 100, Vd (£) = 2,500, and Td (£) set at 10,000. Remaining parameters are base case), #6 (Slowing of disease and reduce clinical events = 0.25, and monthly costs of Ci (£) = 0, Vd (£) = 3000, and Td (£) set at 25,000. Remaining parameters are base case).

3.3.2.5 Probabilistic cost-effectiveness results

The cost-effectiveness in the PSA was £141,092 per QALY, which is slightly lower but close to the deterministic analysis, but higher QALYs and costs were generated as illustrated in Table 17.

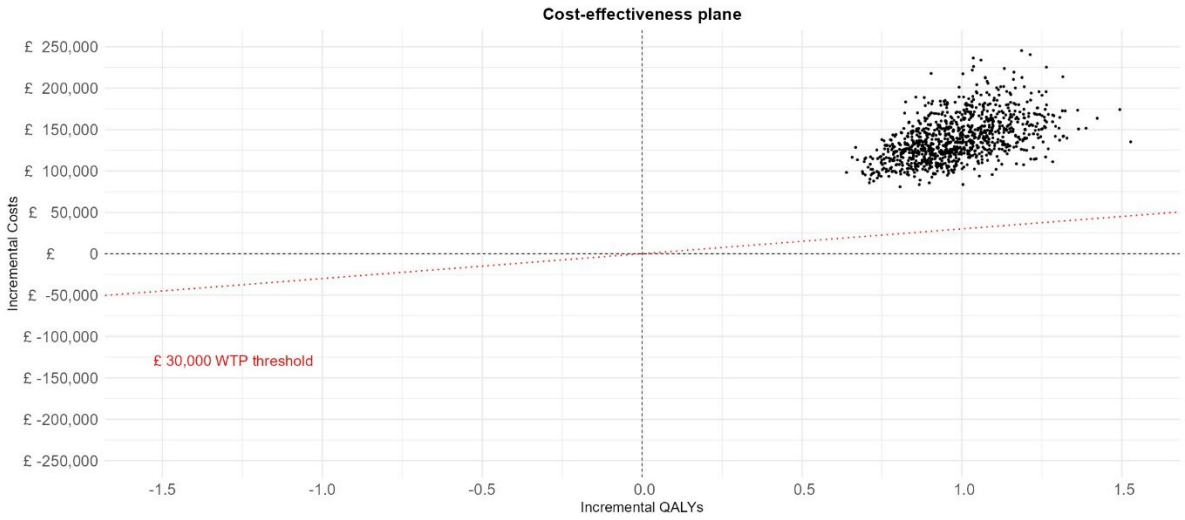
Table 17: Cost-effectiveness results from probabilistic sensitivity analysis

PSA		Costs	QALYs	ICER
Comparators	Hypothetical Intervention & BSC	£304,234	1.81	
	BSC	£165,846	0.83	141,092

* Abbreviations: best supportive care (BSC), quality-adjusted life year (QALY), and incremental cost-effectiveness ratio (ICER).

The cost-effectiveness plane in Figure 9 showed that the PSA simulations were above the cost-effectiveness threshold with the base-case settings in the Shiny app.

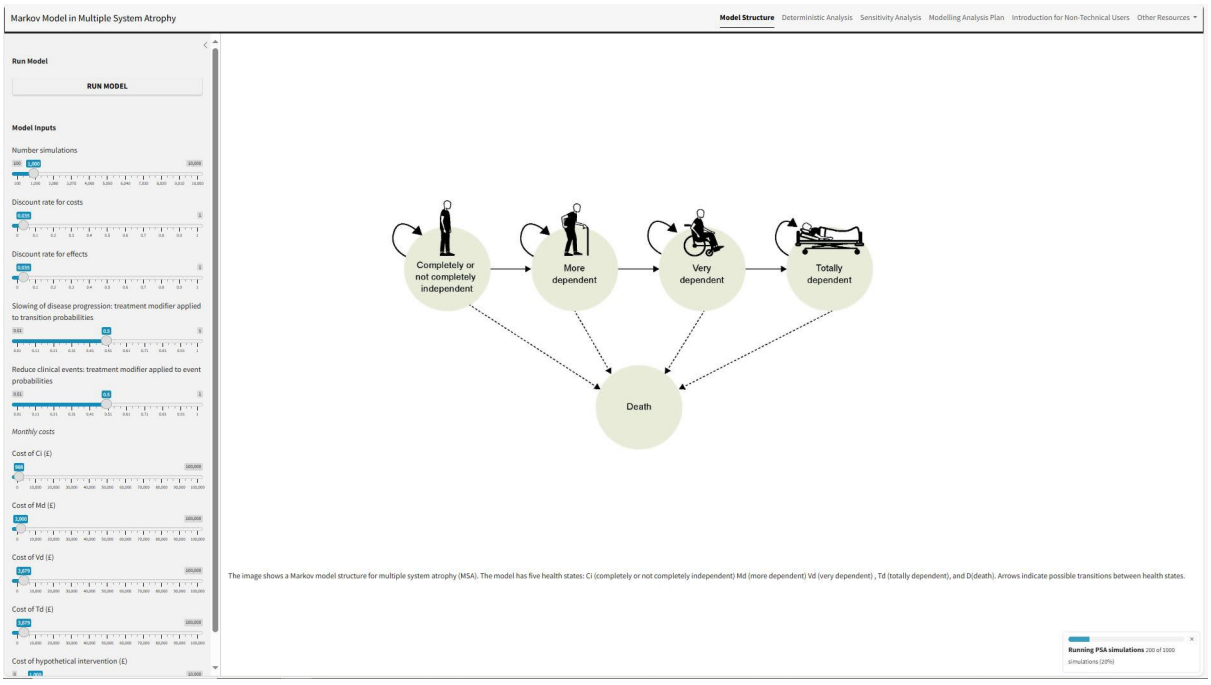
Figure 9: Cost-effectiveness plane for PSA



3.3.3 *Shiny app*

A Shiny app was developed and can be accessed by clicking this link: https://sheffield-university.shinyapps.io/Markov-Model-MSA/. Figure 10 shows an image of the user interface.

Figure 10: Shiny app user interface



3.4 Discussion

3.4.1 *How can researchers better capture the health economic value of multiple system atrophy?*

The literature searches illustrated a lack of cost-effectiveness analyses, and a scarcity of cost-of-illness or burden-of-disease studies which is not unusual for rare diseases such as MSA (Grand *et al.*, 2024b; Pearson *et*

al., 2018). Surprisingly, more research was available on quality-of-life where UMSARS was most often reported (Cao *et al.*, 2018; Foubert-Samier *et al.*, 2020; Golden, McCreary and Vernino, 2022; Jecmenica-Lukic *et al.*, 2021; Jecmenica-Lukic *et al.*, 2018; Kollensperger *et al.*, 2007; Low *et al.*, 2015; May *et al.*, 2007; Wenning *et al.*, 2013; Watanabe *et al.*, 2002; Saulnier *et al.*, 2024). Studies included SF-36 which allows for mapping to EQ-5D, or they directly reported EQ-5D index values for use in cost-effectiveness analyses (Kollensperger *et al.*, 2007; Jecmenica-Lukic *et al.*, 2018; Low *et al.*, 2015; May *et al.*, 2007; Schrag *et al.*, 2006; Schrag *et al.*, 2010; Xiao *et al.*, 2022; Rowen, Brazier and Roberts, 2009). However, EQ-5D was reported for the whole disease, not stratified by scales or health states. Multiple studies reported on natural history of disease, e.g., survival (Cao *et al.*, 2018; Foubert-Samier *et al.*, 2020; Low *et al.*, 2015; Watanabe *et al.*, 2002; Wenning *et al.*, 2013). A variety of disease characteristics were reported such as stridor, falls, orthostatic hypotension, depression, apathy, delusions, and urinary incontinence (Cao *et al.*, 2018; Foubert-Samier *et al.*, 2020; Jecmenica-Lukic *et al.*, 2021; Jecmenica-Lukic *et al.*, 2018; Kollensperger *et al.*, 2007; Saulnier *et al.*, 2024). Few studies reported time-to-event or probability of event analyses, which ought to have implications for early modelling, and they were falls per day, unintelligible speech, and time to aid-required walking, wheelchair confinement, and bedridden state (Watanabe *et al.*, 2002; Wenning *et al.*, 2013). There seems to be little consensus on the inclusion of endpoints, which might be due to the multifaceted clinical representation of people living with MSA.

Establishing decision-analytic model structures for MSA could help guide the inclusion of health-economic endpoints, for example, in natural history of disease studies or clinical trials. This is important for cost and utility data, where recent data and sufficient stratification were missing for clinical events or milestones. Furthermore, reporting resource use rather than country-specific costs is recommendable for transferability. Lastly, research on the caregiver burden, e.g., associated indirect costs, was limited.

3.4.2 Can data gaps be minimised, or cost-effectiveness modelling facilitated by exploration of rare disease analogues?

Cost-effectiveness analyses for disease analogues were plentiful (Ayati *et al.*, 2021a; Ayati *et al.*, 2021b; Bolanos-Diaz, Sanabria-Montanez and Custodio-Capunay, 2019; Furneri *et al.*, 2019; Kremer *et al.*, 2020; Lazzaro *et al.*, 2022; Martins *et al.*, 2023; Michels *et al.*, 2019; Montgomery *et al.*, 2022; Nakhaipour *et al.*, 2020; Navarro and Betancur, 2023; Pinheiro *et al.*, 2020; Poveda *et al.*, 2020; Schur *et al.*, 2021; Spelman *et al.*, 2024; Spelman *et al.*, 2022; Stanisic *et al.*, 2019; Taheri, Sahraian and Yousefi, 2019; Visser *et al.*, 2022; Walter *et al.*, 2019; Xu *et al.*, 2019; Auguste *et al.*, 2020; Baharnoori *et al.*, 2022; Cortesi *et al.*, 2022; Daroudi *et al.*, 2023; Ginestal *et al.*, 2023; Giovannoni *et al.*, 2019; Hua *et al.*, 2019; Lasalvia *et al.*, 2020; Rezaee *et al.*, 2019; Acosta *et al.*, 2021; Hunter *et al.*, 2021; Oppe *et al.*, 2021; Zhao *et al.*, 2022; Palmer *et al.*, 2024; Dimitrova *et al.*, 2023; Mantovani *et al.*, 2020; Michel *et al.*, 2022; Renfrew *et al.*, 2019; Rosato *et al.*, 2021; Huygens and Versteegh, 2021; Versteegh *et al.*, 2022; Chaudhuri *et al.*, 2022; Kalabina *et al.*, 2019; Pahwa *et al.*, 2023; Gajdos *et al.*,

2021; Thakore *et al.*, 2020; Pinto *et al.*, 2023). The 5-year search limit was a minor limitation as a large proportion of cost-effectiveness analyses in advanced or late stage Parkinson had an earlier publication peak (Dams, Zapp and König, 2023). This was a rather contradictory result in comparison with MS and it seemed to be caused by a different publication pattern, which indicated a continuum of publications rather than a peak as for Parkinson (Navarro, Ordóñez-Callamand and Alzate, 2020; Gallehzan *et al.*, 2024).

Health-economic parameters from cost-effectiveness analyses of disease analogues were often not transferable from papers that stratified by model health states. For example, for MS costs and utilities were often stratified by an EDSS, from a variety of studies (Acosta *et al.*, 2021; Auguste *et al.*, 2020; Ayati *et al.*, 2021a; Ayati *et al.*, 2021b; Baharnoori *et al.*, 2022; Cortesi *et al.*, 2022; Furneri *et al.*, 2019; Ginestal *et al.*, 2023; Giovannoni *et al.*, 2019; Hua *et al.*, 2019; Kremer *et al.*, 2020; Lazzaro *et al.*, 2022; Martins *et al.*, 2023; Michels *et al.*, 2019; Montgomery *et al.*, 2022; Nakhaipour *et al.*, 2020; Navarro and Betancur, 2023; Pinheiro *et al.*, 2020; Poveda *et al.*, 2020; Rezaee *et al.*, 2019; Schur *et al.*, 2021; Spelman *et al.*, 2024; Spelman *et al.*, 2022; Stanisic *et al.*, 2019; Taheri, Sahraian and Yousefi, 2019; Walter *et al.*, 2019; Xu *et al.*, 2019). Similarly, utilities were stratified by Hoehn and Yahr and on- and off-states for Parkinson's disease (Kalabina *et al.*, 2019; Chaudhuri *et al.*, 2022). These insights could suggest that data were specifically generated for cost-effectiveness modelling purposes or that patient reported outcomes were captured and once collected, were used to inform health-economic models. Thus, further illustrating the benefit of early conceptual modelling to provide timely guidance for data generation. Useful analyses were those that referred to, or provided thorough micro-costing approaches, or reported relevant clinical events. For example, cost or disutility of urinary disorders, falling, depression, cardiovascular disorders or walking aids, which are relevant events for MSA (Hunter *et al.*, 2021; Walter *et al.*, 2019; Wenning *et al.*, 2013; Watanabe *et al.*, 2002; Cao *et al.*, 2018; Wenning *et al.*, 1994; Jecmenica-Lukic *et al.*, 2021; Foubert-Samier *et al.*, 2020; Low *et al.*, 2015; Kollensperger *et al.*, 2007; Saulnier *et al.*, 2024). These findings further encourage the exploration of sourcing surrogate data from disease analogues, but there is a need for more research (Pearson *et al.*, 2018; Grand *et al.*, 2024b).

However, certain modelling considerations could be carried over such as assessing potential double counting if utility questionnaires were overlapping or if event-driven disutility increments were reported together with health-state utilities (Auguste *et al.*, 2020; Pinheiro *et al.*, 2020). Moreover, there were examples where treatments were differentiated by their probability for various events, which supports the inclusion of a proposed event modifier (rr_trt_mod_e) in conceptual modelling (Poveda *et al.*, 2020; Pinheiro *et al.*, 2020).

The recommendation to investigate disease analogues showed that the strategy was most useful for conceptualising model structures, e.g., the overwhelming use of Markov models based on functional or disability rating scales such as Hoehn and Yahr or EDSS (Grand *et al.*, 2024b). Similarly, for health-economic

parameters, studies tended to be useful for inclusion or comparison if sufficient granularity was reported for costs and utilities, and when assessing applicability of frequent events, e.g., urinary tract disorders or less frequent events such as house or car modifications (Taheri, Sahraian and Yousefi, 2019; Hunter *et al.*, 2021; Wenning *et al.*, 2013; Kollensperger *et al.*, 2007; Saulnier *et al.*, 2024; Low *et al.*, 2015).

By contrast, more targeted searches for missing input parameters (both for rare and prevalent disease analogues) may be more time efficient in situations where model structure, including health-economic parameters, are already known (Grand *et al.*, 2024b; Pearson *et al.*, 2018). However, in all cases, seeking clinical advice is advisable to identify disease analogues and assess data eligibility for cost-effectiveness modelling.

Much work is needed to establish a process for ensuring the validity of disease analogues. We know that a disease analogue was used in a HTA in the UK where quality-of-life data for wheelchair confinement were sourced from MS (prevalent) to inform Duchenne's disease (rare) (Pearson *et al.*, 2018). Many questions remain: how does one define a disease analogue? What is a good disease analogue? How does one ensure validity? What is broader acceptance by payers?

As mentioned, ALS, MS, and loss of ambulation through, e.g., late-stage Parkinson disease or spinal cord injury were chosen because of overlapping symptoms with MSA, e.g., with significant impact on functioning, quality of life and survival, and with higher prevalence because of known evidence scarcities for rare diseases (Kanters *et al.*, 2013; Grand *et al.*, 2024b). Spinal cord injury was later excluded in the publication to sharpen the focus exclusively on diseases of the nervous system. The disease was kept in this document for readers to understand how the process was refined over time and for discussion.

It was the clinical experts that proposed the initial disease analogues. However, as a health economist, I educated them on the issue that is data scarcities in rare diseases and further explained the concept of disease analogues by providing the previous HTA example from the UK.

One may rightly ask, what are the limitations of the chosen disease analogues for MSA? The clinical representation is very similar between MSA and Parkinsons disease, to a degree where misdiagnosis often occur (Krismer *et al.*, 2024). Recently, a case report was published where ALS was misdiagnosed as MSA (Tang, Yao and Wang, 2025). However, disease-modifying therapies are missing for MSA and Parkinson, and it is believed that biomarkers, early diagnostics to ensure sufficient stratification by clinical and genetic features may be the key to overcoming current hurdles (Liu, Wang and Shang, 2024; Zhang *et al.*, 2025). By contrast, treatment for MS has evolved with the introduction of 20 disease-modifying therapies (Reidy *et al.*, 2025). It could conceivably be hypothesised that access to disease-modifying treatments may increase quality of life for patients. Thus, caution must be exercised when considering surrogate data across two diseases with vastly different treatment opportunities despite the clinical similarities between diseases. It

limits the applicability of MS as a disease analogue for MSA. A criterion to implement in future analyses could be that similar treatment opportunities need to exist across diseases for viability.

3.4.3 Was early health technology assessment feasible for multiple system atrophy with currently available evidence?

Multiple definitions exist for early HTA, but an amalgamation of them that aligns with this study is: a systematic approach that syntheses data gaps to guide early evidence generation and to demonstrate an intervention's potential value for payers, patients, providers and other healthcare stakeholders at an early stage of development (Ijzerman *et al.*, 2017; Llorian *et al.*, 2023).

3.4.3.1 Model structure

Markov models constructed by functional, or disability rating scales have seen widespread use in disease analogues, which can support the choice of UMSARS Part IV to inform a model structure in MSA (Auguste *et al.*, 2020; Ayati *et al.*, 2021a; Ayati *et al.*, 2021b; Baharnoori *et al.*, 2022; Chaudhuri *et al.*, 2022; Cortesi *et al.*, 2022; Daroudi *et al.*, 2023; Furneri *et al.*, 2019; Ginestal *et al.*, 2023; Giovannoni *et al.*, 2019; Hua *et al.*, 2019; Kalabina *et al.*, 2019; Kremer *et al.*, 2020; Lasalvia *et al.*, 2020; Martins *et al.*, 2023; Michels *et al.*, 2019; Montgomery *et al.*, 2022; Nakhaipour *et al.*, 2020; Navarro and Betancur, 2023; Pinheiro *et al.*, 2020; Poveda *et al.*, 2020; Rezaee *et al.*, 2019; Schur *et al.*, 2021; Spelman *et al.*, 2024; Spelman *et al.*, 2022; Stanisic *et al.*, 2019; Taheri, Sahraian and Yousefi, 2019; Thakore *et al.*, 2020; Visser *et al.*, 2022; Walter *et al.*, 2019; Xu *et al.*, 2019; Pahwa *et al.*, 2023). Hand searches for less prevalent disease analogues (rare diseases) further corroborated these findings, but this was also important because MSA was opportunistically chosen. Thus, it is interesting to explore cost-effectiveness modelling in other rare diseases. For example, Huntington disease where a total functional score was combined with active functional decline by Shoulson and Fahn categories (Guzauskas *et al.*, 2024). In addition, spinocerebellar ataxia type 1 where early model health states were stratified by a scale for the assessment rating of Ataxia (van Prooije *et al.*, 2023). Moreover, it can conceivably be hypothesised that UMSARS Part IV progression will correlate with increased costs and reduced quality-of-life, which is further indicated by published data and mean index scores derived from EMSA IPD in Table 14 (Winter *et al.*, 2011b). What is more, McCrone and colleagues hold the view that their data support the hypothesis that disease modifying interventions can decelerate or halt progression of disease, and hence, reduce the economic burden (McCrone *et al.*, 2011).

However, there is abundant room for further progress in determining the most appropriate model structure for MSA. How and which non-functional events such as depression should a model account for, or short-term events like hospitalisations due to falling, if they should be accounted for at all? In this study, short-term events were included despite a risk of double-counting. For example, risk of double-counting quality of life as the disutilities were subtracted from the health-state utilities, which facilitated a straightforward

inclusion of an event modifier. However, a multiplicative method is normally recommended when estimating health-state utilities with co-occurring comorbidities (Ara and Brazier, 2017).

Clinical experts highlighted that patients could deteriorate within a day, e.g., due to urinary tract infection, and at times recover quickly if antibiotic treatment was efficient. So, these events were included to highlight sudden drops in quality of life that were not captured if an event happened in between the six-monthly EQ-5D measurements. However, these events were not assigned consequences in terms of costs, because aggregated cost data were used across health states which may account for these events, but this was not clear from the publication (Winter *et al.*, 2011b).

As mentioned, there is little data to suggest which events to include in models. Can or should stopping rules be applied, e.g., the ethical aspects of continued treatment in health states perceived worse than death (Clarke, Ellis and Brownrigg, 2021)? What is the role of biomarkers in responder / non-responder analyses? The clinical representation of MSA is highly dynamic and multifaceted, which makes it difficult to construct a parsimonious model, e.g., balancing simplicity with the underlying clinical reality, and more research is needed here.

What is more, aid-required walking, wheelchair confinement, and bedridden states have previously been reported as important motor-function or clinical milestones (Watanabe *et al.*, 2002). Also, a noteworthy relationship exists between UMSARS Part IV and these clinical milestones. For example, the fifth UMSARS Part IV item states “totally dependent and helpless. Bedridden.” By contrast, aid-required walking and wheelchair confinement are non-explicitly mentioned, but as previously highlighted, they have been illustrated alongside each other (Saulnier *et al.*, 2024). In theory, it should be possible to triangulate the results by partitioned survival modelling (time-to-event analyses) whereby patients move through health states based on parametric equations as opposed to transition probabilities for Markov models (York Health Economics Consortium, 2016e). For example, by digitising available Kaplan-Meier curves and reconstructing patient-level data on clinical progression milestones, but this would require the health state Ci to represent the state “able to walk”, despite being two separate items in UMSARS Part IV (Watanabe *et al.*, 2002). In this analysis UMSARS IV items I (completely independent) and II (not completely independent) were grouped as Ci because cost and utility data indicated similar values, but this should be further investigated once more data become available as there might be an incremental difference in costs and effects that is not captured when these items are collapsed into one health state.

3.4.3.2 Early cost-effectiveness modelling

The early cost-effectiveness modelling only had just enough data to populate all health-economic parameters. For example, the model was bound by available cost data which were aggregated for UMSARS Part IV item 1 and 2, and 4 and 5. The Shiny app allows for variation in costs, but one could expand the

model structure into all five UMSARS Part IV items and generate state specific costs. The natural history of disease or EMSA data, which were captured in 2003 or 2004 with two years of follow-up, were crucial in estimating transition probabilities and utilities for health states. However, I encourage more research on estimation of transition probabilities in datasets with few subjects and short follow-up, as this is often the case for rare diseases. For example, the analyses showed sensitivity to the choice of methods, such as non-parametric count methods versus multistate modelling, which are available from Table 4 and Table 13, respectively.

The utility values in the earlier health states were surprisingly low. For example, collapsing item 1 (12 data points) and item 2 (75 data points) yielded a utility value for UMSARS Part IV item 1 + 2 of 0.58 (Table 14). The utility values from literature varied when the disease as a whole was considered, e.g., from 0.2 to 0.56 (Schrag *et al.*, 2006; Schrag *et al.*, 2010; Xiao *et al.*, 2022). What could explain these discrepancies? One could hypothesise that population characteristics are likely to influence utility values, e.g., whether the population is captured at symptom onset or probable diagnosis, because both patient groups are likely to meet criteria for UMSARS Part IV item 1 or 2, but these two populations may present with different utility values. For example, quality of life may be higher for those captured from symptom onset, because they may be earlier in their disease, e.g., not yet diagnosed. Moreover, if average values for the whole disease are considered, e.g., from 0.2 to 0.56, it is not unlikely that those with milder disease present with a utility value of 0.58. Naturally, because we are comparing the full severity spectrum (mild to severe) with only milder patients (UMSARS Part IV item 1 + 2). It would be interesting to strengthen this analysis by considering a Bayesian approach where utilities from advanced Parkinsons disease were used as priors to increase the evidence base.

Modelling results are highly uncertain at this early stage. However, model conceptualisation guidelines emphasise that models are a simplification of reality and that model specification is an iterative process, and future iterations may further add to a model as the understanding of the decision problem deepens (Kunst *et al.*, 2023a; Roberts *et al.*, 2012). Model development should be independent from considerations on availability of data to inform health-economic parameters, which is why it is important not to solely appraise early economic evaluations of rare diseases based on data availability (Roberts *et al.*, 2012; Caro *et al.*, 2012). I strongly oppose this practice because it is an inherent challenge and will hinder any effort to advance research on economic evaluations of rare diseases (Grand *et al.*, 2024b). The use of a treatment modifier in early HTA could potentially be necessary to determine the efficacy required in clinical studies or for estimation of cost-effectiveness with current evidence or efficacy projections, and subsequently inform go/no-go decisions in early pharmaceutical pipeline development, e.g., against country specific HTA frameworks or cost-effectiveness thresholds (Llorian *et al.*, 2023). It is useful when validating cost-effectiveness analyses during very early HTA to distinguish according to guidelines on multi-application

models: validation in the general sense (the MSA model itself) versus for a specific application (the effect of an intervention in MSA) (Eddy *et al.*, 2012). The focus here was clearly on the model development, which from an early HTA perspective has value, whereas nothing was currently available to inform an intervention, which adds to the uncertainty of the modelling results. It has been stated in guidelines that sensitivity analyses are warranted for model parameters even if data do not exist e.g. for an intervention, to assess a potential influence on results (Roberts *et al.*, 2012). To summarise, my viewpoint on cost-effectiveness modelling in early HTA is that a treatment modifier has merit, and it is useful for determining efficacy levels required in clinical studies for go/no-go decisions, timely data generation and for establishing model structures, despite uncertainty in results.

The PSA illustrated the uncertainty in data, while the scenario analyses illustrated the importance of generating state specific cost data which had high influence on the results. However, cost-effectiveness was achievable depending on user-specified input parameters. Moreover, by eliminating this uncertainty, one could potentially also start to see deterministic results that were cost-effective. To conclude, I propose caution when interpreting results, but conceptual modelling has been useful in determining potential model structures and data gaps. It is advantageous to contextualise early conceptual modelling of MSA in an iterative modelling framework, because here I present only a humble beginning (first iteration) (Kunst *et al.*, 2023a). More research is needed on the most appropriate model structure and data gaps for MSA.

3.4.3.3 Future work for early health technology assessment

Practical frameworks vary for early HTA, but early-stage health-economic modelling is commonly included and has previously made use of hypothetical scenarios (Ijzerman *et al.*, 2017; Wang, Rattanavipapong and Teerawattananon, 2021). Although, the challenges for personalised medicines have been described, there is lack of focus on rare diseases (Ijzerman *et al.*, 2017; Llorian *et al.*, 2023). Interestingly, Wang and colleagues highlight the lack of a unified framework for early HTA as a limitation for wider use by manufacturers and innovators, which may have been partly induced by the inherent methodological affluency in an earlier definition as compared to newer. For example, "*all methods used to…"* (Ijzerman *et al.*, 2017), is perhaps the reason for the suggested change to *"a systematic approach to"* (Llorian *et al.*, 2023), because a systematic approach may result in a framework that is easier to unify. By contrast, a definition that promotes the use of all methods may be more difficult to encompass in a unified framework.

3.4.4 Is it worthwhile developing user-friendly interfaces for early health technology assessment in the context of rare diseases?

Early HTA includes a broad range of stakeholders, which aligns well with my intentions for a Shiny app. Prof. Whitty wrote that "authors of modelling papers can do things to make their work substantially more useful for policy. The best is to provide an interactive interface, where if the policymaker does not agree with the

starting assumptions of the model, they can change them" (Whitty, 2015, p. 4). In my opinion, it is also one of several ways to promote transparency, validation, spark debate, and hopefully encourage others to progress a field of research such as a rare disease like MSA (Eddy *et al.*, 2012). Perhaps Shiny apps, and R in particular, are becoming increasingly popular for these reasons, coupled with computational advantages and script-based coding resulting in less copy-paste errors as compared to spreadsheet-based modelling (Xin *et al.*, 2022; Alarid-Escudero *et al.*, 2019; Hart *et al.*, 2020). Importantly, the independent review of R code for the Shiny app conducted by an experienced modeller found no errors or bugs, and deterministic results were reproduced in Excel. Nonetheless, tutorial papers on cost-effectiveness modelling in R have recently surged (Hart *et al.*, 2020; Xin *et al.*, 2022; Naylor *et al.*, 2023; Krijkamp *et al.*, 2018; Green *et al.*, 2023; Ben *et al.*, 2023; Smith, Mohammed and Schneider, 2023). Interestingly, Hart and colleagues mention early cost-effectiveness modelling, PSA, changes to underlying data, and long run time if not conducted in R, as criteria that may help decide when R will lead to substantive improvements in efficiency and transparency (Hart *et al.*, 2020). I believe that this case-study fulfilled most if not all criteria, and thus, early cost-effectiveness modelling in R outweighed the burden of programming and reviewing on a less familiar platform. To summarise, development of user-friendly interfaces is certainly possible with help from programming tutorials, even with modest R experience, and can be recommended for early conceptual cost-effectiveness modelling in rare diseases, e.g., MSA (Smith and Schneider, 2020; Alarid-Escudero *et al.*, 2023; Krijkamp *et al.*, 2020; Green *et al.*, 2023).

3.4.5 Limitations

A major limitation of this analysis was that cost data were only reported in two cost-of-illness studies by McCrone and Winter et al., both dating back to 2011 (Winter *et al.*, 2011b; McCrone *et al.*, 2011). Winter and colleagues reported costs stratified by UMSARS part IV, but item 1 and 2, together with 4 and 5 were aggregated. A limitation was that patients were recruited in 2005, which meant that the data predates the publication date by six years. The perspective was German which required an assumption of generalisability to a UK setting (Winter *et al.*, 2011b). The ISPOR task force recommends that the consumer price index is used for the local currency before conversion with purchasing power parity or exchange rates (Shi *et al.*, 2010). Germany and UK are both high-income countries, but it would have been more appropriate to follow their guidance, but the data used were UK consumer price inflation data and official exchange rates, but one could have used purchasing power parity. It is a limitation that the approach did not strictly adhere to published guidance on conversion of costs. In addition, a reduction in costs were not modelled despite a decrease in clinical events for the hypothetical intervention arm to avoid double-counting in relation to the use of aggregated cost data. The inclusion of clinical events was based on several factors, e.g., previous inclusion in cost-effectiveness analyses for disease analogues. However, there was no formal study, guidelines or established consensus that outlines the significance of disutility decrements and if events are

sufficiently short to avoid double-counting, e.g., in relation to the use of six-monthly EQ-5D questionnaires to inform health state utilities (Acosta *et al.*, 2021; Ayati *et al.*, 2021b; Hunter *et al.*, 2021; Martins *et al.*, 2023). The disutility values for urinary or bladder, and bowel disorders were estimated from EQ-5D catalogues (Hvidberg and Alava, 2023; Sullivan *et al.*, 2011). On the contrary, hypotension disorders, e.g., orthostatic hypotension, was estimated from an MSA population (Xiao *et al.*, 2022). Lastly, the lack of disease-modifying interventions, led to the use of treatment and event modifiers to assess the potential influence on cost-effectiveness results.

3.5 Conclusions

The literature searches for disease-analogues were useful for providing information on model structures, which showed a widespread use of functional rating scales to inform Markov models. Although, the modelling demonstration incorporated little surrogate data, it could potentially have been important had the natural history of disease data not been available to inform transition probabilities and utilities. It was feasible to conduct a first hypothetical cost-effectiveness analysis for MSA even though health-economic parameters were scarce. Early conceptual modelling is recommendable as it provides import insights on data gaps and evidence-generation needs. Moreover, it is recommendable to conduct iterative modelling, and the proposed model will almost certainly develop over time as new data emerge, such as expansion of health states, stopping rules, responder / non-responders. Shiny apps or user-friendly interfaces should be considered when conducting early conceptual modelling and they are attainable with published tutorials or R courses, but analysts should evaluate published criteria on potential pros and cons if they have less modelling experience in R.

Chapter 4

Conceptual modelling of functional milestones in multiple system atrophy: comparing Markov and partitioned survival models using digitised time-to-event data

This chapter contains my 3rd study, which explored conceptual modelling of functional milestones to inform model health states in MSA using digitised time-to-event data. It is important to highlight that this a model demonstration of using functional milestones in a hypothetical cost-effectiveness analysis. The conceptual modelling is undertaken using both Markov and partitioned survival models.

The latest version of the manuscript is included in Section IV.III as part of III Appendix: , which is intended for submission to PharmacoEconomics Open. The journal editor has already assessed the manuscript for suitability. The remaining step before submission is internal approval at Lundbeck A/S, but the manuscript may change based on reviewer comments.

4.1 Introduction

The introduction outlines the rationale for conducting conceptual modelling with functional milestones in MSA while exploring implications of using different model structures.

4.1.1 Rationale

The inspiration for this study arose from the previous model demonstration in Chapter 3: cost-effectiveness modelling of MSA for early HTA (Grand *et al.*, 2025). The previous study briefly alluded to the potential use of functional (clinical) milestones as opposed to UMSARS Part IV for cost-effectiveness modelling, which in my opinion warranted further research. It was stated in our publication that *"In theory, it should be possible to triangulate our results by partitioned survival modelling (time-to-event analyses) whereby patients move through health states based on parametric equations as opposed to transition probabilities for Markov models" (Grand et al., 2025, p. 11)*.

Chapter 3 illustrated overwhelming evidence and precedence for using functional or disability scales in disease analogues to inform decision-analytic model structures (Auguste *et al.*, 2020; Ayati *et al.*, 2021a; Ayati *et al.*, 2021b; Baharnoori *et al.*, 2022; Chaudhuri *et al.*, 2022; Cortesi *et al.*, 2022; Daroudi *et al.*, 2023; Furneri *et al.*, 2019; Ginestal *et al.*, 2023; Giovannoni *et al.*, 2019; Hua *et al.*, 2019; Kalabina *et al.*, 2019; Kremer *et al.*, 2020; Lasalvia *et al.*, 2020; Martins *et al.*, 2023; Michels *et al.*, 2019; Montgomery *et al.*, 2022; Nakhaipour *et al.*, 2020; Navarro and Betancur, 2023; Pinheiro *et al.*, 2020; Poveda *et al.*, 2020; Rezaee *et al.*, 2019; Schur *et al.*, 2021; Spelman *et al.*, 2024; Spelman *et al.*, 2022; Stanisic *et al.*, 2019; Taheri, Sahraian and Yousefi, 2019; Thakore *et al.*, 2020; Visser *et al.*, 2022; Walter *et al.*, 2019; Xu *et al.*, 2019; Pahwa *et al.*, 2023). It was hypothesised that increased costs and reduced quality-of-life data were correlated with disease

progression on the disease-specific disability scale UMSARS Part IV, which was further supported by clinical experts. It meant that UMSARS Part IV was used to inform a Markov model structure for early HTA in Chapter 3.

What is crucial for the academic merit of this study is that the relationship between UMSARS Part IV and functional milestones such as able to walk, aid-required walking, wheelchair confinement and bedridden were emphasised as a matter that deserves attention in relation to explorative cost-effectiveness modelling (Wenning *et al.*, 2004). It is important for the study objectives that these health states were previously illustrated alongside each other by Saulnier and colleagues (Saulnier *et al.*, 2024).

Exploration of functional milestones were highlighted as a simple way to triangulate previous modelling results by partitioned survival and Markov modelling, and to further evaluate the most appropriate model health states for MSA (York Health Economics Consortium, 2016e). For example, by digitising available Kaplan-Meier curves and reconstructing patient-level data on clinical progression milestones, which allowed for the use of data with up to 17 years of follow-up (Watanabe *et al.*, 2002). The comparative modelling required the first two items: completely independent and not completely independent to represent the state “able to walk”, despite being two separate items in UMSARS Part IV. It is worth noting that these two states were also collapsed in our previous study in Chapter 3 because they represented similar utilities and due to major data limitations, e.g., that only aggregated costs data were available for these two items.

4.1.2 Objectives

The study aimed to address the following research objectives:

- Can patient-level data be reliably digitised and reconstructed from existing published sources?
- Can alternative model structures based on functional milestones be developed to compare the advantages and disadvantages of different modelling strategies, such as Markov and partitioned survival models?
- Is it worthwhile to explore EVPI in the context of rare disease?

4.2 Methods

This section outlines the methods, but generally, I attempt to apply the same assumptions and health-economic parameters as in the previous analysis in Chapter 3 where possible to allow for a comparison of model types. Noteworthy is that modelling strategies, health states and data to inform disease progression were different than in the previous cost-effectiveness analysis in MSA (Grand *et al.*, 2025).

An R project includes code for data analysis and results, which are available from Github: https://github.com/TobiasGrand/PSM-and-Markov-for-milestones-in-MSA. The modelling report is a

Markdown report (html file) with a dropdown button named “show”, which opens relevant code chunks when clicked (Allaire *et al.*, 2023; R Core Team, 2021). The markdown report is available from: https://tobiasgrand.github.io/Report_PSM_Markov_MSA/.

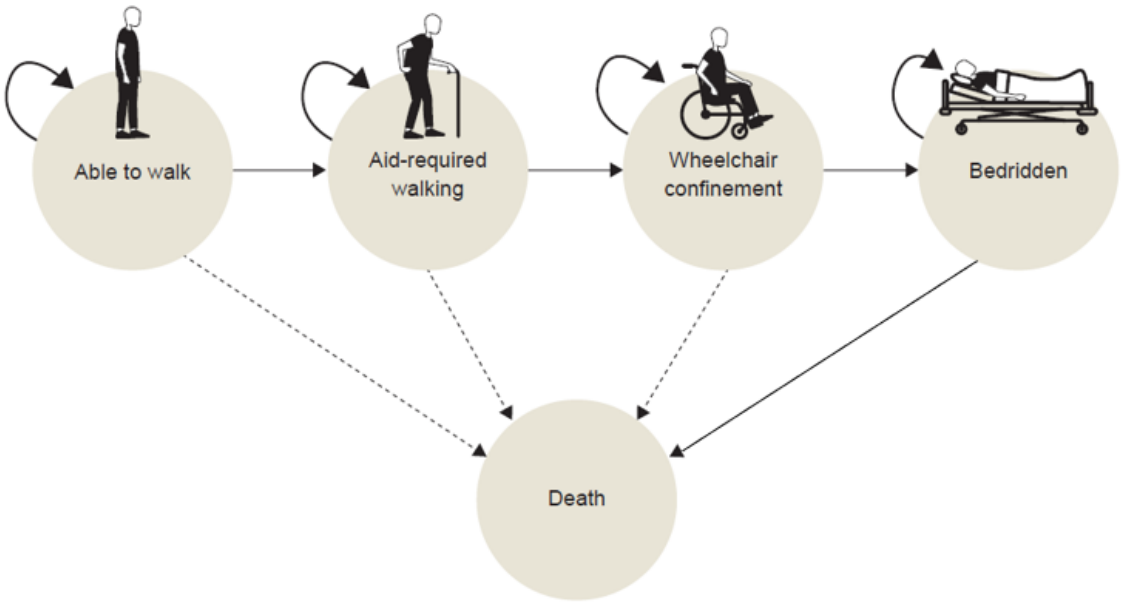
4.2.1 Study population and perspective

The perspective of the analysis was a UK NHS perspective with a cost-effectiveness threshold of £30,000 per QALY (Bouvy, 2024). Discount rates were 3.5% for costs and effects (National Institute for Health and Care Excellence, 2022e). A MAP was not developed for this study because the analysis was primarily pragmatic or conceptual in nature. The diagnosis criteria for the modelled population were the Gilman criteria (Gilman *et al.*, 2008; Gilman *et al.*, 1999). Discrete-time points from Kaplan-Meier curves were reported annually, which guided the decision to use an annual cycle length in the models (Watanabe *et al.*, 2002). The time horizon was 32 years, because few subjects were still at risk of death (meaning they remained alive) as illustrated by published Kaplan-Meier curves at 12 and 25 years, depending on whether their duration of survival was estimated from the time point of probable diagnosis or symptom onset, respectively (Watanabe *et al.*, 2002; Wenning *et al.*, 2013). In addition, a modelled treatment modifier in the study resulted in slowing of disease progression, which meant that a longer time horizon was needed to capture downstream consequences, e.g., effects and costs from the intervention arm. For example, the minimum time horizon to achieve this was 32 years.

4.2.2 Model structure and development

The two models were based on either a partitioned survival or Markov modelling approach. They shared the functional milestones from Figure 11, which had five health states informed by functional milestones: able to walk, aid-required walking, wheelchair confinement, bedridden and death. The progression was sequential through model health states and mortality was only allowed from the bedridden state, but I expand the explanation of the individual model structures in Sections 4.2.5.1 and 4.2.5.2. The comparators were BSC versus a hypothetical drug given alongside BSC because disease modifying treatments are currently missing.

Figure 11: Model structure with functional milestones in multiple system atrophy



* Note that the partitioned survival model is illustrated in Figure 11. Please refer to Figure 6 for the Markov model structure and note that the functional milestones are used instead of UMSARS Part IV. The sequential transitions are illustrated by the unbroken lines. The dotted lines represent clinically plausible transitions that could not be estimated from the digitised published Kaplan-Meier curves. See Sections 4.2.5.1 and 4.2.5.2 for additional details on model attributes.

4.2.3 Estimated health-economic parameters

Health-economic parameters were categorised as either estimated or other health-economic parameters. The estimated health-economic parameters were those that required significant resources and efforts to derive. For example, it was necessary to digitise Kaplan-Meier curves and reconstruct IPD data to estimate transition probabilities for the Markov model and parametric survival models for the partitioned survival model. By contrast, other health-economic parameters were those that could be sourced from published papers, e.g., costs, utilities, and disutilities.

4.2.3.1 Digitising Kaplan-Meier curves for functional milestones

Kaplan-Meier curves for functional milestones were digitised from a natural history of disease study in the absence of access to original IPD (Watanabe *et al.*, 2002). This study was to my knowledge the only published study that reported time to all five functional milestones as Kaplan-Meier curves in MSA (Watanabe *et al.*, 2002). The study used Gilman 1999 diagnosis criteria, and had up to 17 years of follow-up data (Gilman *et al.*, 1999).

The R package IPDfromKM, which is informed by an iterative Kaplan-Meier estimation algorithm developed by Guyot and colleagues, was used to reconstruct pseudo IPD. As only the total number at risk was available

in the source paper, a strong assumption of no censoring was required when reconstructing the pseudo IPD due to the lack of information (Guyot *et al.*, 2012; Lui, Zhou and Lee, 2021). It means that uncertainty in reconstructed hazard rates was underestimated (Guyot *et al.*, 2012).

The event curves were reported in four different figures in the source paper. Data for time to aid-required walking/wheelchair confinement/bedridden/death were reconstructed independently. The reconstructed data were then visually compared for accuracy and revised to improve accuracy if required, e.g., to check for manual clicking errors, against the original figure. See the Markdown report (Section 1.1.3) for illustrations of the reconstructed Kaplan-Meier curves: https://tobiasgrand.github.io/Report_PSM_Markov_MSA/.

4.2.3.2 Efficacy modification for the hypothetical drug

In addition, to constructing a dataset for a hypothetical drug & BSC, an efficacy modifier was implemented. It meant that the time to an event (milestones) was divided by an arbitrary hazard rate of 0.67 to reflect an intervention that delays disease progression by approximately 50%. Efficacy modifications have been used and discussed elsewhere for early modelling in the absence of a licensed disease-modifying treatment (Grand *et al.*, 2025).

4.2.3.3 Reconstructed individual-patient level data

The process for reconstructing IPD had multiple steps. First, subjects (n = 230) were generated using their sequence number (1 to 230) as subject id's, for each of the comparator arms, which resulted in 460 (1-230; 231-460) reconstructed patients assigned to the events. The subjects were assigned their functional milestones (events) using each of the four event datasets for the two interventions. The process for this required looping through annual cycles whilst assigning events based on the cumulative probabilities from the curves and incremental number of subjects with events. The subjects were sequentially assigned these events, e.g., looping from 1 to 230 while assigning events, if any subjects remained unassigned at later cycles, they were assigned the maximum observed time for their event. It meant that subjects with lower subject id's were consistently assigned the earlier events, which resulted in those subjects having faster progression of events.

4.2.3.4 The assumption of sequential progression

This approach meant that subjects were forced sequentially through the model health states, and this made sure that the distributional properties of the reconstructed Kaplan-Meier curves were preserved, and that the reconstructed IPD could be used for fitting distributions prior to cost-effectiveness modelling. This approach allowed for consistency between the Markov and partitioned survival models, as both models used the same sequential progression assumption in their underlying data.

4.2.4 Other health economic parameters

Other health-economic parameters were those that could be sourced from published papers, e.g., costs, utilities and disutilities.

4.2.4.1 Costs

A previous literature review found two cost-of-illness studies from 2011, which meant that health-economic parameters for MSA were scarce (Winter *et al.*, 2011b; McCrone *et al.*, 2011; Grand *et al.*, 2025). However, a recent study focussed on resource utilisation (frequency and odds ratios), which was not sufficient to inform the modelling (Kaplan, Xie and Kumar, 2025). Importantly, stratification of costs were only reported in one study by UMSARS Part IV, which meant that it was necessary to assume that costs stratified by UMSARS Part IV were approximate of functional milestones, which was an apparent limitation (Wenning *et al.*, 2004). The German cost data were adjusted for inflation, euros converted to British pounds, converted from six-monthly to annual costs, and they were assumed generalisable to a UK setting, which was also necessary in the previous cost-effectiveness analysis (Grand *et al.*, 2025). For transparency, an R script was coded for the sourcing and preparation of cost data. The drug costs were £12,000 per year, which was an arbitrary cost estimate. The drug costs were not applied to the bedridden state, which was regarded as a stopping rule by clinical experts. Table 18 includes the cost data across health states.

Table 18: Cost data for functional milestones

State	Mean costs	Distribution	Shape*	Scale*
Able to walk	£11,853	Gamma	46	257.67
Aid-required walking	£24,002			521.79
Wheelchair confinement	£44,152			959.83
Bedridden	£44,152			959.83

*Note that standard deviations to derive shape and scale were estimated based on total costs and number of subjects. Values are rounded.

4.2.4.2 Utilities and disutilities

Utility data were not available for functional milestones, which meant that a similar assumption of UMSARS Part IV approximating functional milestones was necessary. It allowed for the data to be sourced from our previous paper, which are reported in Table 19 (Grand *et al.*, 2025).

Table 19: Mean index scores for functional milestones

Functional milestones	Mean Index Score*

Able to walk	0.58
Aid-required waling	0.43
Wheelchair confinement	0.14
Bedridden	-0.13

*Point estimate for mean EQ-5D-3L index score (UMSARS Part IV). Note that numeric values are rounded to two decimals for illustration**,** but all decimals were used for modelling.

The disutilities from Table 20 were sourced from literature following a previous approach, which selected the most conservative estimates for inclusion in a cost-effectiveness analysis (Grand *et al.*, 2025).

Table 20: Disutilities and percentages for clinical events

Clinical events	Percentages*	References	Disutilities*	References
Hypotension disorders	56.7 %	(Wenning *et al.*, 2013)	-0.07	(Xiao *et al.*, 2022)
Urinary or bladder disorders	51.1 %		-0.01	(Sullivan *et al.*, 2011)
Bowel disorders	58.2 %		-0.19	(Hvidberg and Alava, 2023)

* Note all numeric values are rounded to two decimals. Percentage of cohort with clinical event at baseline.

The disutilities were multiplied by their annual probabilities, e.g., transforming 56.7% to a probability of 0.567, which produced disutility decrements that were subtracted from health-state utilities, which is an approach previously used elsewhere (Grand *et al.*, 2025). However, for the hypothetical drug & BSC, an efficacy modifier was implemented to account for a reduction in clinical events resulting from an expected treatment effect. The event modifier had a value of 0.5, which was multiplied with event probabilities, and the utility decrement subtracted from health-state utilities. A beta distribution was only applied in cases where alpha and beta for the beta distribution were positive, which were otherwise fixed. The derived health states utilities and parameters are shown in Table 21.

Table 21: Mean utilities after adjustment for disutility increments, and distributions

Comparators	Health states	Mean utility	Distribution	Alpha	Beta
BSC	Able to walk	0.43	beta	1.80	2.43
	Aid-required walking	0.27		0.48	1.28

	Wheelchair confinement	-0.01	fixed		
	Bedridden	-0.28			
Hypothetical drug & BSC	Able to walk	0.50	beta	2.18	2.17
	Aid-required walking	0.35		0.76	1.41
	Wheelchair confinement	0.07	fixed		
	Bedridden	-0.28			

*Note that numeric values are rounded to two decimals. BSC (Best Supportive Care).

4.2.5 Comparison of Markov partitioned survival model approaches

The next two paragraphs outline the model development for a Markov model and partitioned survival model.

4.2.5.1 Markov model

The Markov model was hardcoded but required that transition probabilities were estimated from the reconstructed datasets, which were organised for use in the msm R package that requires that time is a continuous process underlying the data (Jackson, 2011). The dataset fulfilled the minimum criteria that were: subject id, time, and state, but were moreover ordered by intervention. This package facilitates multistate modelling that estimates transition probabilities by maximum likelihood estimation (Jackson, 2011). Multistate modelling has not seen widespread use in health economics yet, but a review found that it was the third most frequent method for estimating transition probabilities in National Institute for Health and Care Excellence (NICE) appraisals after survival analysis and count methods (Srivastava, Latimer and Tappenden, 2021).

In the msm R package, the Q matrix (transition intensity matrix) defines which transitions between states that are allowed to reflect the model structure, for example, the assumption of sequential transitions. The transition intensity values were manually calibrated in R until a reasonable relationship between observed and predicted prevalence by time was visually confirmed. A higher intensity value indicates that a transition is more likely to occur whereas zeros indicate disallowed transitions. It meant that the strict assumption of sequential progression was applied in the coding by specifying the Q matrix accordingly. However, it is important to highlight that unobserved transitions may occur at any time, because of the underlying assumption of continuous time. This assumption was partly violated as reconstructed data were panel data, e.g., observed at discrete time intervals. For example, for longer cycle lengths, transitions across multiple health states may be observed, e.g., from able to walk directly to wheelchair confinement despite not being permitted by the Q matrix. They do, however, represent a sequence of unobserved transitions occurring between discrete time points. For example, an intermediate sequence from able to walk to aid-required walking to wheelchair confinement within an annual cycle. This problem could likely be solved with a shorter cycle length, but an annual cycle length was enforced for consistency with the partitioned survival

model (Jackson, 2011). It was decided to reflect the observed transitions as dotted lines in the model structure (Figure 11) despite representing an unobserved sequence of events. It represents a limitation of the Markov modelling approach, because cost and QALYs were not accounted for in the intermediate (unobserved) states. A Dirichlet distribution, which is a multivariate generalisation of the beta distribution, was assigned to transition probabilities for the PSA, which is a method to assess parameter uncertainty through random sampling from probability distributions (Briggs, Sculpher and Claxton, 2006).

4.2.5.2 Partitioned survival model

The partitioned survival model naturally required distinct methodological techniques that differed from those employed in the Markov model. For example, the digitised time-to-event datasets were fitted individually for each event curve using Weibull, logistic, log-normal, and exponential distributions and their fit evaluated by Akaike's information criterion (AIC) and Bayesian information criterion (BIC).

The partitioned survival modelling was conducted with the hesim R package for which the authors had written a tutorial for a 4-state partitioned survival model, which supported the decision to use this R package (Incerti and Jansen, 2021; Incerti, Jansen and Clements, 2025). Importantly, the subjects with faster progressions, e.g., those that caused unobserved transitions in the multistate modelling (direct transitions through intermediate states), were handled with greater flexibility in the partitioned survival modelling. The fractional time spent in a health state within a model cycle was obtained from the area differences in cumulative event probability curves (Williams *et al.*, 2017). This approach ensured that subjects who progressed through multiple states during one model cycle still accumulated costs and effects from their time in these intermediate states, which contrasts with the Markov model approach.

The deterministic analysis made use of a single simulation with no sampling from distributions for the parametric time-to-event models and health-economic parameters. The probabilistic sensitivity analysis ran 500 simulations to limit the computational burden, and a normal distribution was applied for time-to-event models.

4.2.6 Reporting of analyses

Multiple analyses were reported for both models and their treatment strategies. First, line plots informed by Markov traces illustrated the state occupancy across annual model cycles. Second, ICER tables were produced for comparison of deterministic and probabilistic results. Third, PSA results were visually represented by a cost-effectiveness plane, cost-acceptability curves, and used to calculate EVPI.

Lastly, results were explored using a QALY shortfall calculator while following relevant guidance from the decision support unit (McNamara *et al.*, 2023; Wailoo, 2024). Age at disease onset was 55.4 years and female study participation approx. 43% to align with the digitised data (Watanabe *et al.*, 2002). The QALY weight

should be applied to the incremental QALYs, which results in the following formula for calculating weighted ICERs:

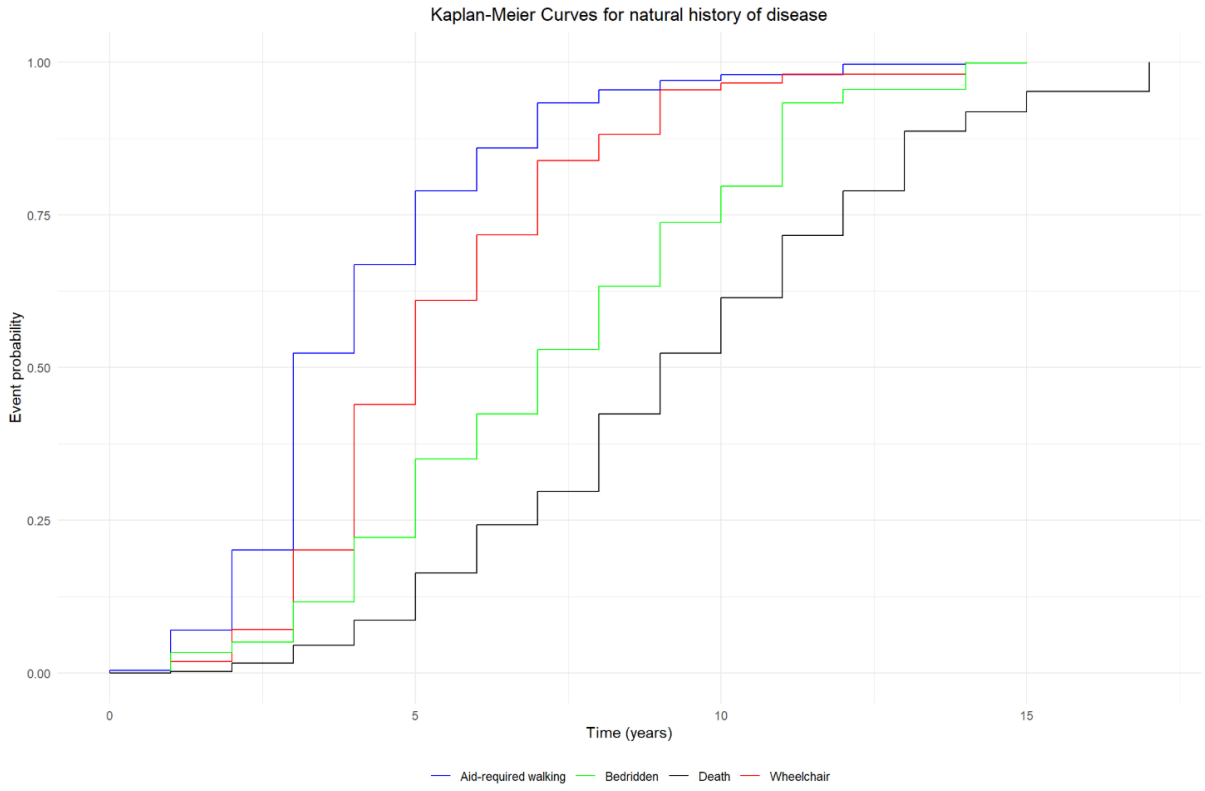
$$ICER_{weighted} = \frac{Costs_{Hypothetical\ Drug\ \&\ BSC} - Costs_{BSC}}{\left(QALYs_{Hypothetical\ Drug\ \&\ BSC} - QALYs_{BSC}\right) x\ QALY\ Weight}$$

4.3 Results

4.3.1 Digitised Kaplan-Meier Curves and fitting of distributions

The Markdown report (https://tobiasgrand.github.io/Report_PSM_Markov_MSA/) contains the full report including code, but all images, e.g., the estimated survival curves are included in Appendix III.I. The digitised Kaplan-Meier curves are shown in Figure 12.

Figure 12: The digitised Kaplan-Meier Curves using IPDfromKM



As mentioned previously, the digitised time-to-event datasets were fitted individually for each event curve using Weibull, logistic, log-normal, and exponential distributions and their fit evaluated by AIC and BIC. Weibull distributions provided the best fit based on these criteria for all event curves, except for aid-required walking across both comparator arms and for wheelchair confinement in the BSC arm where log-normal showed marginally better goodness of fit. However, the log-normal curves had clinically unplausible shapes

(an initial plateau), which meant that Weibull distributions were fitted to each of the events when the event datasets were combined by subject id into parametric models for the two comparators.

4.3.2 Cost-effectiveness results

The cost-effectiveness results in Table 22 show differences between the deterministic and PSAs of £21,000 and £5,000 for the partitioned survival and Markov model, respectively. There was a larger difference when comparing the different model types. For example, across the model types when comparing deterministic analyses, they were £131,380 and £179,362 per QALY, and for probabilistic £151,944 and £183,808 per QALY. It meant that all hypothetical cost-effectiveness estimates were above a cost-effectiveness threshold of £30,000 per QALY.

Table 22: ICER table with cost-effectiveness results across models and analyses

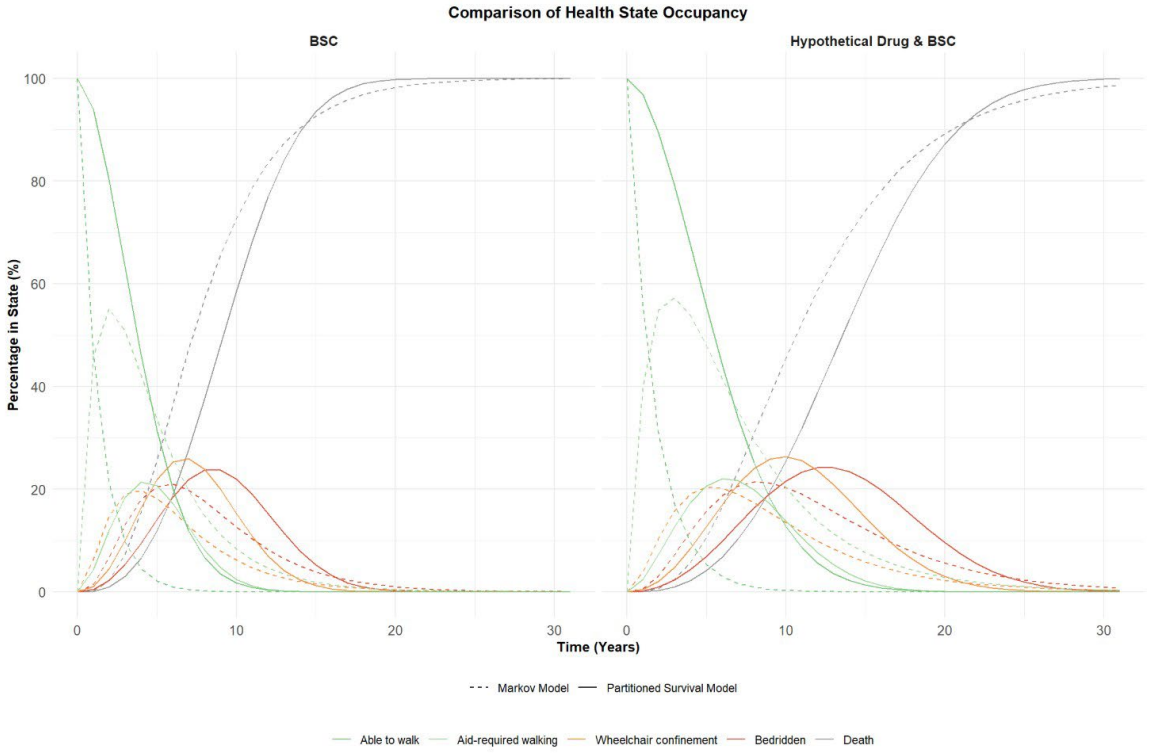
Model	Analysis	Intervention	Costs	QALYs	ICER
Partition Survival model	Deterministic	Hypothetical Drug & BSC	£357,937	2.42	
		BSC	£189,607	1.14	131,380
	Probabilistic	Hypothetical Drug & BSC	£374,840	2.66	
		BSC	£201,966	1.53	151,944
Markov model	Deterministic	Hypothetical Drug & BSC	£375,882	2.08	
		BSC	£208,437	1.14	179,362
	Probabilistic	Hypothetical Drug & BSC	£377,759	2.10	
		BSC	£209,328	1.18	183,808

* Note that the table contains rounded values, but all decimals were used in the R code (calculations). BSC (Best Supportive Care).

4.3.3 State occupancy

Figure 13 contains an illustration of state occupancy by percentage in state and time in years on the y- and x-axis, respectively from both Markov and partitioned survival models. The graphics show that a hypothetical drug & BSC prolongs the time spent in earlier states as compared to BSC alone. In addition, it is also clear that the two model types use different methods to estimate state occupancy resulting in differently shaped curves. For example, the curve for the health state, able to walk, is less steep in the partitioned survival model, which equates to increased time in this state.

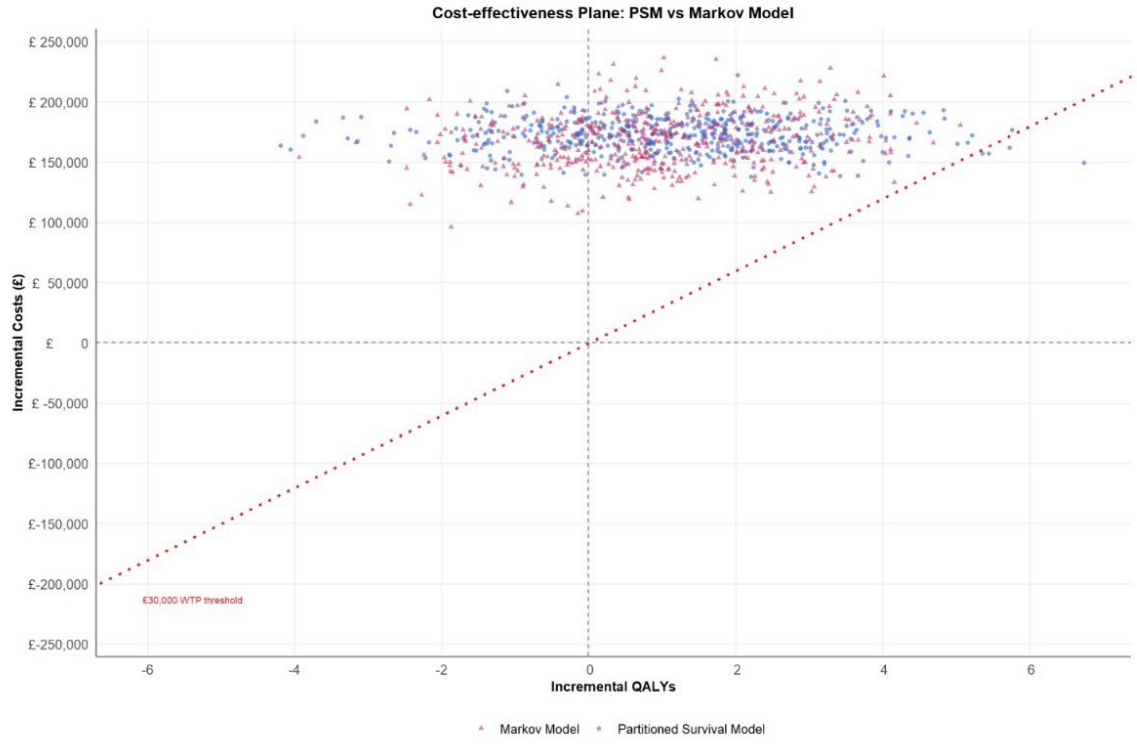
Figure 13: Comparison of health state occupancy



4.3.4 Cost-effectiveness plan and acceptability curves

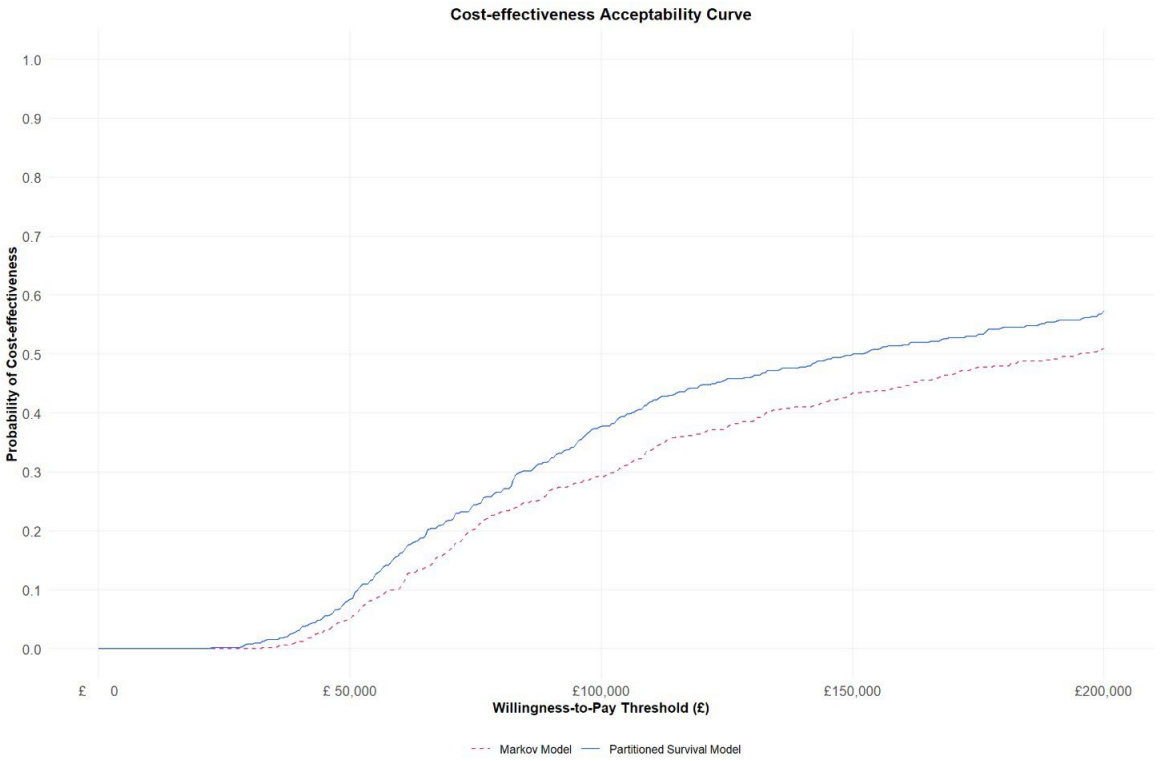
The cost-effectiveness scatterplot is illustrated in Figure 14, which largely showed that the simulations from PSAs were overlapping, but different patterns are noticeable. For example, the partitioned survival model seems to generate simulations with a wider range for incremental QALYs, and less incremental cost differences as compared to the Markov model.

Figure 14: Cost-effectiveness scatterplot



The cost-effectiveness acceptability curves in Figure 15 showed that a 0.5 probability of cost-effectiveness was reached earlier for the partitioned survival model, which was around a £150,000 willingness-to-pay threshold, but closer to a threshold of £195,000 for the Markov model.

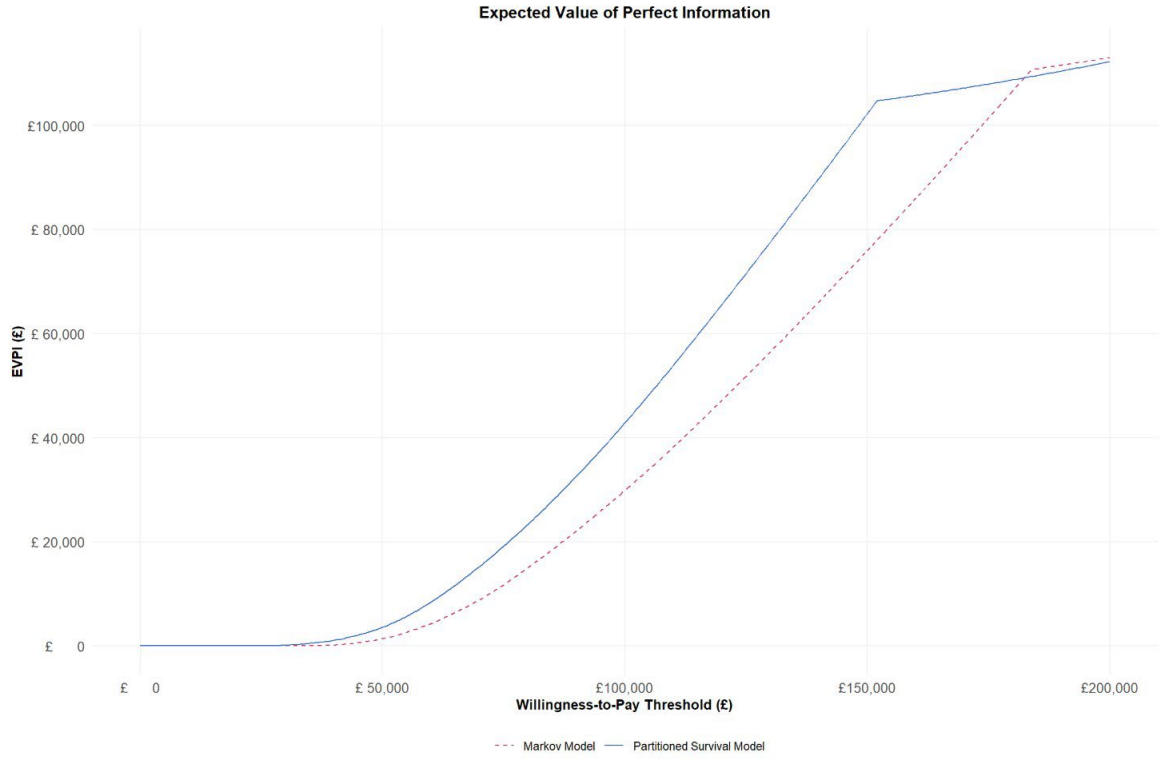
Figure 15: The cost-effectiveness acceptability curves



4.3.5 Expected value of perfect information

Figure 16 shows the EVPI curves, which have similar shapes, but EVPI is slightly higher for the partitioned survival model compared to the Markov model from approx. £30,000 – 170,000 willingness-to-pay thresholds, whereafter the curves converge.

Figure 16: EVPI at the individual level



4.3.6 Quality-adjusted life year shortfall calculator (sensitivity analysis)

NICE guidance states that QALY weights can be applied to reflect disease severity, with the weighting dependent on the lifetime QALY shortfall under BSC (Wailoo, 2024). The QALY shortfall calculator uses outputs from cost-effectiveness analyses and population characteristics to provide a QALY weight, which was 1.2 for all four analyses examined (Schneider *et al.*, 2021; Wailoo, 2024; McNamara *et al.*, 2023). This was then applied to estimate the weighted ICERs. The ICERs with weights applied for disease severity, shown in Table 23, are naturally lower than results without the weighting, shown in Table 22. However, they are still above the cost-effectiveness threshold of £30,000 per QALY.

Table 23: Results with QALY weights from the shortfall calculator

Model	Analysis	Intervention	Costs	QALYs	Weight	$ICER_{weighted}$
Partition Survival model	Deterministic	Hypothetical Drug & BSC	£357,937	2.42		
		BSC	£189,607	1,14	x 1.2	131,508
	Probabilistic	Hypothetical Drug & BSC	£374,840	2.66		
		BSC	£201,966	1.53	x 1.2	127,488

Markov model	Deterministic	Hypothetical Drug & BSC	£375,882	2.08		
		BSC	£208,437	1,14	x 1.2	148,444
	Probabilistic	Hypothetical Drug & BSC	£377,759	2.10		
		BSC	£209,328	1.18	x 1.2	152,564

* Note the QALY shortfall calculator settings: discount rate 3.5%, mean age 54 years, female 43%, scenario "reference case", and remaining discounted QALYs from BSC (called current standard of care in the Shiny app that is the shortfall calculator). In addition, the rounded data from Table 22 were used for these calculations with the shortfall calculator. BSC (Best Supportive Care)

4.4 Discussion

With this present study, I have demonstrated that it is possible to reconstruct and digitise IPD data for hypothetical cost-effectiveness modelling despite significant limitations in terms of data availability. To my knowledge, this is the first study to apply and discuss the potential of EVPI to inform data collection priorities in rare disease using MSA as a proxy.

4.4.1 What are the implications of using alternative model structures and strategies?

It is fundamental for the scientific merit of this study to highlight the differences and similarities between UMSARS Part IV and functional milestones, and why it makes sense to explore the functional milestones through conceptual modelling. The very first study of cost effectiveness in MSA, Grand *et al.* (2025) or Chapter 3, argued for the use of disability scales, e.g., UMSARS Part IV to inform model structures, which was based on a thorough investigation of model structures for disease analogues, which often use disability or functional scales. The disease analogues were progressive diseases of the central nervous system featuring symptoms overlapping with MSA. It was further hypothesised that increased costs and reduced quality-of-life data were correlated with disease progression on the disease-specific disability scale UMSARS Part IV (McCrone *et al.*, 2011; Grand *et al.*, 2025). The relationship between UMSARS Part IV and functional milestones was first illustrated by Saulnier and colleagues, and Table 24 contains a simplified illustration for comparison (Saulnier *et al.*, 2024). It is important to note that the first two items from UMSARS Part IV will need to represent the milestone, able to walk, despite being separate items, for a modelling comparison to be feasible. Noteworthy is that these two items were collapsed in a cost-of-illness study, which could potentially indicate comparable costs, because previously, utilities were found to be of comparable values in a cost-effectiveness analysis (Winter *et al.*, 2011b; Grand *et al.*, 2025). However, collapsing health states may result in small incremental differences being lost. Bedridden appears both in UMSARS Part IV item 5,

but is also a functional milestone, which is interesting since the functional milestones aid-required walking and wheelchair confinement are not explicitly mentioned.

Table 24: Comparison of UMSARS Part IV items and functional milestones

	Item 1	**Item 2**	**Item 3**	**Item 4**	**Item 5**
UMSARS Part IV	Completely independent. Able to do all chores with minimal difficulty or impairment. Essentially normal. Unaware of any difficulty.	Not completely independent. Needs help with some chores.	More dependent. Help with half of chores. Spends a large part of the day with chores.	Very dependent. Now and then does a few chores alone or begins alone. Much help needed.	Totally dependent and helpless. Bedridden.
Functional milestones	Able to walk		Aid-required walking	Wheelchair confinement	Bedridden

Data have shown that UMSARS Part IV responses may fluctuate over shorter periods, e.g., six-monthly intervals, likely due to day-to-day symptoms, but this is to a limited degree (Grand *et al.*, 2025). However, a Markov model allowing backwards transitions would be needed if researchers were to account for this, but it may depend on the dataset at hand. Moreover, this issue may be more pronounced when follow-up intervals are short, because people living with MSA are likely to progress in terms of severity measured by UMSARS Part IV over time. By contrast, functional milestones are likely to have a stronger adherence to a sequential progression assumption, e.g., progression from able to walk to aid-required walking and so forth. Although, e.g., wheelchair use will gradually increase. It is almost certain once this milestone is reached that a wheelchair will never not be needed, only if bedridden, because disease progression will not revert. This aspect has important implications for model development, which is that partitioned survival modelling becomes a viable option, if sequential progression can be assumed, which would be interesting to explore for UMSARS Part IV. Naturally, Markov modelling will remain an option with or without this progression assumption.

4.4.2 How do model structures and data sources influence cost-effectiveness?

The first cost-effectiveness analysis in MSA, Grand *et al.* (2025) or Chapter 3, was a Markov model informed by UMSARS Part IV health states, reported an ICER for the deterministic analysis of £142,036, which is lower than the £179,362 for the Markov model informed by functional milestones but higher than the £131,380 from the partitioned survival model in this analysis (without QALY weights applied) (Grand *et al.*, 2025). ICER estimates were based on hypothetical model demonstrations and were above a cost-effectiveness threshold of £30,000 per QALY.

4.4.2.1 Markov and partitioned survival models

Nonetheless, the result differences that were observed between cost-effectiveness analyses and model types were likely partly attributed to structural and parameter differences, e.g., Markov versus partitioned survival models. For example, the ability of partitioned survival models to handle fractional state occupancy within one model cycle through area differences in cumulative event probability curves. On the other hand, partitioned survival models are dependent on progressive disease trajectories or the availability of survival curves, e.g., problems may occur if survival data is immature. By contrast, Markov models can handle reversal of disease through backwards transitions. The issue of intermediate-state transitions can likely be solved in most cases with a shorter time horizon facilitated by multistate modelling. In addition, Markov models may provide more flexibility for complicated disease pathways, e.g., tunnel states or time-dependency (Briggs and Sculpher, 1998).

In linear models, cost-effectiveness results may differ less, and if the models are simple enough, it may not matter if a Markov or partitioned survival model is used. However, this assumes that the decision problem can be sufficiently answered using the same assumptions and data across the models. In models that are non-linear, e.g., non-linear Markov versus partitioned survival models, one may also expect that EVPI differs across the models. Minor differences in EPVI were observed in this study, but it is an area that would be interesting to explore. In the literature, there is not much research on EVPI in the context of structural uncertainty, but it was found that different assumptions for survival outcomes are critical as they influence cost-effectiveness (Smare *et al.*, 2020).

4.4.2.2 What are the notable differences in comparison to Chapter 3

The previous cost-effectiveness analysis, Grand *et al.* (2025) or Chapter 3, was informed by IPD from EMSA-SG with a 2-year follow-up for UMSARS Part IV whereas this analysis digitised functional milestones where subjects were followed for 17 years (Grand *et al.*, 2025). The digitisation and reconstructed IPD add to the uncertainty of cost-effectiveness results in this present analysis, e.g., assumptions on censoring. Moreover, this analysis required a strict assumption of sequential progression, which was not the case in the previous cost-effectiveness analysis in Chapter 3 since I had access to IPD which was major advantage, e.g., allowing

deaths to occur from all health states. In addition, the previous model in Chapter 3 used a shorter cycle length and thereby avoided intermediate state transitions from multistate modelling (Grand *et al.*, 2025). The efficacy modification was also different in that the method for efficacy incorporation modified transition probabilities directly in a transition matrix. This contrasts with the method used in this present study where time to events were increased in datasets using an arbitrary hazard rate, which meant that the same dataset could later be used in both model types (Chancellor *et al.*, 1997; Briggs, Sculpher and Claxton, 2006; Grand *et al.*, 2025). The hazard rate was applied to all health states, e.g., also for bedridden and death, which was done to avoid an overlap or unrealistic sequence of events, e.g., bedridden occurring before wheelchair confinement. It meant that there was a costing discrepancy in that drug costs were halted, but the treatment modification technically persisted at the bedridden state. Whilst this was a limitation, this assumption was kept to align with the previous cost-effectiveness analysis in Chapter 3 where treatment costs were assumed to be halted for ethical reasons (Grand *et al.*, 2025).

The results are not identical, likely because different approaches (models) and inputs (reconstructed data for functional milestones versus IPD) were used. However, it may be worth exploring partitioned survival approaches in more detail if IPD were available to relax the sequential progression assumption, e.g., to account for censoring or apply this approach using UMSARS Part IV, because this approach seems to provide cost-effectiveness estimates favouring an intervention. At least, as a sensitivity analysis.

4.4.3 The importance of quality-adjusted life year shortfall multipliers

The QALY shortfall multiplier was included in this analysis for two reasons: first, MSA is unlikely to fulfil one of the routing criteria for HST evaluation which is a point prevalence of 1:50,000 cases or less in England, when the prevalence for MSA is estimated at approximately 5 cases per 100,000 people in the UK (NHS Trust, 2023; National Institute for Health and Care Excellence, 2024). Other decision modifiers can potentially be applied to those rare diseases that are selected for HST evaluation, but it depends on the size of benefit. Second, the severity calculated by both proportional and absolute QALY shortfalls indicated sufficient severity to allow for a QALY weight, (multiplier) of 1.2. It is potentially quite important for manufacturers to understand these relatively new methods, e.g., absolute shortfall (tend to favour younger patients) and proportional shortfall (tend to favour older patients). Also, it would be pertinent to know which HTA bodies apply these measures, e.g., Norway, Netherlands, and UK as this might influence cost-effectiveness (Ali, Haines and Lee). For example, the deterministic analysis from the Markov model was reduced from £179,362 to £148,444 per QALY, for the ICER and weighted ICER, respectively.

4.4.4 Is it worth conducting expected value of perfect information analyses for rare diseases?

Pearson *et al.* (2018) did not identify examples of the application of value of information analysis for economic modelling in rare disease nor did the umbrella review of economic evaluations in rare diseases in

Chapter 2, but both studies recommend that it is used to inform data collection priorities (Pearson *et al.*, 2018; Grand *et al.*, 2024b). The EVPI will indicate if it is worthwhile to fund and conduct further research, e.g., if the value of eliminating uncertainty is higher than the costs of collecting new evidence (Briggs *et al.*, 2012; Barton, Briggs and Fenwick, 2008; Kunst *et al.*, 2023a). It was calculated at the individual level (per person) in this analysis, but the population EVPI is more useful, because it provides the upper limit for the value of a study aimed at reducing uncertainty across model parameters. However, it requires that the eligible treatment population is known including the period for which the decision problem is relevant (Heath, Kunst and Jackson, 2024).

Although these analyses may be more useful at the population level, they require additional considerations for rare diseases, especially during early cost-effectiveness modelling. First, a rare disease is often defined partly by its low prevalence, and multiplying individual-level EVPI with a small population will inherently contribute to lower EVPI than for more prevalent diseases (Richter *et al.*, 2015). Thus, minimising the likelihood that additional research will be worthwhile. On the other hand, uncertainty is most likely higher in cost-effectiveness models for rare diseases (Pearson *et al.*, 2018; Grand *et al.*, 2024b; Nestler-Parr *et al.*, 2018). So, an important question remains unanswered, that is “which of these factors contribute most to such a reciprocal relationship?” Second, rare diseases tend to suffer from multiple phenotypes, delayed diagnosis or misdiagnosis, and fewer clinical experts to consult for advice, which can make it difficult to define the eligible population. Furthermore, defining a time horizon for the decision problem may be difficult due to missing natural history of disease data or unknown disease trajectories (Grand *et al.*, 2024b). For these reasons, it may also be worth for researchers to consider if managed-entry agreements could play a role in reducing uncertainty for a specific rare disease prior to investing in additional studies solely based on EVPI analyses.

For increased robustness, it may be worthwhile also exploring EVPPI for a parameter or sets of parameters, which can help determine which parameters contribute most to uncertainty (Briggs, Sculpher and Claxton, 2006; Heath, Kunst and Jackson, 2024). For example, to further evaluate additional research versus managed-entry agreements, as the evidence generation plan may depend on which parameters are causing the uncertainty. Moreover, one can take additional steps and evaluate sample sizes or study designs if the costs of conducting research, e.g., for a set of parameters are lower than EVPPI. These methods are more complex because their methods depend on study design, model type and expertise, and they are called EVSI and ENBS. For example, these methods help determine study types and their optimal design to reduce uncertainty in decision-analytic models (Heath, Kunst and Jackson, 2024). Generally, research on the implications and relevance of these advanced methods in rare diseases are warranted.

4.4.5 Limitations

The reconstruction of IPD was challenging. Obviously, working with IPD of sufficient quality rather than having to reconstruct data from published Kaplan-Meier curves with limited information, is highly preferable. Only the total number at risks were reported, and the reconstruction of IPD relies on the strong assumption that there is no censoring. The subject numbers at risk or number of events would vastly have improved the quality of digitisation, but they were not reported in the original source paper for every event, except, indirectly for aid-required walking if assumed that the study population was event-free at study entry (Watanabe *et al.*, 2002). The only comment on censoring was that subjects (n = 40) undergoing tracheostomy (surgery) were treated as censored data (Watanabe *et al.*, 2002).

The lack of data on censoring led to a likely overestimation of survival. For example, a sequential progression assumption was needed, e.g., death could not occur from all health states, which would be expected from MSA. For example, severe urinary tract infections in earlier disease states, which is one of the leading causes of death (Papapetropoulos *et al.*, 2007). In terms of cost-effectiveness modelling this approach may also overestimate costs, because everyone needs to progress through all states, and accumulate more time in states that represent higher costs or lower quality of life, for example, bedridden. However, it was a necessary simplification given the reconstruction of IPD from available Kaplan-Meier curves, which nonetheless reflect the progressive and fatal nature of MSA. So even if functional milestones to inform partitioned survival models are preferable, much data would need to be generated to increase robustness of results. Future studies can increase their applicability or relevance for modelling by including utilities and cost data for UMSAR Part IV and functional milestones to further compare outcomes. However, there remain overwhelming precedence for using functional or disability rating scales to inform decision-analytic modelling in diseases similar to MSA, as shown in Chapter 3, which means that UMSARS Part IV, as it stands, is currently the most suitable option to define model health states (Grand *et al.*, 2025)

A digitisation approach that explicitly captures censoring (as a vector of censoring times) has been proposed for survival modelling, but this method was not applicable as it requires that censoring times are clearly marked and distinguishable on the curves, and similarly to the IPDfromKM package, it requires as a minimum, the total number at risk at baseline (time = 0) for all event curves (Rogula, Lozano-Ortega and Johnston, 2022). However, this cost-effectiveness analysis nonetheless required merging of multiple Kaplan-Meier curves, as opposed to one event only, e.g., death, to reconstruct a study population. In addition, events were assigned to subjects by looping through subject id's for each of the event datasets, which meant that subjects with lower id numbers were consistently assigned the earlier events, and it resulted in them having faster disease progression. Altogether, these factors added to the complexity of reconstructing IPD for cost-effectiveness modelling. It would be interesting for researchers to further investigate digitisation of multiple event curves and its integration into a unified reconstructed dataset.

Much research is needed to further examine the relationship between functional milestones as opposed to disability or functional scales, e.g., UMSARS Part IV to inform model structures and strategies for MSA. Irrespective of this relationship, a major limitation of this study was that cost and utility data were missing for functional milestones, which meant that these health-economic parameters were sourced as surrogate data from other studies that stratified by UMSARS Part IV (Winter *et al.*, 2011b; Wenning *et al.*, 2013; Grand *et al.*, 2025).

This analysis took a UK NHS perspective, but the time-to-event data were sourced from a Japanese population, which most likely have a different prevalence and distribution of phenotypes (Kaplan, 2024; Watanabe *et al.*, 2002). In addition, it seemed that Winter and colleagues did not include social services costs in their cost categories (inpatient care, drugs, special equipment, home care, and other direct costs) stratified by UMSARS Part IV, which play a large part for end-of-life planning and caregivers. A UK NHS and personal social services perspective would have been preferable (Burns and McFarland, 2020; Flabeau, Meissner and Tison, 2010; Grossauer *et al.*, 2022; Winter *et al.*, 2011b).

4.5 Conclusions

The results from partitioned survival and Markov models did not meet common cost-effectiveness thresholds, but generally, there exists a paucity of health-economic data for functional milestones that make results highly uncertain. In addition, the model demonstration made use of hypothetical interventions. The analysis was populated using reconstructed data for time to functional milestones from a publication. Functional milestones may be modelled using partitioned survival modelling if progression of events is assumed sequential, but more research is needed to assess this model type for UMSARS Part IV. Surrogate data were sourced from UMSARS Part IV for utilities and costs. Lastly, more research is needed to establish the relevance of EVPI analysis in rare diseases. To my knowledge, this was the first analysis to discuss EVPI in the context of rare diseases, and I made use of MSA as proxy to illustrate my points.

Chapter 5 Discussion

The main discussion of this thesis integrates all insights from the individual research projects collectively.

5.1 Summary of the thesis

5.1.1 Aim

The main research aim set out to explore issues, challenges and opportunities for economic evaluation in rare disease. A mix of methods, e.g., literature searches and cost-effectiveness modelling, was applied using a multi-study approach to achieve this aim. The multi-study approach meant that the work was undertaken as a series of research projects, which were reported as independent chapters in this thesis.

Figure 17 illustrates how this approach allowed for selected topics to be examined across multiple studies by differing methods. For example, issues with data scarcities to inform health-economic parameters were first identified as an issue in the umbrella review, but this issue became increasingly pronounced in each of the decision-analytic modelling studies in Chapter 3 and Chapter 4. This is why "data scarcities" is placed in the centre of the Venn diagram where all three circles overlap to indicate that this topic was relevant in all three research projects.

5.1.2 Chapters and research questions

This section outlines the flow of research by binding the chapters and research questions together into a clear and cohesive narrative. Chapter 2 was concerned with an umbrella review. The research question that the study attempted to answer was:

- In health economic research settings are there any issues and challenges for the economic evaluation of orphan drugs in rare diseases which apply less to other drugs?

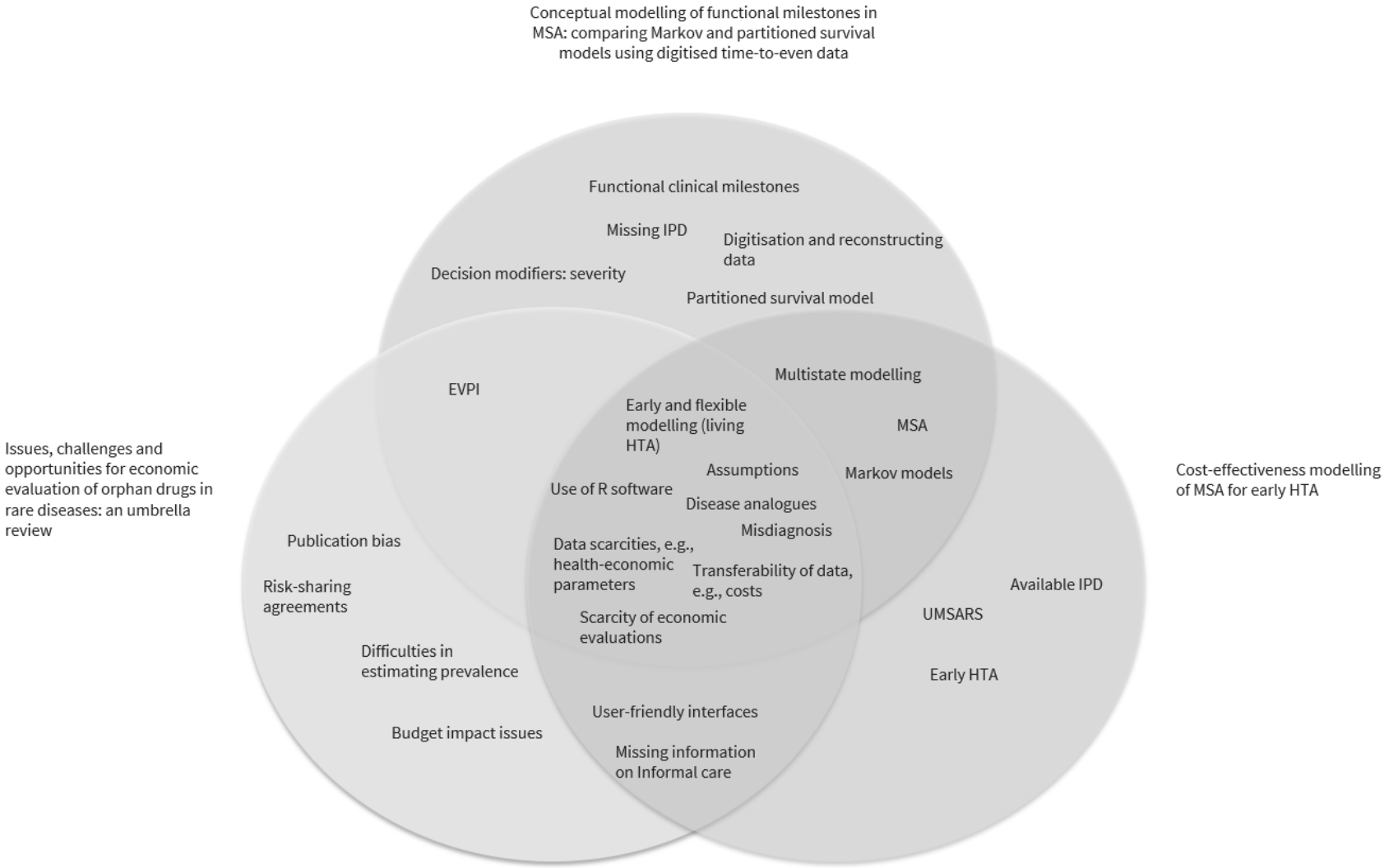
The review identified plenty of issues and challenges for economic evaluation of orphan drugs in rare diseases. Opportunities were difficult to identify and although this was not part of the research question, it was naturally important to identify them. Otherwise, this PhD would make little contribution to the chosen research area.

"*challenges for economic evaluations of orphan drugs in rare diseases are abundant, and solutions are not plentiful and rarely forthcoming"* (Grand *et al.*, 2024b, p. 8)

This statement was increasingly corroborated and evident throughout this PhD journey. However, it proved difficult to compare if challenges and issues applied less to other drugs, which was part of my research

question. Papers tended to focus on a specific rare disease or multiple rare diseases rather than how they compared to more prevalent diseases. Thus, it was not possible with confidence, to assert whether all issues for orphan drugs applied less to other drugs (Grand *et al.*, 2024b). Whilst this study did not sufficiently address if challenges for economic evaluation apply less to other drugs, it did offer significant and novel insights into the health-economic research challenges for economic evaluations of orphan drugs in rare diseases.

Figure 17: Venn diagram that illustrates overlap between topics across research projects (chapters)



Chapter 3 included, to my knowledge, the very first demonstration of an early HTA of MSA. The study sought to answer multiple research questions. The first question was:

- How can a conceptual model be developed for MSA to support cost-effectiveness analysis in early HTA, and what evidence gaps currently limit the reliability of such modelling?

It was important to obtain a vehicle to facilitate further examination of challenges and opportunities identified in the umbrella review. For example, a specific rare disease to be used in early HTA, and to facilitate cost-effectiveness modelling. MSA was chosen as a candidate for the reasons illustrated by the quote below.

"MSA is an excellent research candidate for this purpose due to its archetypal rare-disease attributes, for example, delayed diagnosis, misdiagnosis, multiple phenotypes, and low prevalence. To our knowledge, no economic evaluation exists, little is known about available health-economic parameters for MSA, and it is not clear if current available data are sufficient to construct an early conceptual cost-effectiveness model for early HTA" (Grand et al., 2025, p. 2).

In addition, there was an opportunity to make a significant research contribution to this specific disease. The second research question was:

- Can data gaps in early cost-effectiveness modelling be reduced, or model development facilitated, by conducting literature searches for disease analogues?

The umbrella review in Chapter 2 only discovered a single occasion of sourcing data from disease analogues. In this example, quality-of-life data for wheelchair confinement were sourced from MS (prevalent) to inform Duchenne's disease (rare) (Pearson *et al.*, 2018). This was largely an unexplored opportunity in rare diseases that warranted further exploration using MSA as a facilitator for this.

The diseases analogues were chosen together with clinical experts and apart from higher prevalence, they needed to present overlapping symptoms with MSA, e.g., with significant impact on functioning, quality of life and survival. The research results were disseminated at a conference (ISPOR) and included as part of a publication (Grand *et al.*, 2024a; Grand *et al.*, 2025).

The research questions in Chapter 4 were informed by previous research. For example, the publication that was based on Chapter 3 stated the following:

"In theory, it should be possible to triangulate our results by partitioned survival modelling (time-to-event analyses), whereby patients move through health states based on parametric equations as opposed to transition probabilities for Markov models, for example, by digitising available Kaplan-Meier curves and reconstructing patient-level data on clinical progression milestones…" (Grand et al., 2025, p. 11).

It was an interesting opportunity to both explore other model structures and triangulate previous results. In addition, the quotation from Chapter 3 or Grand *et al.* (2025) was very clear in informing the subsequent research questions that needed to be posed. They were:

- Can patient-level data be reliably digitised and reconstructed from existing published sources?
- Can alternative model structures based on functional milestones be developed to compare the advantages and disadvantages of different modelling strategies, such as Markov and partitioned survival models?

Lastly, early conceptual modelling was highlighted in both Chapter 2 and Chapter 3, but only Chapter 2 focused on the potential benefits of EVPI in rare diseases as an additional way to inform early evidence generation strategies. EVPI was furthermore recommended for modelling in rare diseases in two previous publications, one of them being the umbrella review (Pearson *et al.*, 2018; Grand *et al.*, 2024b). However, to my knowledge, the implications of EVPI in rare diseases have not been thoroughly examined despite the previous recommendations for its use. It was therefore a natural opportunity to consider in Chapter 4, together with the other research questions. Thus, last the research question became:

- Is it worthwhile to explore Expected Value of Perfect Information (EVPI) in the context of rare disease?

5.1.3 Main findings

The projects sought to answer relevant research questions, which could altogether, support the investigation of the main aim. A famous proverb that illustrates the logic behind this strategy is: how do you eat an elephant? One bite at a time.

The umbrella review in Chapter 2 made major research contributions by identifying an abundance of challenges and issues for economic evaluation, but fewer opportunities. The most obvious finding to emerge from this study was the scarcity of evidence, e.g., clinical effectiveness, costs, quality of life and natural history of disease data. This scarcity of evidence was a major cause for the limited quantity of economic evaluations, which could also partly be explained by biases, e.g., publication bias. These findings supported the notion that economic evaluation of orphan drugs in rare diseases is extremely challenging. The main opportunities identified were flexible and comprehensive cost-effectiveness modelling, e.g., living

HTA, user-friendly interface, early or timely modelling, EVPI, sourcing of data from similar diseases (disease analogues), and managed-entry agreements (Grand *et al.*, 2024b).

Early HTA of MSA in Chapter 3 corroborated findings, e.g., that health-economic parameters for rare diseases are scarce. In addition, it illustrated how important access to longitudinal registry data is for estimating disease trajectories, but such data tend to suffer from small sample sizes and limited study-duration for rare disease. The searches for disease-analogues were useful for establishing model structures, because at least for MSA, there was strong precedence for using functional rating or disability scales in disease analogues. The literature searches could potentially had been even more important, if IPD had not been available to inform transition probabilities and utility data. In addition, the very early modelling efforts highlighted data gaps that could be fulfilled by further research, e.g., the need for cost data stratified by all UMSARS Part IV health states. Lastly, it was possible to create user-friendly interfaces, which aligns well with the intentions of early HTA, which is to facilitate inclusion of a broad range of stakeholders (Grand *et al.*, 2025).

Comparative modelling in Chapter 4 illustrated challenges of missing IPD, which required that Kaplan-Meier curves were digitised and IPD reconstructed to inform disease trajectories. A change from UMSARS Part IV to functional milestones to inform health states had large implications for modelling despite an established relationship between these two measures. It meant that data to inform parameters were missing at the study outset for transition probabilities, costs and utilities.

The analysis furthermore highlighted potential opportunities for EVPI, EVPPI, EVSI and ENBS, but these methods should be considered in relation to other initiatives to reduce uncertainty, e.g., managed-entry agreements. In addition, this study was likely the first to discuss implications of EVPI at the population level for rare diseases. For example, highlighting a potential reciprocal relationship caused by low prevalence (small populations) contributing to lower EVPI at the population level whereas higher uncertainty in cost-effectiveness analysis for rare diseases might contribute to higher EVPI, thus, making it unclear if additional research will be worthwhile for rare diseases.

5.1.4 Integrated summary

Taken together, one of the more significant findings to emerge from my research, is that it is notoriously difficult to conduct economic evaluations of orphan drugs in rare diseases. This PhD project extends current knowledge of issues and challenges by providing a comprehensive overview of them in Chapter 2, e.g., grouped by health-economic parameters, economic evaluations, and estimation of value / reimbursement in an umbrella review (Grand *et al.*, 2024b). The present cost-effectiveness analyses of MSA confirm these previous findings and contribute with important corroboration and nuances on these challenges (Grand *et al.*, 2025). Lastly, despite its explorative nature, this PhD project offers some insights into potential

opportunities for economic evaluation of rare diseases that are attainable with modest resources. Whilst these opportunities are not magic bullets (solutions), they were partially substantiated by practical application. For example, timely flexible modelling using user-friendly interfaces to engage stakeholders early and to identify data gaps were promising. For example, searches for disease analogues to reduce data gaps was an opportunity that could be researched further as there was potential for the use of surrogate data, although the applicability varied across studies. In addition, researchers should explore EVPI to guide data collection priorities. I do describe a reciprocal relationship, which is a potential issue, but EVPI remains one of the few data driven approaches to inform data collection, which is why researchers should further explore this opportunity.

5.2 Novel contributions

5.2.1 Cost-effectiveness modelling in rare diseases is challenged by missing data

Ollendorf and colleagues brilliantly capture the current state of affairs " there is simply no magic solution to the conundrum of assessing the evidence on clinical effectiveness, economic impact, and value of drugs to treat rare diseases" (Ollendorf, Chapman and Pearson, 2018). Overall, the research confirmed this statement, but it did also make several noteworthy contributions to enhance and progress the current understanding of the issues, challenges and opportunities.

My viewpoint on the perception of "no magic solution" is that multiple courses of action need to be taken in a coordinated effort. A sensible approach to tackle this issue might be to collectively consider a wide range of opportunities to address this conundrum. Hopefully, a multi-faceted approach may provide enough synergism to get researchers closer to a solution. However, magic will unfortunately not be part of this process. This PhD attempts to substantiate rare-disease challenges, which naturally makes it easier to pinpoint potential areas for improvement, but rare diseases should also be a key political priority, if anything should change, because health economists operate within a framework made up of political boundaries.

The umbrella review in Chapter 2 contributed to existing knowledge on data scarcities for rare diseases by providing a systematic overview of them while discussing the consequences of missing data (Grand *et al.*, 2024b). For example, scarcity of health-parameters remains a major issue and is likely one of the leading causes for missing economic evaluations in rare diseases, together with the issues of publication bias (Ollendorf, Chapman and Pearson, 2018; Pearson *et al.*, 2018; Grand *et al.*, 2024b).

A series of model demonstrations on cost-effectiveness using MSA as an archetypal proxy for rare diseases support the umbrella review findings. The scarcity of health-economic parameters was very apparent in the literature for the health-economic models that were constructed during this PhD (Grand *et al.*, 2025). For

example, chronic progression of rare diseases caused insufficient follow-up in epidemiologic studies, e.g., two years of follow-up in the EMSA-SG dataset (Grand *et al.*, 2024b; Grand *et al.*, 2025). In addition, there were few clinical experts, pronounced issues with delayed diagnosis or misdiagnosis, multiple phenotypes and sparse data on prevalence (Grand *et al.*, 2024b). These attributes supported the decision to use MSA as a proxy for rare diseases and it was concluded that it was an excellent research candidate due to archetypal properties despite being an opportunistically chosen disease, which critics righteously can argue is a key limitation (Krismer and Wenning, 2017; Mészáros *et al.*, 2020; Gilman *et al.*, 2008; Chrysostome *et al.*, 2004; Bower *et al.*, 1997; Osaki *et al.*, 2011; Sakushima *et al.*, 2015; Bjornsdottir *et al.*, 2013; Chiò, Magnani and Schiffer, 1998; Mazzucato *et al.*, 2014; Grand *et al.*, 2024b).

It was far from a plug and play exercise to populate models, even for a minimum of the required health-economic parameters. The use of available data required assumptions and transparent reporting of limitations. Despite this, the data were only just enough to populate all model parameters in Chapter 3 (Grand *et al.*, 2025). The issue of populating models became increasingly difficult in Chapter 4 when functional milestones were used to inform health states, rather than UMSARS Part IV, for Markov and partitioned survival modelling, and much surrogate data were used, which required some very strong assumptions. It was necessary to apply assumptions in all demonstrations of cost-effectiveness modelling. To summarise, the research systematically outlined and confirmed issues of data scarcities for economic evaluation of rare diseases, and that assumptions likely will be necessary in decision-analytic modelling.

5.2.2 Missing health-economic evaluations in rare diseases

The research projects that were undertaken provide additional evidence with respect to the scarcity of economic evaluations, and their quality, but also regarding the difficulties in conducting them. The umbrella review in Chapter 2 concluded that economic evaluations for rare diseases are scarce, which was supported by two systematic reviews. For example, a study of 11 orphan drugs noted that their findings resembled previous research that found a remarkable absence of health-economic evaluations (Kanters *et al.*, 2013). Similarly, in a systematic review of economic evaluations defined as cost-effectiveness and cost-utility analyses none were identified (Lee *et al.*, 2022). In addition, the evidence synthesis for MSA corroborated these findings as no previous economic evaluations were identified (Grand *et al.*, 2025).

In terms of quality, it appears that the majority of economic evaluations have moderate quality and their failure to achieve higher quality is attributed to lack of health-economic parameters, but reporting issues were also reported (Cheng *et al.*, 2012; Weidlich, Kefalas and Guest, 2016; Leonart *et al.*, 2018; Park, Griggs and Suh, 2015; Thorat, Neumann and Chambers, 2018).

A fundamental property of health-economic parameters is their ability to inform economic evaluations. Each of the research chapters illustrated that missing health-economic data is a primary concern for

economic evaluation in rare diseases (Grand *et al.*, 2024b; Grand *et al.*, 2025). It is natural human behaviour if missing data deter analysts from even attempting to undertake economic evaluations in rare diseases as the challenge may seem impossible (Schuller, Hollak and Biegstraaten, 2015).

I do not believe that economic evaluations in rare disease should be primarily judged on the availability of data. First, data scarcities is an inherent issue for rare diseases (Grand *et al.*, 2024b). Second, modelling is an iterative process that can add to a model as the understanding of the decision problem deepens (Kunst *et al.*, 2023a; Roberts *et al.*, 2012). Model development should be independent from initial considerations of data availability, which is why early modelling should be appraised mostly in terms of its model structure, and not available data (Roberts *et al.*, 2012; Caro *et al.*, 2012).

Lastly, analysts should not refrain from conducting economic evaluations in rare diseases as there is an option to use expert opinion if data are missing. Although, experts are fewer in numbers for rare diseases and may be increasingly difficult to recruit (Caro *et al.*, 2012; Briggs *et al.*, 2012). For example, NICE will consider expert elicitation or expert opinion if empirical evidence, e.g., registries or randomised controlled trials are missing, which is an important opportunity not to be overlooked (National Institute for Health and Care Excellence, 2025b).

However, it is admittedly very difficult (not impossible) to conduct cost-effectiveness modelling in rare diseases, which was exemplified in my work. For example, needing to digitise and reconstruct patient-level data for functional milestones, or needing to implement efficacy modifiers to slow disease progression or reduce clinical events. In other words, using innovative techniques where little previous guidance exists. However, there is usually a way forward, and the benefits of doing it are not to be neglected if modelling is considered in an iterative modelling framework, which is to be advocated for, since it can inform evidence generation and subsequently later modelling iterations.

As far as biases are concerned, the umbrella review is likely providing the most comprehensive description so far documenting publication bias and other types of biases as potential causes for missing economic evaluations (Grand *et al.*, 2024b). It is crucial to appropriately evaluate and consider biases in research. For example, publication bias, such as only publishing if cost-effectiveness analyses reach cost-effectiveness thresholds for the favoured intervention. Failure to consider publication bias can lead to conclusions like *"this study demonstrates that orphan drugs can meet traditional reimbursement thresholds*" (Picavet, Cassiman and Simoens, 2015, p. 307). This overambitious claim is discussed in more detail in the umbrella review (Grand *et al.*, 2024b). The conclusion, however, was contested in this PhD.

It was also shown that different types of biases are problematic in economic evaluation of rare diseases as multiple types were identified in systematic reviews (Cheng *et al.*, 2012; Kanters *et al.*, 2013; Picavet, Cassiman and Simoens, 2015; Park, Griggs and Suh, 2015; Woersching, Borrego and Raisch, 2016). For

example, higher frequency of publication in countries with post-marketing obligations (Schuller, Hollak and Biegstraaten, 2015). Publication bias including positive result or industry sponsorship biases, failure to describe magnitude or direction of bias, and selection bias emerged as potential predictors of missing health economic evaluations (Plüddemann, Banerjee and O'Sullivan, 2017; Holman, Bero and Mintzes, 2019). Further research regarding the role of biases would be worthwhile as they can influence the public opinion if decision-makers take conclusions at face value. For example, that decision-analytic modelling is not particularly challenging and that cost-effectiveness thresholds can be met in economic evaluations for rare diseases if evaluated by traditional reimbursement frameworks. Hopefully, my research will help impose healthy scepticism toward economic evaluations that were published resulting in questions such as: Why was this specific study published in the first place? Who was the sponsor? Were the results in favour of the sponsor? What dispositions were previously made for data generation, and how were they sponsored and integrated into the analysis?

5.2.3 Disease-specific contributions: novelty through exploration of multiple system atrophy

The generalisability of disease-specific contributions is naturally subject to certain limitations. For instance, I exclusively focus on MSA in cost-effectiveness modelling. Notwithstanding, this limitation, a large proportion of findings were shared and corroborated between the umbrella review (a review of systematic reviews) and other research projects in MSA. This should improve the confidence in the chosen research candidate (MSA) and results.

5.2.3.1 The very first cost-effectiveness analysis in multiple system atrophy

To my knowledge, this PhD is the very first to conduct a full assessment of evidence for health-economic parameters and to construct a cost-effectiveness analysis for MSA. This is important for multiple reasons, first, it allows researchers to design studies that can inform modelling. Second, it will become important for manufacturers that seek to bring interventions to the market in MSA, as they will naturally have to consider this research. For example, a relevant study to initiate would be a resource utilisation study stratified by all UMSARS Part IV items across relevant geographies. I, nonetheless, make it clear that the proposed model is likely to develop over time.

5.2.3.2 A widespread issue in multiple system atrophy was missing cost data

The umbrella review found that cost-of-illness studies were scarce for rare diseases (Dhawan *et al.*, 2019; Gruhn *et al.*, 2022; Jakes *et al.*, 2021; Kwon *et al.*, 2021; Rice *et al.*, 2018; Weidlich, Kefalas and Guest, 2016; Angelis, Tordrup and Kanavos, 2015; Garcia-Perez *et al.*, 2021; Pearson *et al.*, 2018). This was also the case in MSA, where only two studies on costs were identified during literature searching (Winter *et al.*, 2011b; McCrone *et al.*, 2011). Recently, a paper was published that provided odds ratios for utilisation frequency of certain healthcare resources, e.g., hospitalisation comparing MSA and a control group (Kaplan, Xie and

Kumar, 2025). However, this paper could not directly inform any of the models developed, because odds ratios were provided for very specific outcomes, e.g., an odds ratio for one or more prescriptions at baseline and follow-up between an MSA and control cohort.

An additional review finding from Chapter 2 was also that generalisability of cost data was an issue for cost-of-illness studies, e.g., to use data across geographies, which was influenced by several factors such as perspective, data sources, nomenclature, assumptions, guidelines and value frameworks (Gruhn *et al.*, 2022; Querol *et al.*, 2021; Angelis, Tordrup and Kanavos, 2015; Katsigianni and Petrou, 2022; Garcia-Perez *et al.*, 2021). Again, these challenges were also pronounced in MSA, where cost data were aggregated and grouped by health states without sufficient justification by the authors as to why. (Winter *et al.*, 2011b). The origin of data was German but used to inform a UK NHS perspective, which was obviously not optimal. However, UK cost data were available, which could potentially have informed the analysis (McCrone *et al.*, 2011). So why was it not used?

It could potentially have been used, but this would require e.g., expert elicitation to stratify resource use across the relevant model health states, because results were not stratified across UMSARS Part IV in the publication (McCrone *et al.*, 2011). This would naturally require significant time and resources, e.g., recruitment of experts, and this approach would simply present a different set of limitations. Also, it is important to remember that an important aim of early HTA was to inform data gaps to allow for timely data generation. Considering this, the choice of data source for costs was pragmatic in that the chosen data could directly inform health states. It was most important to communicate the data gaps. For example, the urgent need for up-to-date cost data stratified by all model health states.

The data limitations were unavoidable, but the difference between cost-of-illness studies can be explored further by assessing the implications for cost-effectiveness results in the Shiny app. The app was made publicly available partly to serve this explorative purpose, which is to explore cost-effectiveness through user-friendly interfaces, e.g., easy change of inputs to reflect different data sources.

5.2.3.3 Considering data and structural uncertainty with different model approaches (Markov and partitioned survival models)

The feasibility of cost-effectiveness modelling in economic evaluation of rare diseases is clearly very dependent on the chosen model attributes, e.g., model types or health states. It is something to be aware about if very early cost-effectiveness modelling is required to inform, e.g., go/no-go decision in early pipeline development. For example, Markov models allow much flexibility in terms of deciding on allowed transition, e.g., restricting or allowing certain transitions between health states, like reversed disease progression. By contrast, partitioned survival models are dependent on sequential progression by disease severity.

An important finding was the challenge associated with calculating transition probabilities for Markov models in rare diseases. Disease progression data were collected at six-monthly or yearly intervals, but subjects were few, which makes non-parametric count methods inadequate unless strong assumptions can be reliably made about constant transition rates. In addition, in Chapter 4, intermediate state transitions were observed in the model comparison between Markov and partitioned survival modelling due to the longer (yearly) cycle length, which was elegantly handled by the partitioned survival model through survival equations. However, in Chapter 3, the Markov model for early HTA this was less of an issue because multistate modelling allowed for specification of a cycle length of one month despite having collected data at six-monthly intervals. This was feasible because multistate modelling estimated transition probabilities by maximum likelihood estimation (Jackson, 2011).

The comparison of Markov and partitioned survival models that were informed by functional health states rather than UMSARS Part IV further illustrated the point around the importance of choosing model attributes. For example, UMSARS Part IV is an established measure of disability in MSA and frequently reported in natural history of disease studies (Cao *et al.*, 2018; Foubert-Samier *et al.*, 2020; Golden, McCreary and Vernino, 2022; Jecmenica-Lukic *et al.*, 2021; Jecmenica-Lukic *et al.*, 2018; Kollensperger *et al.*, 2007; Low *et al.*, 2015; May *et al.*, 2007; Wenning *et al.*, 2013; Watanabe *et al.*, 2002; Saulnier *et al.*, 2024).

By contrast, health-economic parameters for functional milestones were largely missing. The data to inform functional health states were digitised from published Kaplan-Meier curves, despite much difficulty. One paper reported all milestones, but rarely were all functional states included in the same paper (Watanabe *et al.*, 2002). Likely because of the extended time horizon needed to capture all events. For example, two studies focused on wheelchair dependency, but not the following state that is bedridden (Zhang *et al.*, 2023; Wenning *et al.*, 2013). What is more, data to inform costs and utilities were not available for functional milestones and this meant that UMSARS Part IV data were used as surrogate data to facilitate a Markov and partitioned survival model comparison.

A main insight was that whether a cost-effectiveness analysis is feasible or not, may be highly influenced by the initial modelling decisions, e.g., choice of health states, because data are very limited. These findings further support the idea that proposing early decision-analytic model structures may help guide the inclusion of health-economic endpoints, for example, in natural history of disease studies or clinical trials (Grand *et al.*, 2025).

It is not that I advocate for deviation from published guidelines. It is highly recommendable to consider model specification as an iterative process that should be independent from initial considerations on data availability to inform health-economic parameters (Kunst *et al.*, 2023a; Roberts *et al.*, 2012; Caro *et al.*, 2012). I merely highlight that model specifications are likely to have a large impact on the feasibility of an analysis. It is not hard to imagine a scenario where one would have thought of cost-effectiveness modelling in MSA as

impossible, resultingly in no further modelling iterations with alternative parameters, if the initial model had been constructed using functional milestones assuming that this was the optimal choice to inform health states.

5.3 Impact of this work

The impact of this work was multiple for me and others. Personally, I am most proud of how well the umbrella review has performed in terms of activity (Grand *et al.*, 2024b). I suspect that it is highly unusual that a publication has been accessed almost 10,000 times before the PhD has even been awarded. It reflects a high interest in economic evaluation of rare diseases but hopefully also that the publication represents a relevant research contribution. It has increased awareness of my persona, which means that I am receiving invitations from journals to volunteer as a reviewer and to present on economic evaluation of rare diseases at conferences and alike. I hope to convince a group of people to host an issue panel at ISPOR, because a session on economic evaluation of rare diseases is missing. The need is there and there were previously sessions on the inflation reduction act and real-world evidence in the context of rare diseases, but my proposed issue panel is needed. Amongst other implications is my personal upskilling, which has naturally taken place over the last three years.

As an industrial PhD student, I have been placed in a pharmaceutical company where I have contributed with knowledge over the years and helped where I could, e.g., with modelling. I am proud to state, but without revealing sensitive information, that the data gaps highlighted through early HTA and cost-effectiveness modelling for MSA will result in substantial investments into data generating studies to fill data gaps identified through this PhD. Innovation Fund Denmark write on their website that industrial PhD students combine new knowledge and concrete challenges, and the results is value-creation for all parties (industry and academia) (Innovation Fund Denmark, 2025). I would like to believe that this objective was met.

5.4 Recommendations for economic evaluations in rare diseases

5.4.1 Recommendations for nomenclature

It is my opinion that nomenclature is very important as it improves conciseness, provides researchers with a common language, and limits misunderstandings.

5.4.1.1 The definition of "orphan drug" and "rare disease"

Richter and colleagues previously conducted a comprehensive review (report of the ISPOR Rare Disease Special Interest Group) of terminology in rare diseases, which showed terminology was highly varying across organisations and jurisdictions including prevalence thresholds to define rare diseases (Richter *et al.*,

2015). To honour their important work, which hopefully will be adopted by others, this document adopted strict nomenclature based on their findings. For example, using these quotes:

"… there is some global consistency in using the terms "rare disease" and "orphan drug" preferentially in defining a rare disease and the technologies associated with rare diseases, respectively. (Richter et al., 2015, p. 913)"

or

"Based on our results, the terminology "rare disease," coupled with a prevalence threshold in the range of 40 to 50 cases/100,000, could present a realistic starting point for a harmonized definition of rare diseases. (Richter et al., 2015, p. 914)"

These quotes were interpreted in the following manner. The first quote states that an orphan drug is a technology associated with a rare disease. In addition, the report recommends using the term rare disease. The terminology for a rare disease included a prevalence threshold, which was a minimum of 40 to 50 patients per 100,000 people. I encourage others to adopt the wording in their future work, like I have done in this PhD, as significant effort was directed towards harmonisation of nomenclature by the special interest group.

5.4.2 Recommendations for dealing with data scarcities in rare diseases

Several opportunities were identified that could help alleviate issues with data scarcities in economic evaluation of rare diseases. However, most, if not all, of these opportunities require additional research to assess and realise their full potential.

5.4.2.1 Is it an opportunity to use disease analogues to facilitate model development?

Importantly, the focus on potential opportunities further extended the current knowledge by proposing recommendations that had potential to alleviate issues related to data scarcities. However, as mentioned previously, there is simply no magic solution (Ollendorf, Chapman and Pearson, 2018). Notwithstanding these limitations, my research suggests to further examine the potential of literature searches for disease analogues. As alluded to earlier, the umbrella review only discovered a single occasion of sourcing of data from disease analogues. Quality-of-life data were sourced from MS (prevalent) to Duchenne's disease (rare) for wheelchair confinement (Pearson *et al.*, 2018). This largely made it a novel opportunity for exploration and although the original idea came from literature, this PhD was the first to systematically assess its full potential.

This opportunity was evaluated as part of the evidence assessment in Chapter 3 by searching for health-economic parameters in economic evaluations of disease analogues. Interestingly, many cost-effectiveness analyses were found for the disease analogues, which was an early positive indicator of feasibility (Ayati *et*

al., 2021a; Ayati *et al.*, 2021b; Bolanos-Diaz, Sanabria-Montanez and Custodio-Capunay, 2019; Furneri *et al.*, 2019; Kremer *et al.*, 2020; Lazzaro *et al.*, 2022; Martins *et al.*, 2023; Michels *et al.*, 2019; Montgomery *et al.*, 2022; Nakhaipour *et al.*, 2020; Navarro and Betancur, 2023; Pinheiro *et al.*, 2020; Poveda *et al.*, 2020; Schur *et al.*, 2021; Spelman *et al.*, 2024; Spelman *et al.*, 2022; Stanisic *et al.*, 2019; Taheri, Sahraian and Yousefi, 2019; Visser *et al.*, 2022; Walter *et al.*, 2019; Xu *et al.*, 2019; Auguste *et al.*, 2020; Baharnoori *et al.*, 2022; Cortesi *et al.*, 2022; Daroudi *et al.*, 2023; Ginestal *et al.*, 2023; Giovannoni *et al.*, 2019; Hua *et al.*, 2019; Lasalvia *et al.*, 2020; Rezaee *et al.*, 2019; Acosta *et al.*, 2021; Hunter *et al.*, 2021; Oppe *et al.*, 2021; Zhao *et al.*, 2022; Palmer *et al.*, 2024; Dimitrova *et al.*, 2023; Mantovani *et al.*, 2020; Michel *et al.*, 2022; Renfrew *et al.*, 2019; Rosato *et al.*, 2021; Huygens and Versteegh, 2021; Versteegh *et al.*, 2022; Chaudhuri *et al.*, 2022; Kalabina *et al.*, 2019; Pahwa *et al.*, 2023; Gajdos *et al.*, 2021; Thakore *et al.*, 2018; Thakore *et al.*, 2020).

Disease analogues proved most useful for conceptualisation of model structures. For example, overwhelming support for Markov models informed by functional or disability scales (Grand *et al.*, 2025). Increasing confidence and robustness of health-economic models is very important, because health-economic models are at the heart of economic evaluation and included in HTA processes. For example, health-economic models decide reimbursement for orphan drugs, and reimbursement will ultimately decide if patients can access treatments (Hughes-Wilson *et al.*, 2012; Nestler-Parr *et al.*, 2018; Gammie, Lu and Babar, 2015). The reimbursement example illustrates the high level of interconnectivity from early identification of health-economic parameters and later consequences for patients. In addition, it may also be easier for clinical experts, which are difficult to find, to advise on model structures in rare diseases if they are provided with reference-cases in the form of economic evaluations for disease analogues, because it is unlikely they will have much experience with model development (Grand *et al.*, 2024b).

My research has posed questions in need of further investigation, but in general, the concept of disease analogues show potential to alleviate challenges for economic evaluation in rare diseases. However, considerably more work needs to be conducted to determine the applicability in other rare diseases. Can searches for disease analogues help with sourcing of surrogate data or help in determining model structures in other rare diseases? Is it primarily useful for model development and less useful for surrogate data? A definition for rare disease analogues in MSA was proposed but the precise process of identifying disease analogues remains to be elucidated. A natural progression of this work is to analyse and propose guidance on establishing disease analogues. How should this guidance look? Who will initiate these efforts? To summarise, I recommend that disease analogues are further investigated to assess their potential for informing model structures and for sourcing surrogate data in rare diseases.

5.4.2.2 Recommendations for early planning of evidence generation

The importance of timely evidence generation should not be overlooked, to a degree, that carried my thoughts to the historical motto of Maersk (a Danish shipping company): *“with timely care” (Lunde, 2020)*.

The motto emphasises the importance of doing things properly and on time, so that preventable loses are avoided.

This point is substantiated by focussing on natural history of disease data. The umbrella review in Chapter 2 found that early planning of registries or natural history of disease studies may alleviate data issues (Grand *et al.*, 2024b). The IPD from EMSA-SG was pivotal for cost-effectiveness modelling in MSA in Chapter 3. It facilitated estimation of utilities and transition probabilities across model health states (UMSARS Part IV), which would not otherwise have been possible as data were not identified during literature searches (Grand *et al.*, 2025). In other words – the analysis would have been very difficult without.

However, in rare diseases it is challenging to set up prospective data collections, and it needs to be done early to have sufficient data at model preparation stages. Patients are difficult to identify due to delayed diagnosis and misdiagnosis, and recruitment will need to happen from specialist centres or from multiple countries. This is likely needed to obtain a sufficient sample size but it increases the costs of research, as it is more troublesome than for more common diseases where recruitment is easier (Grand *et al.*, 2024b). However, manufacturers do not have much choice, it is a do-or-die situation for economic evaluation in rare diseases. To summarise, it is strongly recommended to make early plans for evidence generation in rare diseases. It is, moreover, recommended for researchers to engage in early HTAs to inform early pipeline decision, e.g., go/no-go decisions, but moreover, use early HTA to inform future data collection efforts.

5.4.2.3 Is value of information analysis an overlooked opportunity in rare diseases?

Chapter 4 and Section 1.1.7 explain the concept of value of information analysis, which is why it will not be revisited here. Importantly, value of information analysis was first recommended to inform data collection priorities in a targeted literature review on economic modelling in rare diseases (Pearson *et al.*, 2018). The targeted review did not identify any examples of previous use in the literature, nor did they elaborate on the potential application and complications, apart from their recommendation to use it. This PhD significantly advanced this recommendation through practical application and discussion. The umbrella review did not identify any examples but expanded the recommendation to suggest that both EVPI and EVPPI should be applied as a minimum to inform evidence generation in rare diseases. Only EVPI was applied in a cost-effectiveness analysis of MSA using both Markov and partitioned survival modelling approaches in Chapter 4 as a humble first step to partly enhance the current understanding of this recommendation. It would be interesting to examine the role of structural uncertainty in relation to value of information analysis when comparing different model types.

These findings will doubtless be much scrutinised, but there were some immediate dilemmas that were relevant to rare diseases that needed consideration. For example, small populations translate to lower EVPI at the population level, by contrast, uncertainty may be higher in cost-effectiveness analyses for rare

diseases resulting in an increase of EVPI. It was an undescribed reciprocal relationship before this research commenced. The challenges for EVPI in rare diseases were elucidated in the research projects conducted. For example, the application of EVPI might be further complicated by multiple phenotypes, delayed diagnosis or misdiagnosis, and fewer clinical experts to consult for advice, which can make it difficult to define the eligible population. Furthermore, defining a time horizon for the decision problem may be difficult due to missing natural history of disease data or unknown disease trajectories (Grand *et al.*, 2024b). For these reasons, I do recommend EVPI is evaluated in conjugation with other opportunities that may reduce uncertainty through data collection. For example, EVPI can help determine the potential of managed-entry agreements, before significant investments into further research are made. I recommend additional research on EVPI and related methods, as an important opportunity to inform data collection efforts and design studies, but especially in the context of rare diseases where experience is scarce.

5.4.3 Recommendations for economic evaluation of multiple system atrophy

Further work is required to establish the most appropriate model structure for MSA. The research in Chapter 3 and Chapter 4 will serve as a basis for future studies. There is a paucity of data for functional milestones. By contrast, UMSARS Part IV is an established measure of disability, which means that more data were available. For example, there were plenty of supporting evidence to suggest a disability scale to inform modelling in MSA if disease-analogues were considered (Grand *et al.*, 2025).

However, a major limitation, and opportunity for further research was that models were populated with currently available data, which meant that the model health states were bound by currently available data. For example, for costs. As alluded to previously, an iterative modelling approach would be helpful in this case, e.g., generate cost data across all UMSARS Part IV items to capture all potential incremental differences for disease progression. In other words, a first iteration could show that cost data were missing, which limits the number of health states. The second iteration (after data generation of cost data) could warrant a model update and perhaps a need for further evidence generation is discovered through EVPI. For example, another model iteration with updated registry or phase II data? In this way – there is continuous development. Decision-making can improve over time, e.g., to inform pipeline development. It could also be that a launch plan is continuously informed by early modelling, which adds a country-specific element to modelling. For example, an additional need for country-specific data, e.g., resource utilisation, because generalisability of cost data were shown to be problematic across geographies (Grand *et al.*, 2024b). For example, cost data in MSA exist for Germany, France and UK, which may be far from enough to launch an intervention broadly across HTA markets (McCrone *et al.*, 2011).

Much can be done to further improve modelling in MSA. As previously mentioned, there is abundant room to determine non-functional events for inclusion in modelling, e.g., depression (Grand *et al.*, 2025). It is my hope that by increasing awareness of this disease that people will start to consider biomarkers or

responder/non-responder analyses, or other types of analyses that could have potential implications for cost-effectiveness.

The overarching recommendation is to advance modelling in MSA in an iterative modelling framework, which both informs data generation efforts, but also prompts additional iterations when new data becomes available.

5.5 Limitations

5.5.1 Drawbacks of the work presented

The umbrella review in Chapter 2 had methodological limitations. For example, only one reviewer conducted the quality assessment and full screening of papers despite having two reviewers screen titles and abstracts. In addition, conducting a systematic review of systematic reviews (umbrella review) meant that relevant information could have missed, also because searches only included studies from 2010 and onwards. For example, narrative reviews or studies from before 2010 were excluded (Ollendorf, Chapman and Pearson, 2018). Critics may question the exclusion of ATMPs, which was a difficult decision to make at the very beginning of this PhD. There were certainly arguments for and against, which were equally valid, but the main concern was that very specific issues for ATMPs would dilute or sway results for orphan drugs that were not curative by nature. Especially, because ATMPs were not yet representing the majority of orphan drug designations, which made it a decisive concern (Farkas *et al.*, 2017). For example, one-off costs of curative treatments leading to major affordability issues (Grand *et al.*, 2024b).

The methods used in this PhD present certain limitations, which can all be questioned. Was it the right decision to focus on early HTA or conceptual cost-effectiveness modelling? Would it have been better to focus on specific methods like belief elicitation or EVPI? Why focus on MSA when other diseases are equally worthy?

Ultimately, the aim was to explore issues, challenges and opportunities for economic evaluation in rare disease. It is a question that in my opinion warrants a broad approach. Initially, to identify issues, challenges and opportunities. This was the reasoning for deciding to do a systematic literature review. This decision proved very useful and is typical for a PhD. However, one could have taken other research projects in various directions. However, I decided that the best way forward was to further investigate these initiatives by a disease-specific approach. For example, a demonstration of cost-effectiveness modelling in MSA in Chapter 3 and Chapter 4, each of these studies had multiple limitations that will not be repeated here. One could have decided to conduct additional literature searches on, e.g., EVPI, but generally, it seemed that many opportunities were unexplored in case-studies of rare diseases.

I have previously explained the suitability of MSA as a research candidate, but this disease also presented an opportunity to fill major research gaps which was naturally very appealing. It would have been possible to corroborate results from literature across multiple rare diseases, but one should not underestimate the resources required to conduct a systematic assessment of literature even in a rare disease or for model development where nothing exists. From an academic perspective it was very fulfilling and challenging work to first identify issues, challenges and opportunities in literature and then apply this knowledge in a largely unexplored rare disease like MSA.

5.5.2 What is missing from this PhD

This PhD did not make policy suggestions, but I hope that readers become aware of potential policy implications and their ethical consequences. For example, in the introduction d (Section 1.2) where the consequences of policies are highlighted. The orphan drug act of 1983 is used to illustrate the importance of these incentivising policies. I can reiterate that:

"Health service providers need to reward innovators of medicine because they direct their efforts toward profitable areas. Fail to do so – and they will stop providing the innovations, e.g., orphan drugs" (Towse, 2010).

It is quite simple. Manufacturers of orphan drugs need to be rewarded if they are to continue their research. It means that solutions need to be found for economic evaluations of orphan drugs in rare diseases assuming that this is a political priority. A powerful solution is policy. I do not make suggestions for policies but hope to help decision-makers make informed decision. This PhD focusses on the challenges for researchers.

However, another important point, is that every political decision has consequences and there may be ethical consequences that are important to be aware about. So perhaps – even before policies are discussed, a pertinent question is: what ethical principles govern our decision-making? Do we want to maximise health through a theoretical foundation that is utilitarianism (Dolan, 2001; Coast, 2009)? Or do we want to promote fairness principles, e.g., ethical principles that favour orphan drugs: non-abandonment (rare-disease patients should not be abandoned), rights-based approach (all have the right to high-quality healthcare), and equality of opportunity (all need a fair chance of improving their health) (Zelei *et al.*, 2016)? Or do we want a balanced approach? (Ollendorf, Chapman and Pearson, 2018).

These are critical questions, and it personally bothers me when politicians state that they wish for more orphan drugs to reach patients, when they have not considered these ethical principles and consequently, revised their HTA body's guidelines accordingly. This PhD supports the notion that economic evaluation in rare diseases is notoriously challenging, even with timely planning and large investments into data generating efforts. It is therefore also important to praise and observe those who dare to innovate. For example, an HTA body like NICE that has implemented various initiatives like the HST evaluation that

increases the willingness-to-pay threshold for diseases that are very rare with limited treatment options and where it is difficult to collect evidence (National Institute for Health and Care Excellence, 2024)

5.5.3 What is uncertain about this work

It is uncertain if all challenges, issues and opportunities were captured in the umbrella review since the study focused on systematic reviews. It is plausible that additional issues exist that were not described in the study. What I do know is that orphan drugs classified as ATMPs were not described and a review of them would make for an interesting comparison. Without a comparative review it is unclear if the issues are similar. There might be an overlap as suggested by Gladwell and colleagues, for example, sparse supportive evidence, small sample sizes, insufficient follow-up and single-armed studies (Gladwell *et al.*, 2024). However, it was also suggested to consider economic aspects for curative and non-curative treatments separately (Zelei *et al.*, 2016). Finally, it is highly likely that additional opportunities exist. Health economics is already a mixture of various disciplines, e.g., statistics, operations research, economics etc. It is therefore highly likely that research methods from other disciplines could present opportunities. It is also possible that opportunities exist that were not identified. However, it is certain that those opportunities that were identified need further development.

The cost-effectiveness results reported for MSA are highly uncertain due to the data scarcities. This is both when UMSARS Part IV and functional milestones are used to inform health states, but data scarcities were most pronounced for functional milestones. For example, there was no data to inform a costs and utilities for functional milestones. The focus was on the model development, which was transparently reported including code. The model development in Chapter 3 included data analysis of IPD, which meant that utilities and transition probabilities were calculated. This was novel data that had not been previously reported for MSA.

5.6 Future work

Additional value of information methods could have been applied in Chapter 4 but only EVPI was applied in the analysis and there are naturally reasons why the others were not applied, e.g., time constraints. However, it would make for an interesting research topic, to apply the full range of value of information methods to further understand their potential value in rare diseases, as there are apparent considerations that are pertinent to rare diseases, which to my knowledge, this PhD was the first to outline. As alluded to earlier, for increased robustness, it may be worthwhile also exploring EVPPI for a parameter or sets of parameters, which can help determine which parameters contribute most to uncertainty (Briggs, Sculpher and Claxton, 2006; Heath, Kunst and Jackson, 2024). For example, to further evaluate additional research versus managed-entry agreements, as evidence generation plans may depend on which parameters are causing the uncertainty. Moreover, one can take additional steps and evaluate sample sizes or study designs

if the costs of conducting research are lower than EVPPI. These methods are more complex because their methods depend on study design, model type and expertise, and they are EVSI and ENBS. For example, these methods can help determine study types and their optimal design to reduce uncertainty in decision-analytic models (Heath, Kunst and Jackson, 2024). Generally, research on the implications and relevance of these advanced methods in rare diseases are warranted. From a personal point of view, this would be a highly interesting research project.

Moreover, additional research on disease analogues, e.g., refine the definition and concept, validity, acceptability by payers as opposed to belief elicitation, would also be highly interesting. User-friendly model interfaces in early HTA, and modelling in R, I expect to progress steadily due to very high interest reflected by the surge in tutorials. However, my personality tends to gravitate towards difficult problems and because EVPI is computationally and technically challenging, it would be a natural choice of study if I had three more years for research.

Chapter 6 Conclusions

The main research aim set out to explore issues, challenges and opportunities for economic evaluation in rare disease. Multiple studies and a mixed methods approach were applied to address this research aim. The most obvious finding to emerge from the umbrella review was the scarcity of health-economic parameters. It was, moreover, an issue for cost-effectiveness modelling in MSA where health-economic parameters were scarce.

A limited quantity of economic evaluations was identified in the umbrella review, and the possible causes were publication bias and missing health-economic parameters. In addition, no economic evaluations were identified in MSA, which was possibly due to missing health-economic parameters. It meant that issues and challenges were largely aligned across the research projects. These findings suggest that in general, it is very challenging to conduct economic evaluations of orphan drugs in rare diseases.

What is more, this project was designed to identify opportunities in literature, which were evaluated in a series of cost-effectiveness modelling demonstrations. An opportunity was flexible and comprehensive cost-effectiveness modelling, which was broken down into several research initiatives. For example, early modelling to identify data gaps, preferably in an iterative modelling framework that can guide data collection efforts, e.g., through EVPI. In addition, user-friendly interfaces to include multiple stakeholders were proposed, or disease analogues to reduce data gaps by surrogate inputs. This project has demonstrated, for the first time, the potential of opportunities in economic evaluation of rare diseases, but they should be explored through additional research.

To my knowledge, this study is the very first to systematically assess the full health-economic body of literature and to construct an economic evaluation in MSA. This research had several practical applications: first, it was a vehicle to corroborate identified challenges from the umbrella review; second, it was needed to explore the opportunities identified; lastly it filled an important research gap for MSA.

To conclude, demonstrating the value of orphan drugs in rare diseases is notoriously difficult, because of pertinent challenges such as dealing with data scarcities for health-economic parameters. Solutions are not plentiful and rarely forthcoming; however, opportunities were identified, but they need further development by future researchers.

References

Abdallah, K., Huys, I., Claes, K. and Simoens, S. (2021) 'Methodological Quality Assessment of Budget Impact Analyses for Orphan Drugs: A Systematic Review', *Frontiers in Pharmacology,* 12 (no pagination).

Acosta, C., Gianinazzi, M., Dort, T., Armstrong, N., Ryder, S., Lundqvist, T., Ekelund, M. and Lycke, J. (2021) 'Modeling the cost-effectiveness of prolonged-release fampridine for the treatment of walking impairment in patients with multiple sclerosis in Sweden', *Journal of Medical Economics,* 24(1), pp. 770–780.

Alarid-Escudero, F., Krijkamp, E., Enns, E. A., Yang, A., Hunink, M. G. M., Pechlivanoglou, P. and Jalal, H. (2023) 'A Tutorial on Time-Dependent Cohort State-Transition Models in R Using a Cost-Effectiveness Analysis Example', *Med Decis Making,* 43(1), pp. 21–41.

Alarid-Escudero, F., Krijkamp, E. M., Pechlivanoglou, P., Jalal, H., Kao, S. Z., Yang, A. and Enns, E. A. (2019) 'A Need for Change! A Coding Framework for Improving Transparency in Decision Modeling', *Pharmacoeconomics,* 37(11), pp. 1329–1339.

Ali, S., Haines, A. and Lee, K. 'Severity Weighting of Quality-Adjusted Life-Years for Economic Evaluation: What, How, and Where Next?', *Canadian Journal of Health Technologies*.

Allaire, J., Xie, Y., Dervieux, C., McPherson, J., Luraschi, J., Ushey, K. and Atkins, A. (2023) *Rmarkdown: Dynamic documents for r*. Available at: https://github.com/rstudio/rmarkdown (Accessed: 3 July 2025).

Alsulamy, N., Lee, A., Thokala, P. and Alessa, T. (2020) 'What Influences the Implementation of Shared Decision Making: An Umbrella Review', *Patient Education and Counseling,* 103(12), pp. 2400–2407.

Angelis, A., Naci, H. and Hackshaw, A. (2020) 'Recalibrating Health Technology Assessment Methods for Cell and Gene Therapies', *Pharmacoeconomics,* 38(12), pp. 1297–1308.

Angelis, A., Tordrup, D. and Kanavos, P. (2015) 'Socio-economic burden of rare diseases: A systematic review of cost of illness evidence', *Health Policy,* 119(7), pp. 964–79.

Ara, R. and Brazier, J. (2017) 'Estimating Health State Utility Values for Comorbidities', *PharmacoEconomics,* 35(1), pp. 89–94.

Arno, P. S., Bonuck, K. and Davis, M. (1995) 'Rare Diseases, Drug Development, and AIDS: The Impact of the Orphan Drug Act', *The Milbank Quarterly,* 73(2), pp. 231–252.

Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H. and Tungpunkom, P. (2015) 'Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach', *Int J Evid Based Healthc,,* 13(3), pp. 132–40.

Auguste, P., Colquitt, J., Connock, M., Loveman, E., Court, R., Ciccarelli, O., Counsell, C. and Armoiry, X. (2020) 'Ocrelizumab for Treating Patients with Primary Progressive Multiple Sclerosis: An Evidence Review Group Perspective of a NICE Single Technology Appraisal', *PharmacoEconomics,* 38(6), pp. 527–536.

Australian Government (2018) *Orphan drug designation*. Available at: https://www.tga.gov.au/resources/resource/guidance/orphan-drug-designation (Accessed: 6 December 2023).

Australian Government (2021) *Orphan drug designation eligibility criteria*. Available at: https://www.tga.gov.au/resources/resource/guidance/orphan-drug-designation-eligibility-criteria (Accessed: 6 December 2023).

Ayati, N., Fleifel, L., Sharifi, S., Sahraian, M. A. and Nikfar, S. (2021a) 'Cladribine tablets are a cost-effective strategy in high-disease activity relapsing multiple sclerosis patients in Iran', *Current Journal of Neurology,* 20(3), pp. 146–153.

Ayati, N., Taheri, S., Sahraian, M. A. and Nikfar, S. (2021b) 'Cost-effectiveness of ocrelizumab for treatment of Iranian patients with relapsing multiple sclerosis', *Current Journal of Neurology,* 20(3), pp. 154–161.

Babac, A., Damm, K. and Graf von der Schulenburg, J. M. (2019) 'Patient-reported data informing early benefit assessment of rare diseases in Germany: A systematic review', *Health Economics Review,* 9(1), pp. 34.

Baharnoori, M., Bhan, V., Clift, F., Thomas, K., Mouallif, S., Adlard, N., Cooney, P., Blanchette, F., Patel, B. P. and Grima, D. (2022) 'Cost-Effectiveness Analysis of Ofatumumab for the Treatment of Relapsing-Remitting Multiple Sclerosis in Canada', *PharmacoEconomics - Open,* 6(6), pp. 859–870.

Baran-Kooiker, A., Czech, M. and Kooiker, C. (2018) 'Multi-Criteria Decision Analysis (MCDA) Models in Health Technology Assessment of Orphan Drugs-a Systematic Literature Review. Next Steps in Methodology Development?', *Frontiers in Public Health,* 6, pp. 287.

Barton, G. R., Briggs, A. H. and Fenwick, E. A. (2008) 'Optimal cost-effectiveness decisions: the role of the cost-effectiveness acceptability curve (CEAC), the cost-effectiveness acceptability frontier (CEAF), and the expected value of perfection information (EVPI)', *Value Health,* 11(5), pp. 886–97.

Ben-Shlomo, Y., Wenning, G. K., Tison, F. and Quinn, N. P. (1997) 'Survival of patients with pathologically proven multiple system atrophy', *Neurology,* 48(2), pp. 384.

Ben, Â. J., van Dongen, J. M., El Alili, M., Esser, J. L., Broulíková, H. M. and Bosmans, J. E. (2023) 'Conducting Trial-Based Economic Evaluations Using R: A Tutorial', *Pharmacoeconomics,* 41(11), pp. 1403–1413.

Bentham, J. (1823) 'An introduction to the principles of morals and legislation.-New. ed. rev'.

Bjornsdottir, A., Gudmundsson, G., Blondal, H. and Olafsson, E. (2013) 'Incidence and prevalence of multiple system atrophy: a nationwide study in Iceland', *J Neurol Neurosurg Psychiatry,* 84(2), pp. 136–40.

Bolanos-Diaz, R., Sanabria-Montanez, C. and Custodio-Capunay, N. (2019) 'Cost-effectiveness of interferon therapy for multiple sclerosis in Peru: impact of treatment adherence', *Journal of Pharmaceutical Health Services Research,* 10(1), pp. 125–132.

Booth, A. (2006) 'Clear and present questions: formulating questions for evidence based practice', *Library Hi Tech,* 24(3), pp. 355–368.

Booth, A., Sutton, A. and Papaioannou, D. (2016) 'Searching the literature', *Systematic Approaches to a Successful Literature Review*: Sage, pp. 110.

Bouvy, J. (2024) *Should NICE's cost-effectiveness thresholds change?* Available at: https://www.nice.org.uk/news/blogs/should-nice-s-cost-effectiveness-thresholds-change- (Accessed: 9 January 2024).

Bower, J. H., Maraganore, D. M., McDonnell, S. K. and Rocca, W. A. (1997) 'Incidence of progressive supranuclear palsy and multiple system atrophy in Olmsted County, Minnesota, 1976 to 1990', *Neurology,* 49(5), pp. 1284–1288.

Bramer, W. M., Rethlefsen, M. L., Kleijnen, J. and Franco, O. H. (2017) 'Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study', *Systematic Reviews,* 6(1), pp. 245.

Brent, R. J. (2023) 'Cost-Benefit Analysis versus Cost-Effectiveness Analysis from a Societal Perspective in Healthcare', *International Journal of Environmental Research and Public Health,* 20(5), pp. 4637.

Briggs, A. and Sculpher, M. (1998) 'An Introduction to Markov Modelling for Economic Evaluation', *PharmacoEconomics,* 13(4), pp. 397–409.

Briggs, A., Sculpher, M. and Claxton, K. (2006) *Decision Modelling for Health Economic Evaluation.* Oxford University Press.

Briggs, A. H., Weinstein, M. C., Fenwick, E. A. L., Karnon, J., Sculpher, M. J. and Paltiel, A. D. (2012) 'Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6', *Value in Health,* 15(6), pp. 835–842.

Burns, M. R. and McFarland, N. R. (2020) 'Current Management and Emerging Therapies in Multiple System Atrophy', *Neurotherapeutics,* 17(4), pp. 1582–1602.

CADTH (2022a) *Economic Evaluations & Models - Embase. In: CADTH Search Filters Database.* Available at: https://searchfilters.cadth.ca/link/15 (Accessed: 23 November 2022).

CADTH (2022b) *Economic Evaluations & Models - MEDLINE. In: CADTH Search Filters Database.* Available at: https://searchfilters.cadth.ca/link/16 (Accessed: 17 November 2022).

CADTH (2023a) *Economic - Health Utilities / Quality of Life - MEDLINE*. Available at: https://searchfilters.cda-amc.ca/link/19 (Accessed: 1 May 2024).

CADTH (2023b) *Economic - Health Utilities / Quality of Life - Standard - Embase*. Available at: https://searchfilters.cda-amc.ca/link/18 (Accessed: 1 May 2024).

CADTH (2024) *Economic - Narrow - MEDLINE, Embase*. Available at: https://searchfilters.cda-amc.ca/link/20 (Accessed: 10 May 2024).

Cao, B., Zhang, L., Zou, Y., Wei, Q., Ou, R., Chen, Y. and Shang, H. F. (2018) 'Survival analysis and prognostic nomogram model for multiple system atrophy', *Parkinsonism and Related Disorders,* 54, pp. 68–73.

Carlson, J. J., Chen, S. and Garrison, L. P., Jr. (2017) 'Performance-Based Risk-Sharing Arrangements: An Updated International Review', *Pharmacoeconomics,* 35(10), pp. 1063–1072.

Carlson, J. J., Sullivan, S. D., Garrison, L. P., Neumann, P. J. and Veenstra, D. L. (2010) 'Linking payment to health outcomes: A taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers', *Health Policy,* 96(3), pp. 179–190.

Caro, J. J., Briggs, A. H., Siebert, U. and Kuntz, K. M. (2012) 'Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1', *Med Decis Making,* 32(5), pp. 667–77.

Chancellor, J. V., Hill, A. M., Sabin, C. A., Simpson, K. N. and Youle, M. (1997) 'Modelling the Cost Effectiveness of Lamivudine/Zidovudine Combination Therapy in HIV Infection', *PharmacoEconomics,* 12(1), pp. 54–66.

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A. and Borges, B. (2023) *shiny: Web Application Framework for R*. Available at: https://shiny.rstudio.com/ (Accessed: 1 March 2023).

Chaudhary, A. and Kumar, V. (2025) 'Rare diseases: a comprehensive literature review and future directions', *Journal of Rare Diseases,* 4(1), pp. 33.

Chaudhuri, K. R., Pickard, A. S., Alobaidi, A., Jalundhwala, Y. J., Kandukuri, P. L., Bao, Y., Sus, J., Jones, G., Ridley, C., Oddsdottir, J., Najle-Rahim, S., Madin-Warburton, M., Xu, W. and Schrag, A. (2022) 'The Cost Effectiveness of Levodopa-Carbidopa Intestinal Gel in the Treatment of Advanced Parkinson's Disease in England', *Pharmacoeconomics,* 40(5), pp. 559–574.

Chen, T., Chen, J., Chen, C., Zheng, H., Chen, Y., Liu, M. and Zheng, B. (2021) 'Systematic review and cost-effectiveness of bosentan and sildenafil as therapeutic drugs for pediatric pulmonary arterial hypertension', *Pediatric Pulmonology,* 56(7), pp. 2250–2258.

Cheng, M. M., Ramsey, S. D., Devine, E. B., Garrison, L. P., Bresnahan, B. W. and Veenstra, D. L. (2012) 'Systematic review of comparative effectiveness data for oncology orphan drugs', *American Journal of Managed Care,* 18(1), pp. 47–62.

Chiò, A., Magnani, C. and Schiffer, D. (1998) 'Prevalence of Parkinson's disease in Northwestern Italy: comparison of tracer methodology and clinical ascertainment of cases', *Mov Disord,* 13(3), pp. 400–5.

Chrysostome, V., Tison, F., Yekhlef, F., Sourgen, C., Baldi, I. and Dartigues, J. F. (2004) 'Epidemiology of multiple system atrophy: a prevalence and pilot risk factor study in Aquitaine, France', *Neuroepidemiology,* 23(4), pp. 201–8.

Chung, B. H. Y., Chau, J. F. T. and Wong, G. K. (2021) 'Rare versus common diseases: a false dichotomy in precision medicine', *NPJ Genom Med,* 6(1), pp. 19.

Clarke, S., Ellis, M. and Brownrigg, J. (2021) 'The impact of rarity in NICE's health technology appraisals', *Orphanet Journal of Rare Diseases,* 16(1), pp. 218.

Coast, J. (2009) 'Maximisation in extra-welfarism: A critique of the current position in health economics', *Soc Sci Med,* 69(5), pp. 786–92.

Congress (1994) *H.R.4865 - Orphan Drug Act Amendments of 1994. 103rd Congress (1993-1994)*. Available at: https://www.congress.gov/bill/103rd-congress/house-bill/4865/text (Accessed: 30 June 2023).

Congress (2021) *H.R.1629 - Fairness in Orphan Drug Exclusivity Act. 117th Congress (2021-2022)*. Available at: https://www.congress.gov/bill/117th-congress/house-bill/1629 (Accessed: 30 June 2023).

Connected Papers (2022) *Explore connected papers in a visual graph*. Available at: https://www.connectedpapers.com/ (Accessed: 13 January 2023).

Cortesi, P. A., Antonazzo, I. C., Gasperini, C., Nica, M., Ritrovato, D. and Mantovani, L. G. (2022) 'Cost-effectiveness and budget impact analysis of siponimod in the treatment of secondary progressive multiple sclerosis in Italy', *PLoS ONE,* 17(3 March), pp. e0264123.

Coyle, D., Durand-Zaleski, I., Farrington, J., Garrison, L., Graf von der Schulenburg, J. M., Greiner, W., Longworth, L., Meunier, A., Moutie, A. S., Palmer, S., Pemberton-Whiteley, Z., Ratcliffe, M., Shen, J., Sproule, D., Zhao, K. and Shah, K. (2020) 'HTA methodology and value frameworks for evaluation and policy making for cell and gene therapies', *Eur J Health Econ,* 21(9), pp. 1421–1437.

Dams, J., Zapp, J.-J. and König, H.-H. (2023) 'Modelling the Cost Effectiveness of Treatments for Parkinson's Disease: An Updated Methodological Review', *PharmacoEconomics,* 41(10), pp. 1205–1228.

Daroudi, R., Mousavi, M., Shirazikhah, M., Alizadeh Zarei, M., Hendi, H., Joghataei, F. and Darvishi, A. (2023) 'Cost-utility analysis of multiple sclerosis rehabilitation in Iran', *Expert Review of Pharmacoeconomics and Outcomes Research,* 23(10), pp. 1129–1137.

De Freitas, L., Goodacre, S., O'Hara, R., Thokala, P. and Hariharan, S. (2018) 'Interventions to improve patient flow in emergency departments: an umbrella review', *Emergency Medicine Journal,* 35(10), pp. 626–637.

Denis, A., Simoens, S., Fostier, C., Mergaert, L. and Cleemput, I. (2009) *Policies for rare diseases and orphan drugs*: Brussels: Belgian Health Care Knowledge Centre (KCE)KCE Reports 112C). Available at: https://kce.fgov.be/sites/default/files/2021-11/d20091027332.pdf (Accessed: 9 November 2022).

Dhawan, A., Lawlor, M. W., Mazariegos, G. V., McKiernan, P., Squires, J. E., Strauss, K. A., Gupta, D., James, E. and Prasad, S. (2019) 'Disease burden of Crigler-Najjar syndrome: Systematic review and future perspectives', *Journal of Gastroenterology & Hepatology,* 35(4), pp. 530–543.

Dimitrova, M., Seitaridou, Y., Lazarova, R., Petrova, G., Mitov, K., Milanov, I., Marinov, L., Kamusheva, M. and Panayotov, P. (2023) 'COST-EFFECTIVENESS OF DISEASE-MODIFYING TREATMENTS FOR MULTIPLE SCLEROSIS IN BULGARIA BASED ON EVIDENCE FROM REAL WORLD SETTINGS', *Farmacia,* 71(6), pp. 1324–1332.

Dolan, P. (2001) 'Utilitarianism and the measurement and aggregation of quality--adjusted life years', *Health Care Anal,* 9(1), pp. 65–76.

Drummond, M. (2015) 'When do performance-based risk-sharing arrangements make sense?', *The European Journal of Health Economics,* 16(6), pp. 569–571.

Drummond, M. F., Wilson, D. A., Kanavos, P., Ubel, P. and Rovira, J. (2007) 'Assessing the economic challenges posed by orphan drugs', *Int J Technol Assess Health Care,* 23(1), pp. 36–42.

Eddy, D. M., Hollingworth, W., Caro, J. J., Tsevat, J., McDonald, K. M. and Wong, J. B. (2012) 'Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7', *Med Decis Making,* 32(5), pp. 733–43.

Eichler, H.-G., Kossmeier, M., Zeitlinger, M. and Schwarzer-Daum, B. (2023) 'Orphan drugs' clinical uncertainty and prices: Addressing allocative and technical inefficiencies in orphan drug reimbursement', *Frontiers in Pharmacology,* Volume 14 - 2023.

European Comission (2021) *Horizon Europe. The EU research and innovation programme 2021 - 27*. Available at: https://research-and-innovation.ec.europa.eu/system/files/2022-06/ec_rtd_he-investing-to-shape-our-future_0.pdf (Accessed: 13 December 2022).

European Comission (2022) *Research and innovation*. Available at: https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe_en (Accessed: 13 Decenber 2022).

European Medicines Agency (2015) *Reflection paper on classification of advanced therapy medicinal products*. Available at: https://www.ema.europa.eu/system/files/documents/scientific-guideline/wc500187744_en.pdf (Accessed: 5 December 2023).

European Medicines Agency (2018) *Rare diseases, orphan medicines. Getting the facts straight*. Available at: https://www.ema.europa.eu/en/documents/other/rare-diseases-orphan-medicines-getting-facts-straight_en.pdf (Accessed: 16 December 2022).

European Medicines Agency (2020) *Benefit-risk methodology.* Available at: https://www.ema.europa.eu/en/about-us/what-we-do/regulatory-science-research/benefit-risk-methodology (Accessed: 25 October 2022).

European Medicines Agency (2022a) *Authorisation of medicines*. Available at: https://www.ema.europa.eu/en/about-us/what-we-do/authorisation-medicines (Accessed: 13 December 2022).

European Medicines Agency (2022b) *Orphan Incentives*. Available at: https://www.ema.europa.eu/en/human-regulatory/research-development/orphan-designation/orphan-incentives (Accessed: 20 October 2022).

European Medicines Agency (2022c) *Orphan Medicinal Product Designation: Overview 2000-2001*: Orphan Medicines - Scientific Evidence Generation. Available at: https://www.ema.europa.eu/en/documents/other/orphan-medicines-figures-2000-2021_en.pdf (Accessed: 20 October 2022).

Facey, K. M., Espin, J., Kent, E., Link, A., Nicod, E., O'Leary, A., Xoxi, E., van de Vijver, I., Zaremba, A., Benisheva, T., Vagoras, A. and Upadhyaya, S. (2021) 'Implementing Outcomes-Based Managed Entry Agreements for Rare Disease Treatments: Nusinersen and Tisagenlecleucel', *Pharmacoeconomics,* 39(9), pp. 1021–1044.

Farkas, A. M., Mariz, S., Stoyanova-Beninska, V., Celis, P., Vamvakas, S., Larsson, K. and Sepodes, B. (2017) 'Advanced Therapy Medicinal Products for Rare Diseases: State of Play of Incentives Supporting Development in Europe', *Front Med (Lausanne),* 4, pp. 53.

Faulkner, E., Spinner, D. S., Ringo, M. and Carroll, M. (2019) 'Are Global Health Systems Ready for Transformative Therapies?', *Value in Health,* 22(6), pp. 627–641.

Felzmann, H. (2017) 'Utilitarianism as an approach to ethical decision making in health care', *Key concepts and issues in nursing ethics*: Springer, pp. 29–41.

Ferrario, A. and Kanavos, P. (2015) 'Dealing with uncertainty and high prices of new medicines: A comparative analysis of the use of managed entry agreements in Belgium, England, the Netherlands and Sweden', *Social Science & Medicine,* 124, pp. 39–47.

Fiorenza, S., Ritchie, D. S., Ramsey, S. D., Turtle, C. J. and Roth, J. A. (2020) 'Value and affordability of CAR T-cell therapy in the United States', *Bone Marrow Transplant,* 55(9), pp. 1706–1715.

Flabeau, O., Meissner, W. G. and Tison, F. (2010) 'Multiple system atrophy: current and future approaches to management', *Ther Adv Neurol Disord,* 3(4), pp. 249–63.

Food and Drug Administration (2020) *Frequently asked questions on patents and exclusivity*. Available at: https://www.fda.gov/drugs/development-approval-process-drugs/frequently-asked-questions-patents-and-exclusivity#:~:text=Patents%20can%20be%20issued%20or,have%20just%20one%20or%20neither. (Accessed: 30 June 2023).

Foubert-Samier, A., Pavy-Le Traon, A., Guillet, F., Le-Goff, M., Helmer, C., Tison, F., Rascol, O., Proust-Lima, C. and Meissner, W. G. (2020) 'Disease progression and prognostic factors in multiple system atrophy: A prospective cohort study', *Neurobiology of disease,* 139, pp. 104813.

Fowler, C. (2011) *MSA Trust Research Strategy*. Available at: https://www.msatrust.org.uk/wp-content/uploads/2015/11/MSA-Trust-Research-Strategy.pdf (Accessed: 9 January 2024).

Fox, J. and Bouchet-Valat, M. (2022) *Rcmdr: R Commander. R package version 2.8-0*. Available at: https://socialsciences.mcmaster.ca/jfox/Misc/Rcmdr/ (Accessed: 16 March 2023).

Friedmann, C., Levy, P., Hensel, P. and Hiligsmann, M. (2018) 'Using multi-criteria decision analysis to appraise orphan drugs: a systematic review', *Expert Review of Pharmacoeconomics & Outcomes Research,* 18(2), pp. 135–146.

Furneri, G., Santoni, L., Ricella, C. and Prosperini, L. (2019) 'Cost-effectiveness analysis of escalating to natalizumab or switching among immunomodulators in relapsing-remitting multiple sclerosis in Italy', *BMC health services research,* 19(1), pp. 436.

Gajdos, O., Rozanek, M., Donin, G. and Kamensky, V. (2021) 'Cost-Utility Analysis of Home Mechanical Ventilation in Patients with Amyotrophic Lateral Sclerosis', *Healthcare (Basel, Switzerland),* 9(2).

Gallehzan, N. A., Khosravi, M., Jamebozorgi, K., Mir, N., Jalilian, H., Soleimanpour, S., Hoseini, S., Rezapour, A. and Eshraghi, A. (2024) 'Cost-utility and cost-effectiveness analysis of disease-modifying drugs of relapsing–remitting multiple sclerosis: a systematic review', *Health Economics Review,* 14(1), pp. 12.

Gallop, K., Pham, N., Maclaine, G., Saunders, E., Black, B., Hubig, L. and Acaster, S. (2023) 'Health-related quality-of-life and burden for caregivers of individuals with neurogenic orthostatic hypotension', *Neurodegenerative disease management,* 13(1), pp. 35–45.

Gammie, T., Lu, C. Y. and Babar, Z. U. (2015) 'Access to Orphan Drugs: A Comprehensive Review of Legislations, Regulations and Policies in 35 Countries', *PLoS One,* 10(10), pp. e0140002.

Garcia-Perez, L., Linertova, R., Valcarcel-Nazco, C., Posada, M., Gorostiza, I. and Serrano-Aguilar, P. (2021) 'Cost-of-illness studies in rare diseases: a scoping review', *Orphanet Journal Of Rare Diseases,* 16(1), pp. 178.

Giannuzzi, V., Conte, R., Landi, A., Ottomano, S. A., Bonifazi, D., Baiardi, P., Bonifazi, F. and Ceci, A. (2017) 'Orphan medicinal products in Europe and United States to cover needs of patients with rare diseases: an increased common effort is to be foreseen', *Orphanet J Rare Dis,* 12(1), pp. 64.

Gilman, S., Low, P. A., Quinn, N., Albanese, A., Ben-Shlomo, Y., Fowler, C. J., Kaufmann, H., Klockgether, T., Lang, A. E., Lantos, P. L., Litvan, I., Mathias, C. J., Oliver, E., Robertson, D., Schatz, I. and Wenning, G. K. (1999) 'Consensus statement on the diagnosis of multiple system atrophy', *J Neurol Sci,* 163(1), pp. 94–8.

Gilman, S., Wenning, G. K., Low, P. A., Brooks, D. J., Mathias, C. J., Trojanowski, J. Q., Wood, N. W., Colosimo, C., Dürr, A., Fowler, C. J., Kaufmann, H., Klockgether, T., Lees, A., Poewe, W., Quinn, N., Revesz, T., Robertson, D., Sandroni, P., Seppi, K. and Vidailhet, M. (2008) 'Second consensus statement on the diagnosis of multiple system atrophy', *Neurology,* 71(9), pp. 670.

Ginestal, R., Rubio-Terres, C., Moran, O. D., Rubio-Rodriguez, D., De Los Santos, H., Ordonez, C. and Sanchez-Magro, I. (2023) 'Cost-effectiveness of cladribine tablets and dimethyl fumarate in the treatment of relapsing remitting multiple sclerosis in Spain', *Journal of Comparative Effectiveness Research,* 12(2), pp. e220193.

Giovannoni, G., Brex, P. A., Dhiraj, D., Fullarton, J., Freddi, M., Rodgers-Gray, B. and Schmierer, K. (2019) 'Glatiramer acetate as a clinically and cost-effective treatment of relapsing multiple sclerosis over 10 years of use within the National Health Service: Final results from the UK Risk Sharing Scheme', *Multiple Sclerosis Journal - Experimental, Translational and Clinical,* 5(4).

Gladwell, D., Ciani, O., Parnaby, A. and Palmer, S. (2024) 'Surrogacy and the Valuation of ATMPs: Taking Our Place in the Evidence Generation/Assessment Continuum', *PharmacoEconomics,* 42(2), pp. 137–144.

Golden, E. P., McCreary, M. and Vernino, S. (2022) 'Responsiveness of UMSARS and other clinical measures in a longitudinal structured care clinic for multiple system atrophy', *Clinical Autonomic Research,* 32(6), pp. 477–484.

Grand, T., Shijie, R., Thokala, P., Åström, D., Regnier, S. and Hall, J. (2023) 'Issues, challenges and opportunities for economic evaluation of orphan drugs: an umbrella review protocol'.

Grand, T., Thokala, P., Ren, K., Regnier, S., Oudin Åström, D., Awad, S., Bidani, A. and Krismer, P. (2024a) *Use of Disease Analogues to Inform Health-Economic Parameters for Economic Evaluations of Rare Diseases: A Case-Study of Multiple System Atrophy*. ISPOR Europe 2024. Barcelona, Spain: Value in Health. Available at: https://www.ispor.org/heor-resources/presentations-database/presentation/euro2024-4014/144187 (Accessed: 7 July 2024).

Grand, T. S., Ren, S., Hall, J., Åström, D. O., Regnier, S. and Thokala, P. (2024b) 'Issues, Challenges and Opportunities for Economic Evaluations of Orphan Drugs in Rare Diseases: An Umbrella Review', *PharmacoEconomics*.

Grand, T. S., Ren, S., Thokala, P., Zanigni, S., Åström, D. O., Bidani, A., Höglinger, G. U., Poewe, W., Krismer, F. and Regnier, S. (2025) 'Cost-effectiveness modelling of multiple system atrophy for early health technology assessment', *PharmacoEconomics Open*.

Gray, A., Clarke, P., Wolstenholme, J. and Wordsworth, S. (2011) *Applied methods of cost-effectiveness analysis in health care.* Oxford University Press.

Green, N., Lamrock, F., Naylor, N., Williams, J. and Briggs, A. (2023) 'Health Economic Evaluation Using Markov Models in R for Microsoft Excel Users: A Tutorial', *Pharmacoeconomics,* 41(1), pp. 5–19.

Grossauer, A., Sidoroff, V., Heim, B. and Seppi, K. (2022) 'Symptomatic Care in Multiple System Atrophy: State of the Art', *The Cerebellum*.

Grossman, D. B. (1984) 'The Orphan Drug Act: Adoption or Foster Care?', *Food, Drug, Cosmetic Law Journal,* 39(2), pp. 128–151.

Gruhn, S., Witte, J., Greiner, W., Damm, O., Dietzsch, M., Kramer, R. and Knuf, M. (2022) 'Epidemiology and economic burden of meningococcal disease in Germany: A systematic review', *Vaccine,* 40(13), pp. 1932–1947.

Guyot, P., Ades, A. E., Ouwens, M. J. N. M. and Welton, N. J. (2012) 'Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves', *BMC Medical Research Methodology,* 12(1), pp. 9.

Guzauskas, G. F., Tabrizi, S. J., Long, J. D., Arnesen, A., Hamilton, J. L., Claassen, D. O., Munetsi, L. R., Malik, S., Rodríguez-Santana, I., Ali, T. M. and Zhang, F. (2024) 'Long-Term Health Outcomes of Huntington Disease and the Impact of Future Disease-Modifying Treatments: A Decision-Modeling Analysis', *Neurol Clin Pract,* 14(5), pp. e200340.

Haddaway, N. R., Page, M. J., Pritchard, C. C. and McGuinness, L. A. (2022) 'PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis', *Campbell Systematic Reviews,* 18(2), pp. e1230.

Hart, R., Burns, D., Ramaekers, B., Ren, S., Gladwell, D., Sullivan, W., Davison, N., Saunders, O., Sly, I., Cain, T. and Lee, D. (2020) 'R and Shiny for Cost-Effectiveness Analyses: Why and When? A Hypothetical Case Study', *PharmacoEconomics,* 38(7), pp. 765–776.

Heath, A., Kunst, N. and Jackson, C. (2024) *Value of Information for Healthcare Decision-Making.* CRC Press.

Holman, B., Bero, L. and Mintzes, B. (2019) *Industry Sponsorship Bias.* Catalogue of Bias Collaboration. Available at: https://catalogofbias.org/biases/industry-sponsorship-bias/ (Accessed: 24 June 2023).

Hua, L. H., Hersh, C. M., Morten, P., Kusel, J., Lin, F., Cave, J., Varga, S., Herrera, V. and Ko, J. J. (2019) 'The impact of price reductions after loss of exclusivity in a cost-effectiveness analysis: Fingolimod versus interferon beta-1a for the treatment of relapsing-remitting multiple sclerosis', *Journal of Managed Care and Specialty Pharmacy,* 25(4), pp. 490–498b.

Hughes-Wilson, W., Palma, A., Schuurman, A. and Simoens, S. (2012) 'Paying for the Orphan Drug System: break or bend? Is it time for a new evaluation system for payers in Europe to take account of new rare disease treatments?', *Orphanet J Rare Dis,* 7, pp. 74.

Hughes, D. A., Tunnage, B. and Yeo, S. T. (2005) 'Drugs for exceptionally rare diseases: do they deserve special status for funding?', *Qjm,* 98(11), pp. 829–36.

Hunter, S. F., Bindra, J., Chopra, I., Niewoehner, J., Panaccio, M. P. and Wan, G. J. (2021) 'Cost-effectiveness of repository corticotropin injection for the treatment of acute exacerbations in multiple sclerosis', *ClinicoEconomics and Outcomes Research,* 13, pp. 883–892.

Huygens, S. and Versteegh, M. (2021) 'Modeling the Cost-Utility of Treatment Sequences for Multiple Sclerosis', *Value in Health,* 24(11), pp. 1612–1619.

Hvidberg, M. F. and Alava, M. H. (2023) 'Catalogues of EQ-5D-3L Health-Related Quality of Life Scores for 199 Chronic Conditions and Health Risks for Use in the UK and the USA', *PharmacoEconomics,* 41(10), pp. 1287–1388.

ICD10Data (2024) *Other functional intestinal disorders*. Available at: https://www.icd10data.com/ICD10CM/Codes/K00-K95/K55-K64/K59-.

Ijzerman, M. J., Koffijberg, H., Fenwick, E. and Krahn, M. (2017) 'Emerging Use of Early Health Technology Assessment in Medical Product Development: A Scoping Review of the Literature', *PharmacoEconomics,* 35(7), pp. 727–740.

Incerti, D. and Jansen, J. (2021) *hesim: Health Economic Simulation Modeling and Decision Analysis.* Available at: https://arxiv.org/abs/2102.09437. (Accessed: 3 July 2025).

Incerti, D., Jansen, J. P. and Clements, M. (2025) *Partitioned survival models*. Available at: https://hesim-dev.github.io/hesim/articles/psm.html (Accessed: 15 Jan 2025).

Innovation Fund Denmark (2025) *Innovation Fund Denmark*. Available at: https://innovationsfonden.dk/en/p/industrial-researcher (Accessed: November 6 2025).

Jackson, C. (2011) 'Multi-State Models for Panel Data: The msm Package for R', *Journal of Statistical Software,* 38(8), pp. 1 – 28.

Jakes, R. W., Kwon, N., Nordstrom, B., Goulding, R., Fahrbach, K., Tarpey, J. and Van Dyke, M. K. (2021) 'Burden of illness associated with eosinophilic granulomatosis with polyangiitis: a systematic literature review and meta-analysis', *Clinical Rheumatology,* 40(12), pp. 4829–4836.

Jansen, J. P., Incerti, D. and Linthicum, M. T. (2019) 'Developing Open-Source Models for the US Health System: Practical Experiences and Challenges to Date with the Open-Source Value Project', *Pharmacoeconomics,* 37(11), pp. 1313–1320.

Jecmenica-Lukic, M., Petrovic, I. N., Pekmezovic, T., Tomic, A., Stankovic, I., Svetel, M. and Kostic, V. S. (2021) 'The profile and evolution of neuropsychiatric symptoms in multiple system atrophy: Self-and caregiver report', *Journal of Neuropsychiatry and Clinical Neurosciences,* 33(2), pp. 124–131.

Jecmenica-Lukic, M. V., Pekmezovic, T. D., Petrovic, I. N., Dragasevic, N. T. and Kostic, V. S. (2018) 'Factors associated with deterioration of health-related quality of life in multiple system atrophy: 1-year follow-up study', *Acta Neurologica Belgica,* 118(4), pp. 589–595.

Kalabina, S., Belsey, J., Pivonka, D., Mohamed, B., Thomas, C. and Paterson, B. (2019) 'Cost-utility analysis of levodopa carbidopa intestinal gel (Duodopa) in the treatment of advanced Parkinson's disease in patients in Scotland and Wales', *Journal of Medical Economics,* 22(3), pp. 215–225.

Kanters, T. A., de Sonneville-Koedoot, C., Redekop, W. K. and Hakkaart, L. (2013) 'Systematic review of available evidence on 11 high-priced inpatient orphan drugs', *Orphanet Journal Of Rare Diseases,* 8, pp. 124.

Kaplan, S. (2024) 'Prevalence of multiple system atrophy: A literature review', *Revue Neurologique*.

Kaplan, S., Xie, H. and Kumar, A. (2025) 'Diagnostic symptoms, treatment patterns, and healthcare resource utilization in multiple system atrophy in the United States', *Neurodegenerative Disease Management*, pp. 1–10.

Katsigianni, E. I. and Petrou, P. (2022) 'A systematic review of economic evaluations of enzyme replacement therapy in Lysosomal storage diseases', *Cost Effectiveness & Resource Allocation,* 20(1), pp. 51.

Knoble, N., Nayroles, G., Cheng, C. and Arnould, B. (2018) 'Illustration of patient-reported outcome challenges and solutions in rare diseases: a systematic review in Cushing's syndrome', *Orphanet Journal Of Rare Diseases,* 13(1), pp. 228.

Kollensperger, M., Stampfer-Kountchev, M., Seppi, K., Geser, F., Frick, C., Del Sorbo, F., Albanese, A., Gurevich, T., Giladi, N., Djaldetti, R., Schrag, A., Low, P. A., Mathias, C. J., Poewe, W. and Wenning, G. K. (2007) 'Progression of dysautonomia in multiple system atrophy: A prospective study of self-perceived impairment', *European Journal of Neurology,* 14(1), pp. 66–72.

Kremer, I. E. H., Hiligsmann, M., Carlson, J., Zimmermann, M., Jongen, P. J., Evers, S. M. A. A., Petersohn, S., Pouwels, X. G. L. V. and Bansback, N. (2020) 'Exploring the Cost Effectiveness of Shared Decision Making for Choosing between Disease-Modifying Drugs for Relapsing-Remitting Multiple Sclerosis in the Netherlands: A State Transition Model', *Medical decision making : an international journal of the Society for Medical Decision Making,* 40(8), pp. 1003–1019.

Krijkamp, E. M., Alarid-Escudero, F., Enns, E. A., Jalal, H. J., Hunink, M. G. M. and Pechlivanoglou, P. (2018) 'Microsimulation Modeling for Health Decision Sciences Using R: A Tutorial', *Med Decis Making,* 38(3), pp. 400–422.

Krijkamp, E. M., Alarid-Escudero, F., Enns, E. A., Pechlivanoglou, P., Hunink, M. G. M., Yang, A. and Jalal, H. J. (2020) 'A Multidimensional Array Representation of State-Transition Model Dynamics', *Med Decis Making,* 40(2), pp. 242–248.

Krismer, F., Fanciulli, A., Meissner, W. G., Coon, E. A. and Wenning, G. K. (2024) 'Multiple system atrophy: advances in pathophysiology, diagnosis, and treatment', *Lancet Neurol,* 23(12), pp. 1252–1266.

Krismer, F. and Wenning, G. K. (2017) 'Multiple system atrophy: insights into a rare and debilitating movement disorder', *Nat Rev Neurol,* 13(4), pp. 232–243.

Kunst, N., Burger, E. A., Coupé, V. M. H., Kuntz, K. M. and Aas, E. (2023a) 'A Guide to an Iterative Approach to Model-Based Decision Making in Health and Medicine: An Iterative Decision-Making Framework', *PharmacoEconomics*.

Kunst, N., Siu, A., Drummond, M., Grimm, S. E., Grutters, J., Husereau, D., Koffijberg, H., Rothery, C., Wilson, E. C. F. and Heath, A. (2023b) 'Consolidated Health Economic Evaluation Reporting Standards - Value of Information (CHEERS-VOI): Explanation and Elaboration', *Value in Health,* 26(10), pp. 1461–1473.

Kwon, C. S., Daniele, P., Forsythe, A. and Ngai, C. (2021) 'A Systematic Literature Review of the Epidemiology, Health-Related Quality of Life Impact, and Economic Burden of Immunoglobulin A Nephropathy', *Journal of Health Economics & Outcomes Research,* 8(2), pp. 36–45.

Langford, B., Zhou, Y. and Miyasaki, J. M. (2023) 'Multiple System Atrophy Caregivers' Experience: A Mixed Methods Study', *Canadian Journal of Neurological Sciences,* 50(1), pp. 49–59.

Largent, E. A. and Pearson, S. D. (2012) 'Which Orphans Will Find a Home? The Rule of Rescue in Resource Allocation for Rare Diseases', *Hastings Center Report,* 42(1), pp. 27–34.

Lasalvia, P., Hernandez, F., Castaneda-Cardona, C., Cuestas, J. A. and Rosselli, D. (2020) 'Cost-Effectiveness of Natalizumab Compared With Fingolimod for Relapsing-Remitting Multiple Sclerosis Treatment in Colombia', *Value in Health Regional Issues,* 23, pp. 13–18.

Lasalvia, P., Prieto-Pinto, L., Moreno, M., Castrillon, J., Romano, G., Garzon-Orjuela, N. and Rosselli, D. (2019) 'International experiences in multicriteria decision analysis (MCDA) for evaluating orphan drugs: a scoping review', *Expert Review of Pharmacoeconomics & Outcomes Research,* 19(4), pp. 409–420.

Lawton, S. E. (1991) 'Controversy Under the Orphan Drug Act: Is Resolution on the Way?', *Food, Drug, Cosmetic Law Journal,* 46(2), pp. 327–343.

Lazzaro, C., Bergamaschi, R., Zaffaroni, M., Totaro, R. and Paolicelli, D. (2022) 'Cost-utility analysis of teriflunomide in naive vs. previously treated patients with relapsing-remitting multiple sclerosis in Italy', *Neurological Sciences,* 43(8), pp. 4933–4944.

Le Neveu, M., Nicholson, R., Agrawal, P., Early, M. and Patterson, D. (2023) 'Determining health-related quality of life and health state utility values of recurrent urinary tract infections in women', *International Urogynecology Journal,* 34(8), pp. 1831–1835.

Lee, C., Lam, A., Kangappaden, T., Olver, P., Kane, S., Tran, D. and Ammann, E. (2022) 'Systematic literature review of evidence in amyloid light-chain amyloidosis', *Journal of Comparative Effectiveness Research,* 11(6), pp. 451–472.

Legal Information Institute (2022a) *21 U.S. Code § 360bb - Designation of drugs for rare diseases or conditions*. Available at: https://www.law.cornell.edu/uscode/text/21/360bb (Accessed: 14 December 2022).

Legal Information Institute (2022b) *21 U.S. Code § 360cc - Protection for drugs for rare diseases or conditions*. Available at: https://www.law.cornell.edu/uscode/text/21/360cc (Accessed: 13 December 2022).

Legal Information Institute (2022c) *21 U.S. Code § 360dd - Open protocols for investigations of drugs for rare diseases or conditions*. Available at: https://www.law.cornell.edu/uscode/text/21/360dd (Accessed: 13 December 2022).

Legal Information Institute (2022d) *21 U.S. Code § 360ee–1 - FDA rare neurodegenerative disease grant program*. Available at: https://www.law.cornell.edu/uscode/text/21/360ee-1 (Accessed: 13 December 2022).

Legal Information Institute (2022e) *21 U.S. Code § 360ee - Grants and contracts for development of drugs for rare diseases and conditions*. Available at: https://www.law.cornell.edu/uscode/text/21/360ee (Accessed: 13 December 2022).

Legal Information Institute (2022f) *21 U.S. Code § 360ff - Priority review to encourage treatments for rare pediatric diseases*. Available at: https://www.law.cornell.edu/uscode/text/21/360ff (Accessed: 13 December 2022).

Legal Information Institute (2022g) *21 U.S. Code § 360aa - Recommendations for investigations of drugs for rare diseases or conditions*. Available at: https://www.law.cornell.edu/uscode/text/21/360aa#:~:text=Go!-,21%20U.S.%20Code%20%C2%A7%20360aa%20%2D%20Recommendations%20for%20investigations%20of,for%20rare%20diseases%20or%20conditions&text=if%20the%20drug%20is%20a,section%20262%20of%20title%2042. (Accessed: 13 December 2022).

Legal Information Institute (2022h) *26 U.S. Code § 41 - Credit for increasing research activities*. Available at: https://www.law.cornell.edu/uscode/text/26/41 (Accessed: 13 December 2022).

Legal Information Institute (2022i) *26 U.S. Code § 45C - Clinical testing expenses for certain drugs for rare diseases or conditions*. Available at: https://www.law.cornell.edu/uscode/text/26/45C (Accessed: 13 December 2022).

Leonart, L. P., Borba, H. H. L., Ferreira, V. L., Riveros, B. S. and Pontarolo, R. (2018) 'Cost-effectiveness of acromegaly treatments: a systematic review', *Pituitary,* 21(6), pp. 642–652.

Liu, M., Wang, Z. and Shang, H. (2024) 'Multiple system atrophy: an update and emerging directions of biomarkers and clinical trials', *Journal of Neurology,* 271(5), pp. 2324–2344.

Llorian, E., Waliji, L. A., Dragojlovic, N., Michaux, K. D., Nagase, F. and Lynd, L. D. (2023) 'Frameworks for Health Technology Assessment at an Early Stage of Product Development: A Review and Roadmap to Guide Applications', *Value in Health,* 26(8), pp. 1258–1269.

Low, P. A., Reich, S. G., Jankovic, J., Shults, C. W., Stern, M. B., Novak, P., Tanner, C. M., Gilman, S., Marshall, F. J., Wooten, F., Racette, B., Chelimsky, T., Singer, W., Sletten, D. M., Sandroni, P. and Mandrekar, J. (2015) 'Natural history of multiple system atrophy in the USA: a prospective cohort study', *The Lancet Neurology,* 14(7), pp. 710–719.

Lui, N., Zhou, Y. and Lee, J. (2021) 'IPDfromKM: reconstruct individual patient data from published Kaplan-Meier survival curves', 2021.

Lunde, N. (2020) 'Lundes julekalender: Sådan fik A.P. Møller sit gennembrud i 1920'erne.', *Boersen*. Available at: https://borsen.dk/nyheder/virksomheder/ap-moller-ved-rettidig-omhu.

Mantovani, L. G., Cozzolino, P., Cortesi, P. A. and Patti, F. (2020) 'Cost-Effectiveness Analysis of Cannabinoid Oromucosal Spray Use for the Management of Spasticity in Subjects with Multiple Sclerosis', *Clinical drug investigation,* 40(4), pp. 319–326.

Marseille, E. and Kahn, J. G. (2019) 'Utilitarianism and the ethical foundations of cost-effectiveness analysis in resource allocation for global health', *Philosophy, Ethics, and Humanities in Medicine,* 14(1), pp. 5.

Martins, P., Vandewalle, B., Felix, J., Capela, C. M., Cerqueira, J. J., Salgado, A. V., Ferreira, D. G. and Monteiro, I. (2023) 'Cost-effectiveness Analysis of Ocrelizumab for the Treatment of Relapsing and Primary Progressive Multiple Sclerosis in Portugal', *PharmacoEconomics - Open,* 7(2), pp. 229–241.

Mauskopf, J. A., Sullivan, S. D., Annemans, L., Caro, J., Mullins, C. D., Nuijten, M., Orlewska, E., Watkins, J. and Trueman, P. (2007) 'Principles of good practice for budget impact analysis: report of the ISPOR Task Force on good research practices--budget impact analysis', *Value Health,* 10(5), pp. 336–47.

May, S., Gilman, S., Sowell, B. B., Thomas, R. G., Stern, M. B., Colcher, A., Tanner, C. M., Huang, N., Novak, P., Reich, S. G., Jankovic, J., Ondo, W. G., Low, P. A., Sandroni, P., Lipp, A., Marshall, F. J., Wooten, F. and Shults, C. W. (2007) 'Potential outcome measures and trial design issues for multiple system atrophy', *Movement disorders : official journal of the Movement Disorder Society,* 22(16), pp. 2371–7.

Mazzucato, M., Visonà Dalla Pozza, L., Manea, S., Minichiello, C. and Facchin, P. (2014) 'A population-based registry as a source of health indicators for rare diseases: the ten-year experience of the Veneto Region's rare diseases registry', *Orphanet J Rare Dis,* 9, pp. 37.

McCrone, P., Payan, C. A., Knapp, M., Ludolph, A., Agid, Y., Leigh, P. N. and Bensimon, G. (2011) 'The economic costs of progressive supranuclear palsy and multiple system atrophy in France, Germany and the United Kingdom', *PLoS One,* 6(9), pp. e24369.

McNamara, S., Schneider, P. P., Love-Koh, J., Doran, T. and Gutacker, N. (2023) 'Quality-Adjusted Life Expectancy Norms for the English Population', *Value in Health,* 26(2), pp. 163–169.

Meissner, W. G., Fernagut, P. O., Dehay, B., Péran, P., Traon, A. P., Foubert-Samier, A., Lopez Cuina, M., Bezard, E., Tison, F. and Rascol, O. (2019) 'Multiple System Atrophy: Recent Developments and Future Perspectives', *Mov Disord,* 34(11), pp. 1629–1642.

Mészáros, L., Hoffmann, A., Wihan, J. and Winkler, J. (2020) 'Current Symptomatic and Disease-Modifying Treatments in Multiple System Atrophy', *Int J Mol Sci,* 21(8).

Michel, M., Le Page, E., Laplaud, D. A., Wardi, R., Lebrun, C., Zagnoli, F., Wiertlewski, S., Coustans, M., Edan, G., Chevreul, K. and Veillard, D. (2022) 'Cost-utility of oral methylprednisolone in the treatment of multiple sclerosis relapses: Results from the COPOUSEP trial', *Revue neurologique,* 178(3), pp. 241–248.

Michels, R. E., de Fransesco, M., Mahajan, K., Hengstman, G. J. D., Schiffers, K. M. H., Budhia, S., Harty, G. and Krol, M. (2019) 'Cost Effectiveness of Cladribine Tablets for the Treatment of Relapsing-Remitting Multiple Sclerosis in The Netherlands', *Applied Health Economics and Health Policy,* 17(6), pp. 857–873.

Miller, K. L., Fermaglich, L. J. and Maynard, J. (2021) 'Using four decades of FDA orphan drug designations to describe trends in rare disease drug development: substantial growth seen in development of drugs for rare oncologic, neurologic, and pediatric-onset diseases', *Orphanet J Rare Dis,* 16(1), pp. 265.

Miller, K. L. and Lanthier, M. (2018) 'Investigating the landscape of US orphan product approvals', *Orphanet J Rare Dis,* 13(1), pp. 183.

Miyashita, M., Narita, Y., Sakamoto, A., Kawada, N., Akiyama, M., Kayama, M., Suzukamo, Y. and Fukuhara, S. (2009) 'Care burden and depression in caregivers caring for patients with intractable neurological diseases at home in Japan', *Journal of the neurological sciences,* 276(1-2), pp. 148–52.

Mohammadshahi, M., Olyaeemanesh, A., Ehsani-Chimeh, E., Mobinizadeh, M., Fakoorfard, Z., Akbari Sari, A. and Aghighi, M. (2022) 'Methods and criteria for the assessment of orphan drugs: a scoping review', *International Journal of Technology Assessment in Health Care,* 38(1), pp. e59.

Montgomery, S., Woodhouse, F., Vudumula, U., Gudala, K., Duddy, M. and Kroes, M. (2022) 'Stick or twist? Cost-effectiveness of siponimod compared with continuing existing disease-modifying therapies in the treatment of active secondary progressive multiple sclerosis in the UK', *Journal of Medical Economics,* 25(1), pp. 669–678.

Monzio Compagnoni, G. and Di Fonzo, A. (2019) 'Understanding the pathogenesis of multiple system atrophy: state of the art and future perspectives', *Acta Neuropathologica Communications,* 7(1), pp. 113.

Morton, F. and Singh, J. (2024) *eq5d: Methods for Analysing 'EQ-5D' Data and Calculating 'EQ-5D' Index Scores*. Available at: https://cran.r-project.org/web/packages/eq5d/index.html (Accessed: 31 October 2024).

Murakami, M. and Narukawa, M. (2016) 'Matched analysis on orphan drug designations and approvals: cross regional analysis in the United States, the European Union, and Japan', *Drug Discov Today,* 21(4), pp. 544–9.

Nakhaipour, H. R., Vudumula, U., Khurana, V., Sebire, G., Mah, J. K., Pohl, D., Schecter, R. and Adlard, N. (2020) 'Cost-effectiveness of fingolimod versus interferon-beta1a for the treatment of pediatric-onset multiple sclerosis in Canada', *Journal of Medical Economics,* 23(12), pp. 1525–1533.

National Cancer Institute (2022) *NCI Dictionary of Cancer Terms: Orphan drug designation*. Available at: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/orphan-drug-designation (Accessed: 16 December 2022).

National Health Service England (2016) *Appraisal and Funding of Cancer Drugs from July 2016 (including the new Cancer Drug Fund)*. Available at: https://www.england.nhs.uk/wp-content/uploads/2013/04/cdf-sop.pdf (Accessed: 27 October 2022).

National Health Service England (2021) *NHS England announces new Innovative Medicines Fund to fast-track promising new drugs*. Available at: https://www.england.nhs.uk/2021/07/nhs-england-announces-new-innovative-medicines-fund-to-fast-track-promising-new-drugs/ (Accessed: 8 December 2023).

National Institute for Health and Care Excellence (2017a) *Interim Process and Methods of the Highly Specialised Technologies Programme Updated to reflect 2017 changes*. Available at: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-highly-specialised-technologies-guidance/HST-interim-methods-process-guide-may-17.pdf (Accessed: 24 June 2023).

National Institute for Health and Care Excellence (2017b) *Parkinson's disease in adults. NICE guideline [NG71]. Appendix F - HE report.* Available at: https://www.nice.org.uk/guidance/ng71/evidence/appendix-f-he-report-pdf-4538466259 (Accessed: 22 April 2024).

National Institute for Health and Care Excellence (2022a) *Elosulfase alfa for treating mucopolysaccharidosis type 4A*. Available at: https://www.nice.org.uk/guidance/hst19/chapter/3-Committee-discussion (Accessed: 8 December 2023).

National Institute for Health and Care Excellence (2022b) *The Innovatives Medicines Fund Principles.* Available at: https://www.england.nhs.uk/wp-content/uploads/2022/06/B1686-the-innovate-medicines-fund-principles-june-2022.pdf (Accessed: 27 October 2023).

National Institute for Health and Care Excellence (2022c) *NICE health technology evaluation topic selection: the manual*. Available at: https://www.nice.org.uk/process/pmg37/chapter/highly-specialised-technologies (Accessed: 2023 13 December).

National Institute for Health and Care Excellence (2022d) *NICE health technology evaluations: the manual*. Available at: https://www.nice.org.uk/process/pmg36/chapter/committee-recommendations (Accessed: 11 November 2023).

National Institute for Health and Care Excellence (2022e) *NICE health technology evaluations: the manual*. Available at: https://www.nice.org.uk/process/pmg36/chapter/economic-evaluation-2 (Accessed: 23 January 2025).

National Institute for Health and Care Excellence (2023) *Ataluren for treating Duchenne muscular dystrophy with a nonsense mutation in the dystrophin gene*. Available at: https://www.nice.org.uk/guidance/hst22/chapter/3-Committee-discussion (Accessed: 8 December 2023).

National Institute for Health and Care Excellence (2024) *NICE process and methods. Appendix 1: highly specialised technologies*. Available at: https://www.nice.org.uk/process/pmg46/resources/highly-specialised-technologies-nice-prioritisation-board-routing-criteria-pdf-19830673472965 (Accessed: 28 August 2025).

National Institute for Health and Care Excellence (2025a) *Changes to NICE's cost-effectiveness thresholds confirmed*. Available at: https://www.nice.org.uk/news/articles/changes-to-nice-s-cost-effectiveness-thresholds-confirmed (Accessed: 6 February 2026).

National Institute for Health and Care Excellence (2025b) *NICE health technology evaluations: the manual* Available at: https://www.nice.org.uk/process/pmg36/chapter/evidence-2#guiding-principles-for-evidence (Accessed: November 7 2025).

National Institute for Health and Care Excellence (2026) *Our principles*. Available at: https://www.nice.org.uk/about-us/our-principles (Accessed: 4 February 2026).

National Institutes of Biomedical Innovation, Health and Nutrition (2022) *Orphan Products Development Support Program*. Available at: https://www.nibiohn.go.jp/en/activities/orphan-support.html (Accessed: 5 December 2022).

Navarro, C. E. and Betancur, J. E. (2023) 'Cost-Utility Analysis Comparing Ocrelizumab Versus Rituximab in the Treatment of Relapsing-Remitting Multiple Sclerosis: The Colombian Perspective', *Value in health regional issues,* 36, pp. 83–91.

Navarro, C. E., Ordóñez-Callamand, E. and Alzate, J. P. (2020) 'Disease modifying therapies in multiple sclerosis: cost-effectiveness systematic review', *Farm Hosp,* 44(2), pp. 68–76.

Naylor, N. R., Williams, J., Green, N., Lamrock, F. and Briggs, A. (2023) 'Extensions of Health Economic Evaluations in R for Microsoft Excel Users: A Tutorial for Incorporating Heterogeneity and Conducting Value of Information Analyses', *Pharmacoeconomics,* 41(1), pp. 21–32.

Nestler-Parr, S., Korchagina, D., Toumi, M., Pashos, C. L., Blanchette, C., Molsen, E., Morel, T., Simoens, S., Kalo, Z., Gatermann, R. and Redekop, W. (2018) 'Challenges in Research and Health Technology Assessment of Rare Disease Technologies: Report of the ISPOR Rare Disease Special Interest Group', *Value Health,* 21(5), pp. 493–500.

NHS Trust (2023) *Multiple System Atrophy* Leeds Centre for Neurosciences. Available at: https://flipbooks.leedsth.nhs.uk/LN005500.pdf (Accessed: 15 January 2025).

Nicod, E. (2017) 'Why do health technology assessment coverage recommendations for the same drugs differ across settings? Applying a mixed methods framework to systematically compare orphan drug decisions in four European countries', *Eur J Health Econ,* 18(6), pp. 715–730.

Ollendorf, D. A., Chapman, R. H. and Pearson, S. D. (2018) 'Evaluating and Valuing Drugs for Rare Conditions: No Easy Answers', *Value in Health,* 21(5), pp. 547–552.

Onakpoya, I. J., Spencer, E. A., Thompson, M. J. and Heneghan, C. J. (2015) 'Effectiveness, safety and costs of orphan drugs: an evidence-based review', *BMJ Open,* 5(6), pp. e007199.

Oppe, M., Ortin-Sulbaran, D., Vila Silvan, C., Estevez-Carrillo, A. and Ramos-Goni, J. M. (2021) 'Cost-effectiveness of adding Sativex spray to spasticity care in Belgium: using bootstrapping instead of Monte Carlo simulation for probabilistic sensitivity analyses', *European Journal of Health Economics,* 22(5), pp. 711–721.

Osaki, Y., Morita, Y., Kuwahara, T., Miyano, I. and Doi, Y. (2011) 'Prevalence of Parkinson's disease and atypical parkinsonian syndromes in a rural Japanese district', *Acta Neurol Scand,* 124(3), pp. 182–7.

Our world in data (2015) *Death rate by age group in England and Wales*. Available at: https://ourworldindata.org/grapher/death-rate-by-age-group-in-england-and-wales?tab=table (Accessed: 23 October 2024).

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P. and Moher, D. (2021) 'The PRISMA 2020 statement: an updated guideline for reporting systematic reviews', *Systematic Reviews,* 10(1), pp. 89.

Pahwa, R., Merola, A., Soileau, M., Alobaidi, A., Pickard, A. S., Kandukuri, P. L., Bao, Y., Strezewski, J., Oddsdottir, J., Xu, W. and Standaert, D. (2023) 'Cost-Effectiveness of Carbidopa-Levodopa Enteral Suspension for Advanced Parkinson's Disease in the United States', *Mov Disord,* 38(12), pp. 2308–2312.

Palmer, A. J., Zhao, T., Taylor, B. V., van der Mei, I. and Campbell, J. A. (2024) 'Exploring the cost-effectiveness of EBV vaccination to prevent multiple sclerosis in an Australian setting', *J Neurol Neurosurg Psychiatry,* 95(5), pp. 401–409.

Papapetropoulos, S., Tuchman, A., Laufer, D., Papatsoris, A. G., Papapetropoulos, N. and Mash, D. C. (2007) 'Causes of death in multiple system atrophy', *Journal of Neurology, Neurosurgery & Psychiatry,* 78(3), pp. 327.

Papatsoris, A. G., Papapetropoulos, S., Singer, C. and Deliveliotis, C. (2008) 'Urinary and erectile dysfunction in multiple system atrophy (MSA)', *Neurourol Urodyn,* 27(1), pp. 22–7.

Park, T., Griggs, S. K. and Suh, D. C. (2015) 'Cost Effectiveness of Monoclonal Antibody Therapy for Rare Diseases: A Systematic Review', *Biodrugs,* 29(4), pp. 259–74.

Pearson, I., Rothwell, B., Olaye, A. and Knight, C. (2018) 'Economic Modeling Considerations for Rare Diseases', *Value Health,* 21(5), pp. 515–524.

Picavet, E., Cassiman, D. and Simoens, S. (2015) 'What is known about the cost-effectiveness of orphan drugs? Evidence from cost-utility analyses', *Journal of Clinical Pharmacy & Therapeutics,* 40(3), pp. 304–7.

Pinheiro, B., Guerreiro, R., Costa, J. and Miguel, L. S. (2020) 'Cost-effectiveness of cladribine tablets versus fingolimod in patients with highly active relapsing multiple sclerosis in Portugal', *Journal of Medical Economics,* 23(5), pp. 484–491.

Pinto, D., Heinemann, A. W., Chang, S.-H., Charlifue, S., Field-Fote, E. C., Furbish, C. L., Jayaraman, A., Tefertiller, C., Taylor, H. B. and French, D. D. (2023) 'Cost-effectiveness analysis of overground robotic training versus conventional locomotor training in people with spinal cord injury', *Journal of neuroengineering and rehabilitation,* 20(1), pp. 10.

Plüddemann, A., Banerjee, A. and O'Sullivan, J. (2017) *Positive results bias.* Catalogue of Bias Collaboration. Available at: https://catalogofbias.org/biases/positive-results-bias/ (Accessed: 19 June 2023).

Poveda, J. L., Trillo, J. L., Rubio-Terres, C., Rubio-Rodriguez, D., Polanco, A. and Torres, C. (2020) 'Cost-effectiveness of Cladribine Tablets and fingolimod in the treatment of relapsing multiple sclerosis with high disease activity in Spain', *Expert Review of Pharmacoeconomics and Outcomes Research,* 20(3), pp. 295–303.

Pulsinelli, G. A. (1999) 'The Orphan Drug Act: What's Right with It', *Santa Clara High Technology Law Journal,* 15(2).

Querol, L., Crabtree, M., Herepath, M., Priedane, E., Viejo Viejo, I., Agush, S. and Sommerer, P. (2021) 'Systematic literature review of burden of illness in chronic inflammatory demyelinating polyneuropathy (CIDP)', *Journal of Neurology,* 268(10), pp. 3706–3716.

R Core Team (2021) *R: A language and environment for statistical computing*: R Foundation. Available at: https://www.r-project.org/ (Accessed: 1 March 2023).

Ramsey, S., Willke, R., Briggs, A., Brown, R., Buxton, M., Chawla, A., Cook, J., Glick, H., Liljas, B., Petitti, D. and Reed, S. (2005) 'Good Research Practices for Cost-Effectiveness Analysis Alongside Clinical Trials: The ISPOR RCT-CEA Task Force Report', *Value in Health,* 8(5), pp. 521–533.

Raut, M., Singh, G., Hiscock, I., Sharma, S. and Pilkhwal, N. (2022) 'A systematic literature review of the epidemiology, quality of life, and economic burden, including disease pathways and treatment patterns of relapsed/refractory classical Hodgkin lymphoma', *Expert Review of Hematology,* 15(7), pp. 607–617.

Reidy, M., Khan, M., Mills, E. A., Wu, Q., Garton, J., Draayer, D. E., Zahoor, I., Giri, S., Axtell, R. C. and Mao-Draayer, Y. (2025) 'New Frontiers in Multiple Sclerosis Treatment: From Targeting Costimulatory Molecules to Bispecific Antibodies', *International Journal of Molecular Sciences,* 26(8), pp. 3880.

Reinholdt, L., Chesler, E., Pera, M. and Rosenthal, N. (2025) 'The rare-to-common disease journey: a winding road to new therapies', *Trends in Genetics,* 41(9), pp. 762–773.

Renfrew, L. M., Paul, L., McFadyen, A., Rafferty, D., Moseley, O., Lord, A. C., Bowers, R. and Mattison, P. (2019) 'The clinical- and cost-effectiveness of functional electrical stimulation and ankle-foot orthoses for foot drop in Multiple Sclerosis: a multicentre randomized trial', *Clinical rehabilitation,* 33(7), pp. 1150–1162.

Rezaee, M., Izadi, S., Keshavarz, K., Borhanihaghighi, A. and Ravangard, R. (2019) 'Fingolimod versus natalizumab in patients with relapsing remitting multiple sclerosis: a cost-effectiveness and cost-utility study in Iran', *Journal of Medical Economics,* 22(4), pp. 297–305.

Rice, J., White, A., Scarpati, L., Philbin, M., Wan, G. and Nelson, W. (2018) 'Burden of noninfectious inflammatory eye diseases: A systematic literature review', *Journal of Managed Care and Specialty Pharmacy,* 23(3-A SUPPL.), pp. S67.

Richter, T., Nestler-Parr, S., Babela, R., Khan, Z. M., Tesoro, T., Molsen, E. and Hughes, D. A. (2015) 'Rare Disease Terminology and Definitions - A Systematic Global Review: Report of the ISPOR Rare Disease Special Interest Group', *Value Health,* 18(6), pp. 906–14.

Rittenhouse, B. E. and Nicod, E. (2024) 'Rare Diseases: Economic Evaluation and Policy Considerations', *Value in Health,* 27(12), pp. 1635–1636.

Roberts, M., Russell, L. B., Paltiel, A. D., Chambers, M., McEwan, P. and Krahn, M. (2012) 'Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2', *Med Decis Making,* 32(5), pp. 678–89.

Rodriguez-Monguio, R., Spargo, T. and Seoane-Vazquez, E. (2017) 'Ethical imperatives of timely access to orphan drugs: is possible to reconcile economic incentives and patients' health needs?', *Orphanet Journal of Rare Diseases,* 12(1).

Rogula, B., Lozano-Ortega, G. and Johnston, K. M. (2022) 'A Method for Reconstructing Individual Patient Data From Kaplan-Meier Survival Curves That Incorporate Marked Censoring Times', *MDM Policy & Practice,* 7(1), pp. 23814683221077643.

Rosato, R., Pagano, E., Giordano, A., Farinotti, M., Ponzio, M., Veronese, S., Confalonieri, P., Grasso, M. G., Patti, F. and Solari, A. (2021) 'Living with severe multiple sclerosis: Cost-effectiveness of a palliative care intervention and cost of illness study', *Multiple Sclerosis and Related Disorders,* 49, pp. 102756.

Rowen, D., Brazier, J. and Roberts, J. (2009) 'Mapping SF-36 onto the EQ-5D index: how reliable is the relationship?', *Health and Quality of Life Outcomes,* 7(1), pp. 27.

Sakushima, K., Nishimoto, N., Nojima, M., Matsushima, M., Yabe, I., Sato, N., Mori, M. and Sasaki, H. (2015) 'Epidemiology of Multiple System Atrophy in Hokkaido, the Northernmost Island of Japan', *The Cerebellum,* 14(6), pp. 682–687.

Sakushima, K., Takeda, H. and Aoi, Y. (2021) 'Orphan drug designation and development in Japan: 25 years of experience and assessment', *Nature Reviews Drug Discovery,* 20(12), pp. 893–894.

Salleh, S., Thokala, P., Brennan, A., Hughes, R. and Booth, A. (2017) 'Simulation Modelling in Healthcare: An Umbrella Review of Systematic Literature Reviews', *Pharmacoeconomics,* 35(9), pp. 937–949.

Saulnier, T., Fabbri, M., Le Goff, M., Helmer, C., Pavy-Le Traon, A., Meissner, W. G., Rascol, O., Proust-Lima, C. and Foubert-Samier, A. (2024) 'Patient-perceived progression in multiple system atrophy: natural history of quality of life', *Journal of neurology, neurosurgery, and psychiatry*.

Schlander, M., Dintsios, C. M. and Gandjour, A. (2018) 'Budgetary Impact and Cost Drivers of Drugs for Rare and Ultrarare Diseases', *Value in Health,* 21(5), pp. 525–531.

Schneider, P., McNamara, S., Love-Koh, J., Doran, T. and Gutacker, N. (2021) *QALY Shortfall Calculator*. Available at: https://shiny.york.ac.uk/shortfall (Accessed: 1 September 2025).

Schrag, A., Geser, F., Stampfer-Kountchev, M., Seppi, K., Sawires, M., Kollensperger, M., Scherfler, C., Quinn, N., Pellecchia, M. T., Barone, P., Del Sorbo, F., Albanese, A., Ostergaard, K., Dupont, E., Cardozo, A., Tolosa, E., Nilsson, C. F., Widner, H., Lindvall, O., Giladi, N., Gurevich, T., Daniels, C., Deuschl, G., Coelho, M., Sampaio, C., Abele, M., Klockgether, T., Schimke, N., Eggert, K. M., Oertel, W., Djaldetti, R., Colosimo, C., Meco, G., Poewe, W. and Wenning, G. K. (2006) 'Health-related quality of life in multiple system atrophy', *Movement disorders : official journal of the Movement Disorder Society,* 21(6), pp. 809–15.

Schrag, A., Selai, C., Mathias, C., Low, P., Hobart, J., Brady, N. and Quinn, N. P. (2007) 'Measuring health-related quality of life in MSA: The MSA-QoL', *Movement Disorders,* 22(16), pp. 2332–2338.

Schrag, A., Sheikh, S., Quinn, N. P., Lees, A. J., Selai, C., Mathias, C., Litvan, I., Lang, A. E., Bower, J. H., Burn, D. J., Low, P. and Jahanshahi, M. (2010) 'A comparison of depression, anxiety, and health status in patients with progressive supranuclear palsy and multiple system atrophy', *Movement Disorders,* 25(8), pp. 1077–1081.

Schuller, Y., Hollak, C. E. and Biegstraaten, M. (2015) 'The quality of economic evaluations of ultra-orphan drugs in Europe - a systematic review', *Orphanet Journal Of Rare Diseases,* 10, pp. 92.

Schur, N., Gudala, K., Vudumula, U., Vadapalle, S., Bhadhuri, A., Casanova, A., Adlard, N. and Schwenkglenks, M. (2021) 'Cost Effectiveness and Budget Impact of Siponimod Compared to Interferon Beta-1a in the Treatment of Adult Patients with Secondary Progressive Multiple Sclerosis with Active Disease in Switzerland', *PharmacoEconomics,* 39(5), pp. 563–577.

Sequeira, A. R., Mentzakis, E., Archangelidi, O. and Paolucci, F. (2021) 'The economic and health impact of rare diseases: A meta-analysis', *Health Policy and Technology,* 10(1), pp. 32–44.

Sharif, M., Sharif, F., Ali, H. and Ahmed, F. (2013) 'Systematic Reviews Explained: AMSTAR-How to Tell the Good From the Bad and the Ugly', *Oral health and dental management,* 12, pp. 9–16.

Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., Porter, A. C., Tugwell, P., Moher, D. and Bouter, L. M. (2007) 'Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews', *BMC Medical Research Methodology,* 7(1), pp. 10.

Shi, L., Hodges, M., Drummond, M., Ahn, J., Li, S. C., Hu, S., Augustovski, F., Hay, J. W. and Smeeding, J. (2010) 'Good Research Practices for Measuring Drug Costs in Cost-Effectiveness Analyses: An International Perspective: The ISPOR Drug Cost Task Force Report—Part VI', *Value in Health,* 13(1), pp. 28–33.

Shimohata, T., Ozawa, T., Nakayama, H., Tomita, M., Shinoda, H. and Nishizawa, M. (2008) 'Frequency of nocturnal sudden death in patients with multiple system atrophy', *Journal of Neurology,* 255(10), pp. 1483–1485.

Short, H., Stafinski, T. and Menon, D. (2015) 'A National Approach to Reimbursement Decision-Making on Drugs for Rare Diseases in Canada? Insights from Across the Ponds', *Healthcare Policy = Politiques de sante,* 10(4), pp. 24–46.

Shourick, J., Wack, M. and Jannot, A.-S. (2021) 'Assessing rare diseases prevalence using literature quantification', *Orphanet Journal of Rare Diseases,* 16(1), pp. 139.

SIGN (2021a) *Economic studies*. Available at: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.sign.ac.uk%2Fassets%2Fsearch-filters-economic-studies.docx&wdOrigin=BROWSELINK (Accessed: 10 November 2022).

SIGN (2021b) *Systematic Reviews*. Available at: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.sign.ac.uk%2Fassets%2Fsearch-filters-systematic-reviews.docx&wdOrigin=BROWSELINK (Accessed: 10 January 2023).

Simmonds, M., Elliott, J. H., Synnot, A. and Turner, T. (2022) 'Living Systematic Reviews', *Methods Mol Biol,* 2345, pp. 121–134.

Sinclair, S. (2022) 'Explaining rule of rescue obligations in healthcare allocation: allowing the patient to tell the right kind of story about their life', *Med Health Care Philos,* 25(1), pp. 31–46.

Smare, C., Lakhdari, K., Doan, J., Posnett, J. and Johal, S. (2020) 'Evaluating Partitioned Survival and Markov Decision-Analytic Modeling Approaches for Use in Cost-Effectiveness Analysis: Estimating and Comparing Survival Outcomes', *PharmacoEconomics,* 38(1), pp. 97–108.

Smith, R. and Schneider, P. (2020) 'Making health economic models Shiny: A tutorial', *Wellcome Open Res,* 5, pp. 69.

Smith, R. A., Mohammed, W. and Schneider, P. P. (2023) 'Packaging cost-effectiveness models in R: A tutorial. [version 1; peer review: 2 approved with reservations]', *Wellcome Open Res 2023,* (8), pp. 419.

Smith, R. A., Schneider, P. P. and Mohammed, W. (2022) 'Living HTA: Automating Health Technology Assessment with R', *Wellcome Open Research,* 7.

Sonnenberg, F. A. and Beck, J. R. (1993) 'Markov models in medical decision making: a practical guide', *Medical decision making,* 13(4), pp. 322–338.

Spelman, T., Herring, W. L., Acosta, C., Hyde, R., Jokubaitis, V. G., Pucci, E., Lugaresi, A., Laureys, G., Havrdova, E. K., Horakova, D., Izquierdo, G., Eichau, S., Ozakbas, S., Alroughani, R., Kalincik, T., Duquette, P., Girard, M., Petersen, T., Patti, F., Csepany, T., Granella, F., Grand'Maison, F., Ferraro, D., Karabudak, R., Jose Sa, M., Trojano, M., van Pesch, V., Van Wijmeersch, B., Cartechini, E., McCombe, P., Gerlach, O., Spitaleri, D., Rozsa, C., Hodgkinson, S., Bergamaschi, R., Gouider, R., Soysal, A., Castillo, T., Prevost, J., Garber, J., de Gans, K., Ampapa, R., Simo, M., Sanchez-Menoyo, J. L., Iuliano, G., Sas, A., van der Walt, A., John, N., Gray, O., Hughes, S., De Luca, G., Onofrj, M., Buzzard, K., Skibina, O., Terzi, M., Slee, M., Solaro, C., Oreja, G., Ramo-Tello, C., Fragoso, Y., Shaygannejad, V., Moore, F., Rajda, C., Aguera Morales, E. and Butzkueven, H. (2024) 'Comparative effectiveness and cost-effectiveness of natalizumab and fingolimod in rapidly evolving severe relapsing-remitting multiple sclerosis in the United Kingdom', *Journal of medical economics,* 27(1), pp. 109–125.

Spelman, T., Herring, W. L., Zhang, Y., Tempest, M., Pearson, I., Freudensprung, U., Acosta, C., Dort, T., Hyde, R., Havrdova, E., Horakova, D., Trojano, M., De Luca, G., Lugaresi, A., Izquierdo, G., Grammond, P., Duquette, P., Alroughani, R., Pucci, E., Granella, F., Lechner-Scott, J., Sola, P., Ferraro, D., Grand'Maison, F., Terzi, M., Rozsa, C., Boz, C., Hupperts, R., Van Pesch, V., Oreja-Guevara, C., van der Walt, A., Jokubaitis, V. G., Kalincik, T. and Butzkueven, H. (2022) 'Comparative Effectiveness and Cost-Effectiveness of Natalizumab and Fingolimod in Patients with Inadequate Response to Disease-Modifying Therapies in Relapsing-Remitting Multiple Sclerosis in the United Kingdom', *PharmacoEconomics,* 40(3), pp. 323–339.

Srivastava, T., Latimer, N. R. and Tappenden, P. (2021) 'Estimation of Transition Probabilities for State-Transition Models: A Review of NICE Appraisals', *Pharmacoeconomics,* 39(8), pp. 869–878.

Stanisic, S., Bertolotto, A., Berto, P., Di Procolo, P. and Morawski, J. (2019) 'The cost-effectiveness of alemtuzumab in the management of relapse-remitting multiple sclerosis in Italy', *Global and Regional Health Technology Assessment,* 2019.

Stein, J. D., Brown, G. C., Brown, M. M., Sharma, S., Hollands, H. and Stein, H. D. (2002) 'The quality of life of patients with hypertension', *J Clin Hypertens (Greenwich),* 4(3), pp. 181–8.

Sullivan, P. W., Slejko, J. F., Sculpher, M. J. and Ghushchyan, V. (2011) 'Catalogue of EQ-5D scores for the United Kingdom', *Med Decis Making,* 31(6), pp. 800–4.

Taheri, S., Sahraian, M. A. and Yousefi, N. (2019) 'Cost-effectiveness of alemtuzumab and natalizumab for relapsing-remitting multiple sclerosis treatment in Iran: decision analysis based on an indirect comparison', *Journal of Medical Economics,* 22(1), pp. 71–84.

Tang, H., Yao, J. and Wang, Z. (2025) 'Amyotrophic Lateral Sclerosis Masquerading as Multiple System Atrophy with Parkinsonism and Anxiety as Initial Manifestations', *Degener Neurol Neuromuscul Dis,* 15, pp. 95–100.

Thakore, N. J., Lapin, B. R., Kinzy, T. G. and Pioro, E. P. (2018) 'Deconstructing progression of amyotrophic lateral sclerosis in stages: a Markov modeling approach', *Amyotroph Lateral Scler Frontotemporal Degener,* 19(7-8), pp. 483–494.

Thakore, N. J., Pioro, E. P., Udeh, B. L., Lapin, B. R. and Katzan, I. L. (2020) 'A Cost-Effectiveness Framework for Amyotrophic Lateral Sclerosis, Applied to Riluzole', *Value in Health,* 23(12), pp. 1543–1551.

The European Parliament (2000) *REGULATION (EC) No 141/2000 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 16 December 1999 on orphan medicinal products.* Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32000R0141&from=EN (Accessed: 20 October 2022).

The Senate and House of Representatives (1984) *Health Promotion and Disease Prevention Amendments of 1984.* 98th Congress: Public Law 98-551.

Thokala, P., Srivastava, T., Smith, R., Ren, S., Whittington, M. D., Elvidge, J., Wong, R. and Uttley, L. (2023) 'Living Health Technology Assessment: Issues, Challenges and Opportunities', *PharmacoEconomics,* 41(3), pp. 227–237.

Thompson, J., Marijam, A., Mitrani-Gold, F. S., Wright, J. and Joshi, A. V. (2023) 'Activity impairment, health-related quality of life, productivity, and self-reported resource use and associated costs of uncomplicated urinary tract infection among women in the United States', *PLoS One,* 18(2), pp. e0277728.

Thorat, T., Neumann, P. J. and Chambers, J. D. (2018) 'Hemophilia Burden of Disease: A Systematic Review of the Cost-Utility Literature for Hemophilia', *Journal of Managed Care & Specialty Pharmacy,* 24(7), pp. 632–642.

Towse, A. (2010) 'Value based pricing, research and development, and patient access schemes. Will the United Kingdom get it right or wrong?', *Br J Clin Pharmacol,* 70(3), pp. 360–6.

van Prooije, T., Ruigrok, S., van den Berkmortel, N., Maas, R. P. P. W. M., Wijn, S., van Roon-Mom, W. M. C., van de Warrenburg, B. and Grutters, J. P. C. (2023) 'The potential value of disease-modifying therapy in patients with spinocerebellar ataxia type 1: an early health economic modeling study', *Journal of Neurology,* 270(8), pp. 3788–3798.

Van Wilder, L., Rammant, E., Clays, E., Devleesschauwer, B., Pauwels, N. and De Smedt, D. (2019) 'A comprehensive catalogue of EQ-5D scores in chronic disease: results of a systematic review', *Quality of Life Research,* 28(12), pp. 3153–3161.

Versteegh, M. M., Huygens, S. A., Wokke, B. W. H. and Smolders, J. (2022) 'Effectiveness and Cost-Effectiveness of 360 Disease-Modifying Treatment Escalation Sequences in Multiple Sclerosis', *Value in Health,* 25(6), pp. 984–991.

Visser, L. A., Folcher, M., Delgado Simao, C., Gutierrez Arechederra, B., Escudero, E., Uyl-de Groot, C. A. and Redekop, W. K. (2022) 'The Potential Cost-Effectiveness of a Cell-Based Bioelectronic Implantable Device Delivering Interferon-beta1a Therapy Versus Injectable Interferon-beta1a Treatment in Relapsing-Remitting Multiple Sclerosis', *PharmacoEconomics,* 40(1), pp. 91–108.

Wailoo, A. (2024) *A guide to calculating burden of illness for NICE Evaluations: Technical Support Document 23*. Available at: https://sheffield.ac.uk/nice-dsu/tsds/full-list (Accessed: August 25 2025).

Walter, E., Berger, T., Bajer-Kornek, B. and Deisenhammer, F. (2019) 'Cost-utility analysis of alemtuzumab in comparison with interferon beta, fingolimod, and natalizumab treatment for relapsing-remitting multiple sclerosis in Austria', *Journal of Medical Economics,* 22(3), pp. 226–237.

Wang, Y., Rattanavipapong, W. and Teerawattananon, Y. (2021) 'Using health technology assessment to set priority, inform target product profiles, and design clinical study for health innovation', *Technological Forecasting and Social Change,* 172, pp. 121000.

Watanabe, H., Riku, Y., Hara, K., Kawabata, K., Nakamura, T., Ito, M., Hirayama, M., Yoshida, M., Katsuno, M. and Sobue, G. (2018) 'Clinical and Imaging Features of Multiple System Atrophy: Challenges for an Early and Clinically Definitive Diagnosis', *J Mov Disord,* 11(3), pp. 107–120.

Watanabe, H., Saito, Y., Terao, S., Ando, T., Kachi, T., Mukai, E., Aiba, I., Abe, Y., Tamakoshi, A., Doyu, M., Hirayama, M. and Sobue, G. (2002) 'Progression and prognosis in multiple system atrophy: An analysis of 230 Japanese patients', *Brain,* 125(5), pp. 1070–1083.

Weidlich, D., Kefalas, P. and Guest, J. F. (2016) 'Healthcare costs and outcomes of managing beta-thalassemia major over 50 years in the United Kingdom', *Transfusion,* 56(5), pp. 1038–45.

Weinstein, M. C., O'Brien, B., Hornberger, J., Jackson, J., Johannesson, M., McCabe, C. and Luce, B. R. (2003) 'Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices—Modeling Studies', *Value in Health,* 6(1), pp. 9–17.

Wellman-Labadie, O. and Zhou, Y. (2010) 'The US Orphan Drug Act: rare disease research stimulator or commercial opportunity?', *Health Policy,* 95(2-3), pp. 216–28.

Wenning, G. K., Geser, F., Krismer, F., Seppi, K., Duerr, S., Boesch, S., Köllensperger, M., Goebel, G., Pfeiffer, K. P., Barone, P., Pellecchia, M. T., Quinn, N. P., Koukouni, V., Fowler, C. J., Schrag, A., Mathias, C. J., Giladi, N., Gurevich, T., Dupont, E., Ostergaard, K., Nilsson, C. F., Widner, H., Oertel, W., Eggert, K. M., Albanese, A., del Sorbo, F., Tolosa, E., Cardozo, A., Deuschl, G., Hellriegel, H., Klockgether, T., Dodel, R., Sampaio, C., Coelho, M., Djaldetti, R., Melamed, E., Gasser, T., Kamm, C., Meco, G., Colosimo, C., Rascol, O., Meissner, W. G., Tison, F. and Poewe, W. (2013) 'The natural history of multiple system atrophy: a prospective European cohort study', *Lancet Neurol,* 12(3), pp. 264–74.

Wenning, G. K., Shlomo, Y. B., Magalhães, M., Danie, S. E. and Quinn, N. P. (1994) 'Clinical features and natural history of multiple system atrophy: An analysis of 100 cases', *Brain,* 117(4), pp. 835–845.

Wenning, G. K., Tison, F., Seppi, K., Sampaio, C., Diem, A., Yekhlef, F., Ghorayeb, I., Ory, F., Galitzky, M., Scaravilli, T., Bozi, M., Colosimo, C., Gilman, S., Shults, C. W., Quinn, N. P., Rascol, O. and Poewe, W. (2004) 'Development and validation of the Unified Multiple System Atrophy Rating Scale (UMSARS)', *Mov Disord,* 19(12), pp. 1391–402.

Whitty, C. J. M. (2015) 'What makes an academic paper useful for health policy?', *BMC Medicine,* 13(1), pp. 301.

Williams, C., Lewsey, J. D., Mackay, D. F. and Briggs, A. H. (2017) 'Estimation of Survival Probabilities for Use in Cost-effectiveness Analyses: A Comparison of a Multi-state Modeling Survival Analysis Approach with Partitioned Survival and Markov Decision-Analytic Modeling', *Med Decis Making,* 37(4), pp. 427–439.

Winter, Y., Spottke, A. E., Stamelou, M., Cabanel, N., Eggert, K., Hoglinger, G. U., Sixel-Doering, F., Herting, B., Klockgether, T., Reichmann, H., Oertel, W. H. and Dodel, R. (2011a) 'Health-related quality of life in multiple system atrophy and progressive supranuclear palsy', *Neurodegenerative Diseases,* 8(6), pp. 438–446.

Winter, Y., Stamelou, M., Cabanel, N., Sixel-Döring, F., Eggert, K., Höglinger, G. U., Herting, B., Klockgether, T., Reichmann, H., Oertel, W. H., Dodel, R. and Spottke, A. E. (2011b) 'Cost-of-illness in multiple system atrophy and progressive supranuclear palsy', *J Neurol,* 258(10), pp. 1827–34.

Woersching, A. L., Borrego, M. E. and Raisch, D. W. (2016) 'Assessing the Quality of Economic Evaluations of FDA Novel Drug Approvals: A Systematic Review', *Annals of Pharmacotherapy,* 50(12), pp. 1028–1040.

Wynne, O., Szewczyk, Z., Hollis, J., Farragher, E., Doherty, E., Tully, B., Paolucci, F., Gillham, K., Reeves, P., Wiggers, J. and Kingsland, M. (2023) 'Study protocol for an economic evaluation and budget impact of implementation strategies to support routine provision of antenatal care for gestational weight gain: a stepped-wedge cluster trial', *Implementation Science Communications,* 4(1), pp. 40.

Xiao, Y., Zhang, L., Wei, Q., Ou, R., Hou, Y., Liu, K., Lin, J., Yang, T. and Shang, H. (2022) 'Health-related quality of life in patients with multiple system atrophy using the EQ-5D-5L', *Brain and Behavior,* 12(10), pp. e2774.

Xin, Y., Gray, E., Robles-Zurita, J. A., Haghpanahan, H., Heggie, R., Kohli-Lynch, C., Briggs, A., McAllister, D. A., Lawson, K. D. and Lewsey, J. (2022) 'From Spreadsheets to Script: Experiences From Converting a Scottish Cardiovascular Disease Policy Model into R', *Applied Health Economics and Health Policy,* 20(2), pp. 149–158.

Xu, Y., Mao, N., Chirikov, V., Du, F., Yeh, Y.-C., Liu, L., Liu, R. and Gao, X. (2019) 'Cost-effectiveness of Teriflunomide Compared to Interferon Beta-1b for Relapsing Multiple Sclerosis Patients in China', *Clinical drug investigation,* 39(3), pp. 331–340.

York Health Economics Consortium (2016a) *Budget Impact Analysis [online]*. York. Available at: https://yhec.co.uk/glossary/budget-impact-analysis/ (Accessed: 29 November 2023).

York Health Economics Consortium (2016b) *Economic Evaluation [online]*. Available at: https://yhec.co.uk/glossary/economic-evaluation/ (Accessed: 29 November 2023).

York Health Economics Consortium (2016c) *Health Technology Assessment [online]*. Available at: http://www.yhec.co.uk/glossary/health-technology-assessment/ (Accessed: 25 October 2022).

York Health Economics Consortium (2016d) *Highly Specialised Technologies (UK NICE) [online]*. Available at: https://yhec.co.uk/glossary/highly-specialised-technologies-uk-nice/ (Accessed: 6 March 2024).

York Health Economics Consortium (2016e) *Partitioned Survival Model [online]*. Available at: https://yhec.co.uk/glossary/partitioned-survival-model/ (Accessed: 11 October 2024).

York Health Economics Consortium (2016f) *Preferences-Based Measures [online]*. Available at: https://yhec.co.uk/glossary/preference-based-measures/ (Accessed: 2 May 2025).

York Health Economics Consortium (2016g) *Probabilistic/Stochastic Sensitivity Analysis [online]*. Available at: https://yhec.co.uk/glossary/probabilistic-stochastic-sensitivity-analysis/ (Accessed: 12 June 2025).

York Health Economics Consortium (2025a) *Cost-benefit analysis*. Available at: https://www.yhec.co.uk/glossary-term/cost-benefit-analysis/ (Accessed: 4 February 2026).

York Health Economics Consortium (2025b) *Cost-effectiveness analysis*. Available at: https://www.yhec.co.uk/glossary-term/cost-effectiveness-analysis/ (Accessed: 29 January 2026).

York Health Economics Consortium (2025c) *Expected value of sample information (EVSI)*. Available at: https://www.yhec.co.uk/glossary-term/expected-value-of-sample-information-evsi/ (Accessed: 9 February 2026).

York Health Economics Consortium (2025d) *Opportunity costs*. Available at: https://www.yhec.co.uk/glossary-term/opportunity-cost/ (Accessed: 5 February 2026).

Zelei, T., Mendola, N. D., Elezbawy, B., Nemeth, B. and Campbell, J. D. (2021) 'Criteria and Scoring Functions Used in Multi-criteria Decision Analysis and Value Frameworks for the Assessment of Rare Disease Therapies: A Systematic Literature Review', *PharmacoEconomics Open,* 5(4), pp. 605–612.

Zelei, T., Molnar, M. J., Szegedi, M. and Kalo, Z. (2016) 'Systematic review on the evaluation criteria of orphan medicines in Central and Eastern European countries', *Orphanet Journal Of Rare Diseases,* 11(1), pp. 72.

Zhang, L., Hou, Y., Gu, X., Cao, B., Wei, Q., Ou, R., Liu, K., Lin, J., Yang, T., Xiao, Y., Zhao, B. and Shang, H. (2023) 'Prediction of early-wheelchair dependence in multiple system atrophy based on machine learning algorithm: A prospective cohort study', *Clin Park Relat Disord,* 8, pp. 100183.

Zhang, S., Shao, G., Wu, B., Xia, L., Wang, L., Li, L., Jin, K., Zou, Y. and Sun, C. (2025) 'Progress in Disease-Modifying Therapies for Parkinson's Disease', *Aging and disease*, pp. 0–.

Zhao, X., Yang, H., Wei, T., Zhao, J., Liu, J., Huang, Z., Zhu, L., Zhao, Y. and Li, W. (2022) 'Cost-effectiveness analysis of prolonged-release fampridine to treat walking disability of multiple sclerosis in China', *Journal of Comparative Effectiveness Research,* 11(14), pp. 1057–1069.

Appendices

Appendices are included on the following pages.

I Appendix (Umbrella Review)

Appendix A is divided by two headings: tables and figures.

I.I Appendix: Tables

Table 25: Scoping searches in Centre for Reviews and Dissemination (CRD) database conducted on the 10th of January 2023.

Search strategy	Hits
"MesH DESCRIPTOR Rare Diseases EXPLODE ALL TREES"	10

Table 26: Scoping searches in PROSPERO conducted on 10th of January 2023.

Search strategy	Hits
"rare disease" OR "orphan disease"	388
"rare disease" OR "orphan disease" AND economic*	24

Table 27: Scoping searches in MEDLINE (1946 to January 9, 2023) by OVID conducted on the 10th of January 2023.

Topic	Number	Search	Hits
Search filter: Economic studies	1	Economics/ OR costs and cost analysis/ OR cost allocation/ OR cost-benefit analysis/ OR cost control/ OR cost savings/ OR cost of illness/ OR cost sharing/ OR "Deductibles and Coinsurance"/ OR medical savings accounts/ OR health care costs/ OR direct service costs/ OR drug costs/ OR employer health costs/ OR hospital costs/ OR health expenditures/ OR capital expenditures/ OR value of life/ OR Exp economics, hospital/ OR Exp economics, medical/ OR Economics, nursing/ OR	881,241

		Economics, pharmaceutical/ OR exp "fees and charges"/ OR exp budgets/ OR (low adj cost).mp. OR (high adj cost).mp. OR (health?care adj cost*).mp. OR (fiscal or funding or financial or finance).tw. OR (cost adj estimate*).mp. OR (cost adj variable).mp. OR (unit adj cost*).mp. OR (economic* or pharmacoeconomic* or price* or pricing).tw.	
Rare Diseases	2	(disease, orphan.mp OR disease, rare.mp OR orphan disease.mp OR orphan diseases.mp OR rare disease.mp OR rare diseases.mp)	41,794
Rare diseases and economic studies	3	1 AND 2	1,565

Table 28: MEDLINE (1946 to January 9, 2023) searches on Ovid platform performed on the 10th of January 2023.

	Parameter	**Number**	**Search**	**Hits**
Setting	Health Economics (Search filter: Economic evaluations and Models)	1	Economics/	27,484
		2	exp "Costs and Cost Analysis"/	261,968
		3	Economics, Nursing/	4,013
		4	Economics, Medical/	9,232
		5	Economics, Pharmaceutical/	3,092
		6	exp Economics, Hospital/	25,665
		7	Economics, Dental/	1,920
		8	exp "Fees and Charges"/	31,278
		9	exp Budgets/	14,065
		10	budget*.ti,ab,kf.	34,902
		11	(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or	272,212

			financial or finance or finances or financed).ti,kf.	
		12	(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ab. /freq=2	7,590
		13	(cost* adj2 (effective* or utilit* or benefit* or minimi* or analy* or outcome or outcomes)).ab,kf.	201,151
		14	(value adj2 (money or monetary)).ti,ab,kf.	2,938
		15	exp models, economic/	16,168
		16	economic model*.ab,kf.	4,040
		17	markov chains/	15,880
		18	markov.ti,ab,kf.	27,949
		19	monte carlo method/	31,841
		20	monte carlo.ti,ab,kf.	58,282
		21	exp Decision Theory/	12,994
		22	(decision* adj2 (tree* or analy* or model*)).ti,ab,kf.	35,018
		23	or/1-22	726,567
Interest	Rare diseases	**24**	(disease, orphan OR disease, rare OR orphan disease OR orphan diseases OR rare disease OR rare diseases OR orphan condition* OR rare condition*).mp	62,743
Comparison	Orphan drugs	**25**	(orphan, drug OR rare, drug OR orphan, medicine or ultra,	2,310

			orphan or ultra-orphan or orphan, treatment).mp	
Evaluation	Systematic reviews (Search filter: Systematic reviews)	26	Meta-Analysis as Topic/	22,022
		27	meta analy$.tw.	254,567
		28	metaanaly$.tw.	2,500
		29	Meta-Analysis/	173,479
		30	(systematic adj (review$1 or overview$1)).tw.	270,513
		31	exp Review Literature as Topic/	21,484
		32	or/26-31	430,029
		33	cochrane.ab.	125,511
		34	embase.ab.	143,219
		35	(psychlit or psyclit).ab.	917
		36	(psychinfo or psycinfo).ab.	54,840
		37	(cinahl or cinhal).ab.	42,823
		38	science citation index.ab.	3,669
		39	bids.ab.	648
		40	cancerlit.ab.	638
		41	or/33-40	229,344
		42	reference list$.ab.	21,491
		43	bibliograph$.ab.	21,793
		44	hand-search$.ab.	8,319
		45	relevant journals.ab.	1,328
		46	manual search$.ab.	5,775
		47	or/42-46	52,730
		48	selection criteria.ab.	35,074
		49	data extraction.ab	30,427
		50	48 or 49	62,906
		51	Review/	3,099,076
		52	50 and 51	33,660
		53	Comment/	991,959
		54	Letter/	1,203,767
		55	Editorial/	632,217

		56	animal/	7,216,051
		57	human/	20,975,192
		58	56 not (56 and 57)	19,430
		59	or/53-55,58	2,141,848
		60	32 or 41 or 47 or 52	513,986
		61	60 not 59	494,979
Other	Combining searches	62	23 AND 61	21,974
		63	24 OR 25	64,164
		64	62 AND 63	76
Other	Publications from 2010 to 2023.	65	limit 64 to yr="2010 - 2023"	67

Table 29: EMBASE (1974 to Week 1, 2023) searches performed on the 11th of January 2023.

	Parameter	**Number**	**Search**	**Hits**
Setting	Health Economics (Search filter: Economic evaluations and Models)	1	Economics/	244,024
		2	Cost/	61,776
		3	exp Health Economics/	992,192
		4	Budget/	32,641
		5	budget*.ti,ab,kf.	46,129
		6	(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ti,kf.	334,090
		7	(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures	119

			or expense or expenses or financial or finance or finances or financed).ab. /freq=2	
		8	(cost* adj2 (effective* or utilit* or benefit* or minimi* or analy* or outcome or outcomes)).ab,kf.	278,888
		9	(value adj2 (money or monetary)).ti,ab,kf.	3,950
		10	Statistical Model/	172,500
		11	economic model*.ab,kf.	6,019
		12	Probability/	136,956
		13	markov.ti,ab,kf.	36,588
		14	monte carlo method/	48,334
		15	monte carlo.ti,ab,kf.	60,618
		16	Decision Theory/	1,815
		17	Decision Tree/	19,566
		18	(decision* adj2 (tree* or analy* or model*)).ti,ab,kf.	47,922
		19	or/1-18	1,781,668
Interest	Rare diseases	**20**	(disease, orphan OR disease, rare OR orphan disease OR orphan diseases OR rare disease OR rare diseases OR orphan condition* OR rare condition*).mp	108,795
Comparison	Orphan drugs	**21**	(orphan, drug OR rare, drug OR orphan, medicine or ultra, orphan or ultra-orphan or orphan, treatment).mp	4,963
Evaluation	Systematic reviews	22	exp Meta Analysis/	268,096
		23	((meta adj analy$) or metaanalys$).tw.	326,625

	(Search filter: Systematic reviews)	24	(systematic adj (review$1 or overview$1)).tw.	327,183
		25	or/22-24	538,104
		26	cancerlit.ab.	738
		27	cochrane.ab.	158,685
		28	embase.ab.	177,646
		29	(psychlit or psyclit).ab.	1,004
		30	(psychinfo or psycinfo).ab.	52,294
		31	(cinahl or cinhal).ab.	49,918
		32	science citation index.ab.	4,206
		33	bids.ab.	817
		34	or/26-33	273,205
		35	reference lists.ab.	22,959
		36	bibliograph$.ab.	27,875
		37	hand-search$.ab.	10,169
		38	manual search$.ab.	6,763
		39	relevant journals.ab.	1,563
		40	or/35-39	62,551
		41	data extraction.ab.	36,946
		42	selection criteria.ab.	43,268
		43	41 or 42	77,650
		44	review.pt.	2,997,996
		45	43 and 44	36,240
		46	letter.pt.	1,252,462
		47	editorial.pt.	748,254
		48	animal/	1,599,075
		49	human/	24,377,335
		50	48 not (48 and 49)	51,330
		51	or/46-47,50	2,052,033
		52	25 or 34 or 40 or 45	635,467
		53	52 not 51	621,078
Other	Combining searches	54	19 and 53	50,105
		55	20 or 21	112,165
		56	54 and 55	226

Other	Publications from 2010 to 2023.	57	limit 56 to yr="2010 - 2023"	211

Table 30: Cochrane Database of Systematic Reviews (Issue 1 of 12, January 23) searches performed on the 11th of January 2023.

Search strategy	Number	Hits
MeSH descriptor: [Rare Diseases] explode all trees	1	4 Cochrane reviews
From January 2012 to January 2023	2	4 Cochrane reviews

Table 31: Papers included for full screening.

Authors	Year	Title	Papers included after full screening: 36	Comment
Studies from literature search				
K. Abdallah; I. Huys; K. Claes; S. Simoens	2021	Methodological Quality Assessment of Budget Impact Analyses for Orphan Drugs: A Systematic Review	Included	Quality assessment of budget impact analyses for orphan drugs
A. Angelis; D. Tordrup; P. Kanavos	2015	Socio-economic burden of rare diseases: A systematic review of cost of illness evidence	Included	Systematic review of cost-of-illness studies for ten rare diseases
A. Babac; K. Damm; J. M. Graf von der Schulenburg	2019	Patient-reported data informing early benefit assessment of rare diseases in Germany: A systematic review	Included	How patient perspectives are considered as part of early benefit assessment of rare diseases in Germany
A. Baran-Kooiker; M. Czech; C. Kooiker	2018	Multi-Criteria Decision Analysis (MCDA) Models in Health Technology Assessment of Orphan Drugs-a Systematic Literature Review. Next Steps in Methodology Development?	Included	Multi-criteria decision analysis in health technology assessment of orphan drugs
T. Chen; J. Chen; C. Chen; H. Zheng; Y. Chen; M. Liu; B. Zheng	2021	Systematic review and cost-effectiveness of bosentan and sildenafil as therapeutic drugs for paediatric pulmonary arterial hypertension	Included	Systematic review of drugs and economic evaluation for rare hypertension in children

M. M. Cheng; S. D. Ramsey; E. B. Devine; L. P. Garrison; B. W. Bresnahan; D. L. Veenstra	2012	Systematic review of comparative effectiveness data for oncology orphan drugs	Included	Systematic review of trials and economic evaluations for orphan oncology drugs
O. Dabbous; L. Chachoua; S. Aballea; M. Sivignon; D. Malone; S. Petrou; J. Richardson; S. Simoens; M. Toumi	2021 (Changed to 2023 - Full text was available)	Valuation of Treatments for Rare Diseases: A Systematic Literature Review of Societal Preference Studies	Excluded	The systematic review was concerned with the valuation of elements such as prevalence, severity, unmet need etc. compared to other diseases. This paper did not describe them in the context of economic evaluation. Although, perspectives are important for deciding on resource allocation and can potentially be used for weighting in MCDA, the article was nonetheless excluded.
A. Dhawan; M. W. Lawlor; G. V. Mazariegos; P. McKiernan; J. E. Squires; K. A. Strauss; D. Gupta;	2020	Disease burden of Crigler-Najjar syndrome: Systematic review and future perspectives	Included	Systematic review and meta-analysis to describe the burden of Crigler-Najjar Syndrome.

E. James; S. Prasad				
E. Faulkner; D. S. Spinner; M. Ringo; M. Carroll	2019	Are Global Health Systems Ready for Transformative Therapies?	Included	Systematic review of transformative therapies, some of which are orphan drugs, which meet inclusion criteria
N. Franzen; V. P. Retel; W. Schats; W. H. van Harten	2020	Evidence Underlying Policy Proposals for Sustainable Anticancer Drug Prices: A Systematic Review	Excluded	Review is concerned with policies for sustainable anti-cancer drugs. There is no emphasis on orphan drugs in relation to economic evaluation
T. Friede; C. Rover; S. Wandel; B. Neuenschwander	2017	Meta-analysis of two studies in the presence of heterogeneity with applications in rare diseases	Excluded	A meta-analysis, but not a systematic literature review. I will exclude it for this reason, although a meta-analysis can from part of an economic evaluation. This is one is not described in the context of economic evaluation.
C. Friedmann; P. Levy; P. Hensel; M. Hiligsmann	2018	Using multi-criteria decision analysis to appraise orphan drugs: a systematic review	Included	Systematic review for MCDA to appraise orphan drugs

L. Garcia-Perez; R. Linertova; C. Valcarcel-Nazco; M. Posada; I. Gorostiza; P. Serrano-Aguilar	2021	Cost-of-illness studies in rare diseases: a scoping review	Included	Scoping review of cost-of-illness studies in rare diseases. Included because scoping review claimed to have a systematic approach to research and report
S. Gruhn; J. Witte; W. Greiner; O. Damm; M. Dietzsch; R. Kramer; M. Knuf	2022	Epidemiology and economic burden of meningococcal disease in Germany: A systematic review	Included	Systematic review of burden of diseases for invasive meningococcal disease in Germany
R. W. Jakes; N. Kwon; B. Nordstrom; R. Goulding; K. Fahrbach; J. Tarpey; M. K. Van Dyke	2021	Burden of illness associated with eosinophilic granulomatosis with polyangiitis: a systematic literature review and meta-analysis	Included	Systematic review and meta-analysis to describe incidence, prevalence, and disease burden
T. A. Kanters; C. de Sonneville-Koedoot; W. K. Redekop; L. Hakkaart	2013	Systematic review of available evidence on eleven high-priced inpatient orphan drugs	Included	Systematic review of evidence for high priced orphan drugs

E. I. Katsigianni; P. Petrou	2022	A systematic review of economic evaluations of enzyme replacement therapy in Lysosomal storage diseases	Included	Systematic review of economic evaluations
N. Knoble; G. Nayroles; C. Cheng; B. Arnould	2018	Illustration of patient-reported outcome challenges and solutions in rare diseases: a systematic review in Cushing's syndrome	Included	Systematic review of patient reported outcomes
C. S. Kwon; P. Daniele; A. Forsythe; C. Ngai	2021	A Systematic Literature Review of the Epidemiology, Health-Related Quality of Life Impact, and Economic Burden of Immunoglobulin A Nephropathy	Included	Systematic review of QoL and cost-of-illness
P. Lasalvia; L. Prieto-Pinto; M. Moreno; J. Castrillon; G. Romano; N. Garzon-Orjuela; D. Rosselli	2019	International experiences in multicriteria decision analysis (MCDA) for evaluating orphan drugs: a scoping review	Included	Scoping review of MCDA for orphan drugs It was included despite being a scoping review
C. Lee; A. Lam; T. Kangappaden; P. Olver; S. Kane; D. Tran; E. Ammann	2022	Systematic literature review of evidence in amyloid light-chain amyloidosis	Included	Systematic review for clinical, HRQoL and economics

L. P. Leonart; H. H. L. Borba; V. L. Ferreira; B. S. Riveros; R. Pontarolo	2018	Cost-effectiveness of acromegaly treatments: a systematic review	Included	Systematic review of cost-effectiveness analyses
D. Menon; D. Clark; T. Stafinski	2015	Reimbursement of drugs for rare diseases through the public healthcare system in Canada: Where are we now?	Excluded	This was a systematic review of funding pathways for orphan drugs and variation of access. Although this is interesting, the focus was not on the components of economic evaluation, why it was ultimately excluded.
P. A. Miller; S. M. Mulla; T. Adams-Webber; Y. Sivji; G. H. Guyatt; B. C. Johnston	2016	Patient-reported outcomes in rare lysosomal storage diseases: Key informant interviews and a systematic review protocol	Excluded	This was a systematic review protocol, why it was excluded.
M. Mohammadshahi; A. Olyaeemanesh; E. Ehsani-Chimeh; M. Mobinizadeh; Z. Fakoorfard; A.	2022	Methods and criteria for the assessment of orphan drugs: a scoping review	Included	Scoping review on criteria used for assessment of orphan drugs It was included despite being a scoping review

Akbari Sari; M. Aghighi				
S. Nestler-Parr; D. Korchagina; M. Toumi; C. L. Pashos; C. Blanchette; E. Molsen; T. Morel; S. Simoens; Z. Kalo; R. Gatermann; W. Redekop	2018	Challenges in Research and Health Technology Assessment of Rare Disease Technologies: Report of the ISPOR Rare Disease Special Interest Group	Excluded	Although this paper is full of valuable information, it is mainly informed by experts and the literature review mainly served to corroborate their opinions. Despite a targeted review component, the information on methods was sparse and do not indicate that it was systematic. Why this paper needs to be excluded.
I. J. Onakpoya; E. A. Spencer; M. J. Thompson; C. J. Heneghan	2015	Effectiveness, safety and costs of orphan drugs: an evidence-based review	Included	Systematic review to explore the effectiveness and costs of orphan drugs
T. Park; S. K. Griggs; D. C. Suh	2015	Cost Effectiveness of Monoclonal Antibody Therapy for Rare Diseases: A Systematic Review	Included	Systematic review of cost-effectiveness for monoclonal antibody-based orphan drugs

E. Picavet; D. Cassiman; S. Simoens	2015	What is known about the cost-effectiveness of orphan drugs? Evidence from cost-utility analyses	Included	Systematic review of cost-utility analyses for orphan drugs
L. Querol; M. Crabtree; M. Herepath; E. Priedane; I. Viejo; S. Agush; P. Sommerer	2020	Systematic literature review of burden of illness in chronic inflammatory demyelinating polyneuropathy (CIDP)	Included	Systematic review of burden of disease for chronic inflammatory demyelinating polyneuropathy (CIDP)
J. Rajpura; F. E. Shafer; K. M. Sivamurthy; C. Kollmer; N. Vendetti	2012	Health outcomes in vaso occlusive crisis of sickle cell disease-retrospective indicators	Excluded	It is an abstract
M. Raut; G. Singh; I. Hiscock; S. Sharma; N. Pilkhwal	2022	A systematic literature review of the epidemiology, quality of life, and economic burden, including disease pathways and treatment patterns of relapsed/refractory classical Hodgkin lymphoma	Included	Systematic review of the burden of disease for R/R classical Hodgkin lymphoma. It does include cell therapy, but also drugs that are not considered ATMPs.
J. Rice; A. White; L. Scarpati; M. Philbin; G. Wan; W. Nelson	2017	Burden of non-infectious inflammatory eye diseases: A systematic literature review	Included	Systematic review of burden of disease for non-infectious intraocular inflammatory eye disease

M. Schlander; C. M. Dintsios; A. Gandjour	2018	Budgetary Impact and Cost Drivers of Drugs for Rare and Ultrarare Diseases	Included	Systematic review of budget impact and cost drivers for rare diseases
Y. Schuller; C. E. Hollak; M. Biegstraaten	2015	The quality of economic evaluations of ultra-orphan drugs in Europe - a systematic review	Included	Systematic review of economic evaluations
A. R. Sequeira; E. Mentzakis; O. Archangelidi; F. Paolucci	2021	The economic and health impact of rare diseases: A meta-analysis	Included	Systematic review and meta-analysis of burden of disease
H. Short; T. Stafinski; D. Menon	2015	A National Approach to Reimbursement Decision-Making on Drugs for Rare Diseases in Canada? Insights from Across the Ponds	Included	Systematic review of rare disease decision-making policies Included because these policies can include economic evaluations
S. Simoens; M. Dooms	2011	How much is the life of a cancer patient worth? A pharmaco-economic perspective	Excluded	Excluded as this is not a systematic review in the technical sense "The aim is not to undertake a systematic review, in its usual technical sense, but to comment on particularly illustrative examples"

T. Thorat; P. J. Neumann; J. D. Chambers	2018	Hemophilia Burden of Disease: A Systematic Review of the Cost-Utility Literature for Hemophilia	Included	Systematic review of cost-utility analyses
M. Ursino; N. Stallard	2021	Bayesian approaches for confirmatory trials in rare diseases: Opportunities and challenges	Excluded	Concerned with clinical trials and how to mitigate the small patient numbers for rare diseases, but not economic evaluation. Thus, it was excluded
C. Watson; A. Barlev; J. Worrall; S. Duff; R. Beckerman	2019	The Economic Burden of Short-Term Adverse Events Associated with the CHOP Chemotherapy Regimen in Patients with Lymphoproliferative Disorders in the United States: A Comprehensive Literature Review	Excluded	Excluded. Conference abstract
D. Weidlich; P. Kefalas; J. F. Guest	2016	Healthcare costs and outcomes of managing beta-thalassemia major over 50 years in the United Kingdom	Included	Systematic review of cost and effectiveness
A. L. Woersching; M. E. Borrego; D. W. Raisch	2016	Assessing the Quality of Economic Evaluations of FDA Novel Drug Approvals: A Systematic Review	Included	Systematic review economic evaluations (Focus is on novel new drugs, some of which are orphan)

T. Zelei; N. D. Mendola; B. Elezbawy; B. Nemeth; J. D. Campbell	2021	Criteria and Scoring Functions Used in Multi-criteria Decision Analysis and Value Frameworks for the Assessment of Rare Disease Therapies: A Systematic Literature Review	Included	Systematic review of MCDA
T. Zelei; M. J. Molnar; M. Szegedi; Z. Kalo	2016	Systematic review on the evaluation criteria of orphan medicines in Central and Eastern European countries	Included	Systematic review of economic evaluations
B. M. Zimmermann; J. Eichinger; M. R. Baumgartner	2021	A systematic review of moral reasons on orphan drug reimbursement	Excluded	Excluded. Not about economic evaluation, but rather a study of the moral reasons for and against reimbursement of orphan drugs
Studies from literature search				
I. Pearson; B. Rothwell; A. Olaye; C. Knight	2018	Economic modelling considerations for rare diseases	Included	The methods are described as a series of targeted literature reviews, but they have been transparently reported, with the quality similar to other identified systematic reviews. Why it was decided that it did indeed meet the criteria for inclusion.

A. Kavvada; N. Syrigos; A. Charpidou; G. Kourlaba	2023	The role of epidemiology in health technology assessment and reimbursement of new medicines: A review	Excluded	Not on orphan drugs
L. Wiedmann; J. Cairns	2023	Review of economic modelling evidence from NICE appraisals of rare disease treatments for spinal muscular atrophy	Excluded	Relevant, but no systematic literature review component.
B. Johnston, P. Miller; A. Agarwal; S. Mulla; R. Khokhar; K. Oliveira; C. Hitchcock; B. Sadeghirad; M. Mohiuddin; N. Sekercioglu; M. Seweryn; M. Koperny; M. Bala; T. Adams-Webber; A. Granados; A. Hamed, M Crawford; A. van	2016	Limited responsiveness related to the minimal important difference of patient-reported outcomes in rare diseases	Excluded	Not on economic evaluation

der Ploeg; G. Guyatt				
D. Ollendorf; R. Chapman; S. Pearson	2018	Evaluating and valuing drugs for rare conditions: no easy answers	Excluded	Relevant, but no systematic literature review component.

Table 32: Summary of study characteristics.

	Citation details	Type of review	Objective(s)	Date range of database searches	Number of studies	JBI Checklist	Themes

#1	K. Abdallah; I. Huys; K. Claes; S. Simoens	Systematic review of conference abstracts and peer-reviewed literature for budget impact analyses	This research aims to evaluate the methodological quality of budget impact analyses for orphan drugs and to provide suggestions for future analyses	Unknown - Jan 2020	90	Moderate	Budget impact
#2	A. Angelis; D. Tordrup; P. Kanavos	Systematic literature review for cost of ilness studies	The objective of this study is to systematically review the relevant literature on the socio-economic burden of RDs and identify all costs, both direct and indirect, related to ten specifically identified RDs from the	Unknown - March 2011	77	Moderate	Cost-of-illness

			perspective of patients, families, and society.				
#3	A. Babac; K. Damm; J. M. Graf von der Schulenburg	Systematic review of orphan drug value dossiers in Germany	The research question is whether patient perspectives are considered as part of early benefit assessments for rare diseases and how patient perspectives are methodologically elicited and presented?	1st Jan 2011 - March 1 2019	81	Moderate	Patient perspectives

#4	A. Baran-Kooiker; M. Czech; C. Kooiker	Systematic review of MCDA and rare diseases	The research objective is to provide an overview of the current state of knowledge and latest developments in the field of MCDA in HTA for orphan drugs, to review existing models, their design characteristics, as well as to identify opportunities for further model improvement.	Unknown (no time limit) - January 2018	29	Moderate	MCDA

#5	T. Chen; J. Chen; C. Chen; H. Zheng; Y. Chen; M. Liu; B. Zheng	Systematic review of drugs and economic evaluation for rare hypertension in children	Unclear, but the study evaluated the economics of medication for PAH in children; provided a reference for patients to choose a more cost-effective treatment.	Unknown (from establishment of databases) - December 2019	2	Moderate	Economic Evaluation
#6	M. M. Cheng; S. D. Ramsey; E. B. Devine; L. P. Garrison; B. W. Bresnahan; D. L. Veenstra	Systematic review of clinical and economic evidence for oncology orphan drugs in United States	To systematically assess clinical and economic evidence for oncology orphan drugs marketed in the United States and to highlight the challenges and opportunities for evidence development	Unknown - June 2010	81	Moderate	Economic Evaluation

			within the pharmaceutical category.				
#7	A. Dhawan; M. W. Lawlor; G. V. Mazariegos; P. McKiernan; J. E. Squires; K. A. Strauss; D. Gupta; E. James; S. Prasad	Systematic review and meta-analysis for burden of disease	To more fully characterize the clinical, humanistic, and economic burden of Crigler–Najjar syndrome attributable to both chronic hyperbilirubinemia and its extant treatments.	Unknown - December 2017	28	Moderate	Cost-of-illness

#8	E. Faulkner; D. S. Spinner; M. Ringo; M. Carroll	A systematic review of value assessment processes including one hundred health technology assessments	To consider what is transformative, challenges for transformative therapies, and downstream health ecosystem effects.	2012 - 2018	222	Moderate	Economic Evaluation
#9	C. Friedmann; P. Levy; P. Hensel; M. Hiligsmann	A systematic review of MCDA for orphan drugs	The objective of this study was to systematically review existing evidence on the application of MCDA in the assessment and appraisal of orphan drugs worldwide.	Jan 2000 - April 2017	13	Moderate	MCDA

#10	L. Garcia-Perez; R. Linertova; C. Valcarcel-Nazco; M. Posada; I. Gorostiza; P. Serrano-Aguilar	Systematic review of cost-of-illness studies for rare diseases.	The aim of the review was to overview the cost-of-illness studies conducted in rare diseases.	Jan 2007 - December 2018	63	Moderate	Cost-of-illness
#11	S. Gruhn; J. Witte; W. Greiner; O. Damm; M. Dietzsch; R. Kramer; M. Knuf	Systematic review of epidemiological and economic parameters	To describe the key epidemiological and economic parameters of IMD in Germany to support national decision-making process for implementing enhanced prevention measures	Jan 1990 - May 2020	36	Moderate	Cost-of-illness

#12	R. W. Jakes; N. Kwon; B. Nordstrom; R. Goulding; K. Fahrbach; J. Tarpey; M. K. Van Dyke	Systematic review and meta-analysis of prevalence, incidence, and burden of disease	To describe the available published evidence on the incidence, prevalence, morbidity, and HRCU.	Unknown - June 2019	100	Moderate	Cost-of-illness
#13	T. A. Kanters; C. de Sonneville-Koedoot; W. K. Redekop; L. Hakkaart	Systematic review of evidence for high priced orphan drugs	To review the available evidence on clinical effectiveness, cost-effectiveness, and budget impact for orphan drugs	Unknown - August 2012	338	Moderate	Economic Evaluation: budget impact, effectiveness, and cost-effectiveness
#14	E. I. Katsigianni; P. Petrou	Systematic review of economic evaluations	To assess the economic profile of Enzyme replacement therapy	Unknown - March 2022	8	Moderate	Economic evaluation

#15	N. Knoble; G. Nayroles; C. Cheng; B. Arnould	Systematic review of patient reported outcomes	To investigate and describe PRO measurement challenges, and what has been done to address the issue, in the specific context of Cushing's Syndrome, a rare disease	Unknown - December 2016	36	Moderate	Patient reported outcomes
#16	C. S. Kwon; P. Daniele; A. Forsythe; C. Ngai	Systematic review of QoL and cost-of-illness	To gather and analyse published evidence on epidemiology, burden, and current treatments patterns of IgAN	Jan 2010 - June 2020	123	Moderate	Epi, QoL and cost-of-illness

#17	P. Lasalvia; L. Prieto-Pinto; M. Moreno; J. Castrillon; G. Romano; N. Garzon-Orjuela; D. Rosselli	Scoping review of MCDA for orphan drugs	To explore the use of MCDA in the assessment of orphan drugs	Unknown - May 2018	16	Low	MCDA
#18	C. Lee; A. Lam; T. Kangappaden; P. Olver; S. Kane; D. Tran; E. Ammann	Systematic review for clinical and economic evidence	To evaluate the current evidence-base in AL amyloidosis	Unknown (not restricted)	84	High	Clinical, QoL and costs
#19	L. P. Leonart; H. H. L. Borba; V. L. Ferreira; B. S. Riveros; R. Pontarolo	Systematic review of cost-effectiveness analyses	To describe and critically appraise the comparative economic studies on acromegaly treatments	Unknown (no restriction) - May 2018	16	High	Economic evaluation

#20	M. Mohammadshahi; A. Olyaeemanesh; E. Ehsani-Chimeh; M. Mobinizadeh; Z. Fakoorfard; A. Akbari Sari; M. Aghighi	Scoping review on criteria for economic evaluation	To identify different criteria for priority setting of rare diseases to help policy makers in making evidence-informed decisions	2020 - 2021	11	Moderate	Criteria for economic evaluation
#21	I. J. Onakpoya; E. A. Spencer; M. J. Thompson; C. J. Heneghan	Systematic review of effectiveness and costs	To evaluate the clinical effectiveness of orphan drugs that have been granted marketing licenses in Europe, determine the annual costs of each drug, compare the cost of branded orphan drugs against their	Unknown	74 drugs included Number of studies not clear	High	Effectiveness and costs

			generic equivalents, and explore any relationships between orphan drug disease prevalence and annual costs				
#22	T. Park; S. K. Griggs; D. C. Suh	Systematic review of cost-effectiveness analyses	To systematically review cost-effectiveness evidence of monoclonal antibody-based orphan drugs	Unknown - June 2014	16	High	Economic evaluation
#23	E. Picavet; D. Cassiman; S. Simoens	Systematic review of cost-utility analyses	To provide an overview of the available evidence on cost-effectiveness of orphan drugs	Unknown - December 2013	45 Cost-utility analyses	Moderate	Economic evaluation

#24	L. Querol; M. Crabtree; M. Herepath; E. Priedane; I. Viejo; S. Agush; P. Sommerer	Systematic review of burden of disease	To provide a comprehensive literature review of the burden of illness for chronic inflammatory demyelinating polyneuropathy (CIDP)	May 2009 - May 2019	66 Although Figure 1 shows n=64	Moderate	Burden of disease
#25	M. Raut; G. Singh; I. Hiscock; S. Sharma; N. Pilkhwal	Systematic review of burden of disease	To understand and quantify the burden of disease in patients with relapsed-refractory classical Hodgkin lymphoma	Jan 2001 - February 2020 for disease burden and treatment patterns 2015 - 2020 for economic evaluations	144	Moderate	Burden of disease

#26	J. Rice; A. White; L. Scarpati; M. Philbin; G. Wan; W. Nelson	Systematic review of burden of disease	To conduct a systematic literature review on the burden of IEDs with a focus on recent studies reporting clinical and economic outcomes to highlight the clinical and subsequent economic burden of these heterogenous and complex diseases and capture unmet needs under the current treatment landscape by which that burden may be reduced.	Jan 2009 - August 2016	39	Moderate	Burden of disease

#27	M. Schlander; C. M. Dintsios; A. Gandjour	Systematic review of budget impact	To review recent studies reporting health care expenditures for orphan medicinal products in Europe and to contribute to our understanding of the cost drivers of non-oncology orphan medicinal products by means of an empirical analysis in Germany.	Unknown - December 2016	8 in publication and Figure 1 9 in abstract text	Moderate	Budget impact
#28	Y. Schuller; C. E. Hollak; M. Biegstraaten	Systematic review of economic evaluations	To assess the economic evaluations of ultra-orphan drugs marketed in the EU that have been performed	Unknown - April 2015	16	High	Economic evaluation

			so far with specific focus on the methods used and the quality of the studies.				
#29	A. R. Sequeira; E. Mentzakis; O. Archangelidi; F. Paolucci	Systematic review and meta-analysis for burden of disease	To fill his gap (economic impact of rare diseases) by evaluating the economic impact of rare diseases and exploring potential factors associated with the societal burden of rare diseases.	January 2010 - February 2017 in abstract January 2010 - January 2017 in publication	19	High	Burden of disease

#30	H. Short; T. Stafinski; D. Menon	Systematic review of rare disease decision-making policies	To develop an inventory of existing policies and processes for making coverage decisions on drugs for rare diseases around the world.	2004 - Feb 2013 (From Appendix A: Embase by Ovid)	Unclear	Moderate	Policies
#31	T. Thorat; P. J. Neumann; J. D. Chambers	Systematic review of cost-utility analyses	To systematically review the cost-utility analysis literature of haemophilia treatments and demonstrate the challenges in producing cost-utility evidence compared with other rare diseases	2000 - 2015	11	High	Economic evaluation

#32	D. Weidlich; P. Kefalas; J. F. Guest	Systematic review of cost and effectiveness	To provide an update to the earlier analysis and estimate the incidence-based cost of beta-thalassemia major to the United Kingdom's National Health Service (NHS) over a period of 50 years, based on current practice	2009 - Unknown (Study submitted 2015, published 2016)	18	Moderate	Economic evaluation
#33	A. L. Woersching; M. E. Borrego; D. W. Raisch	Systematic review of economic evaluations	To assess the quality of the economic evaluation literature considering novel drug approvals. Our aims were to appraise	1946 - December 2011 Considering FDA approved drugs from 2008 - 2009	36	High	Economic evaluation

			individual aspects of quality, derive a summary score for each article, and test for associations between study features and quality				
#34	T. Zelei; N. D. Mendola; B. Elezbawy; B. Nemeth; J. D. Campbell	Systematic review of MCDA	To investigate the criteria and scoring functions applied in value frameworks and MCDA tools relevant to the evaluation of rare disease therapies	January 2013 - October 2019	15	Moderate	MCDA

#35	T. Zelei; M. J. Molnar; M. Szegedi; Z. Kalo	Systematic review of economic evaluations (value drivers)	To explore and summarize the recent scientific evidence on value drivers related to the health technology assessments of orphan drugs with a special focus on the perspective of third-party payers in Central and Eastern European countries	January 2000 - April 2015	87	Moderate	Economic evaluation

#36	I. Pearson; B. Rothwell; A. Olaye; C. Knight	Systematic review (a series of targeted reviews) of economic evaluations (economic models)	To identify challenges that affect the feasibility and rigor of economic models in rare diseases and strategies that manufacturers have employed in health technology assessment submissions to demonstrate the value of new orphan products that have limited study data	Jan 2007 - Feb 2017	19	Moderate	Economic evaluation

Table 33: Answers to each quality assessment question for all studies.

	Citation details	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**Q11**	**Score**	**Quality**

	Answers: **(Y = Yes, No = No, U = Unclear and NA = Not applicable)**													
#1	K. Abdallah; I. Huys; K. Claes; S. Simoens	Y	Y	Y	Y	Y	U	U	NA	No	Y	Y	7	Moderate
#2	A. Angelis; D. Tordrup; P. Kanavos	Y	Y	Y	U	U	No	No	NA	No	Y	Y	5	Moderate
#3	A. Babac; K. Damm; J. M. Graf von der Schulenburg	Y	Y	Y	Y	U	U	U	NA	No	Y	Y	6	Moderate
#4	A. Baran-Kooiker; M. Czech; C. Kooiker	Y	Y	Y	Y	NA	No	Y	NA	No	Y	Y	7	Moderate
#5	T. Chen; J. Chen; C. Chen; H. Zheng; Y. Chen; M. Liu; B. Zheng	U	Y	Y	Y	Y	Y	Y	NA	No	Y	Y	7	Moderate
#6	M. M. Cheng; S. D. Ramsey; E. B. Devine; L. P. Garrison; B. W. Bresnahan; D. L. Veenstra	Y	Y	Y	Y	Y	No	No	NA	No	U	U	5	Moderate
#7	A. Dhawan; M. W. Lawlor; G. V. Mazariegos; P. McKiernan; J. E. Squires; K. A. Strauss; D. Gupta; E. James; S. Prasad	Y	Y	U	Y	Y	No	No	NA	No	U	U	4	Moderate
#8	E. Faulkner; D. S. Spinner; M. Ringo; M. Carroll	Y	Y	Y	Y	Y	U	No	NA	No	Y	Y	7	Moderate
#9	C. Friedmann; P. Levy; P. Hensel; M. Hiligsmann	Y	Y	Y	Y	NA	No	Y	NA	No	Y	Y	7	Moderate
#10	L. Garcia-Perez; R. Linertova; C. Valcarcel-Nazco; M. Posada; I. Gorostiza; P. Serrano-Aguilar	Y	Y	Y	Y	No	No	No	NA	No	Y	Y	6	Moderate

#11	S. Gruhn; J. Witte; W. Greiner; O. Damm; M. Dietzsch; R. Kramer; M. Knuf	Y	U	Y	Y	No	No	Y	NA	No	Y	Y	6	Moderate
#12	R. W. Jakes; N. Kwon; B. Nordstrom; R. Goulding; K. Fahrbach; J. Tarpey; M. K. Van Dyke	Y	Y	Y	Y	Y	U	Y	Y	No	NA	NA	7	Moderate
#13	T. A. Kanters; C. de Sonneville-Koedoot; W. K. Redekop; L. Hakkaart	Y	Y	Y	Y	NA	NA	Y	NA	No	Y	Y	7	Moderate
#14	E. I. Katsigianni; P. Petrou	Y	U	Y	Y	Y	No	No	NA	No	Y	Y	6	Moderate
#15	N. Knoble; G. Nayroles; C. Cheng; B. Arnould	Y	Y	Y	Y	U	U	Y	NA	No	Y	Y	7	Moderate
#16	C. S. Kwon; P. Daniele; A. Forsythe; C. Ngai	Y	Y	Y	Y	U	U	U	Y	No	Y	Y	7	Moderate
#17	P. Lasalvia; L. Prieto-Pinto; M. Moreno; J. Castrillon; G. Romano; N. Garzon-Orjuela; D. Rosselli	Y	Y	U	Y	U	U	No	NA	No	U	U	3	Low
#18	C. Lee; A. Lam; T. Kangappaden; P. Olver; S. Kane; D. Tran; E. Ammann	Y	Y	Y	Y	Y	Y	U	NA	No	Y	Y	8	High
#19	L. P. Leonart; H. H. L. Borba; V. L. Ferreira; B. S. Riveros; R. Pontarolo	Y	Y	Y	Y	Y	Y	Y	NA	No	Y	Y	9	High
#20	M. Mohammadshahi; A. Olyaeemanesh; E. Ehsani-Chimeh; M. Mobinizadeh; Z. Fakoorfard; A. Akbari Sari; M. Aghighi	Y	U	Y	Y	U	U	No	NA	No	Y	Y	5	Moderate
#21	I. J. Onakpoya; E. A. Spencer; M. J. Thompson; C. J. Heneghan	Y	Y	Y	Y	Y	Y	Y	U	No	Y	Y	9	High
#22	T. Park; S. K. Griggs; D. C. Suh	Y	Y	Y	Y	Y	Y	Y	NA	Y	Y	Y	10	High

#23	E. Picavet; D. Cassiman; S. Simoens	Y	Y	Y	Y	Y	U	Y	U	Y	U	U	7	Moderate
#24	L. Querol; M. Crabtree; M. Herepath; E. Priedane; I. Viejo; S. Agush; P. Sommerer	U	Y	Y	Y	No	No	Y	NA	Y	Y	Y	7	Moderate
#25	M. Raut; G. Singh; I. Hiscock; S. Sharma; N. Pilkhwal	Y	Y	Y	Y	U	No	Y	U	No	Y	Y	7	Moderate
#26	J. Rice; A. White; L. Scarpati; M. Philbin; G. Wan; W. Nelson	Y	Y	U	Y	No	No	Y	NA	No	Y	Y	6	Moderate
#27	M. Schlander; C. M. Dintsios; A. Gandjour	Y	Y	Y	Y	No	No	No	NA	No	U	U	4	Moderate
#28	Y. Schuller; C. E. Hollak; M. Biegstraaten	Y	Y	Y	Y	Y	Y	Y	NA	Y	Y	Y	10	High
#29	A. R. Sequeira; E. Mentzakis; O. Archangelidi; F. Paolucci	Y	Y	Y	Y	U	U	Y	Y	No	Y	Y	8	High
#30	H. Short; T. Stafinski; D. Menon	Y	Y	Y	Y	U	U	Y	NA	Y	U	U	6	Moderate
#31	T. Thorat; P. J. Neumann; J. D. Chambers	Y	Y	Y	Y	U	U	Y	NA	Y	Y	Y	8	High
#32	D. Weidlich; P. Kefalas; J. F. Guest	Y	U	Y	Y	U	U	U	U	No	Y	Y	5	Moderate
#33	A. L. Woersching; M. E. Borrego; D. W. Raisch	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	U	9	High
#34	T. Zelei; N. D. Mendola; B. Elezbawy; B. Nemeth; J. D. Campbell	Y	Y	Y	Y	U	U	Y	NA	No	Y	Y	7	Moderate
#35	T. Zelei; M. J. Molnar; M. Szegedi; Z. Kalo	Y	Y	Y	Y	No	U	U	NA	Y	Y	Y	7	Moderate
#36	I. Pearson; B. Rothwell; A. Olaye; C. Knight	Y	Y	Y	Y	U	No	U	NA	No	Y	Y	6	Moderate

Comments: #23 Q5; EMBASE studies were not appraised, but the criteria used in Tuft registry were reasonable. Q6; Appraisals from the Tufts registry were modified. EMBASE studies were not appraised. #24 Q3; Supplement material 1 is named online resource 1 in paper. #26 Q3; Could not identify the supplementary appendix A, not even from the in-text hyperlink.

Table 34: Issues for economic evaluation of rare diseases.

	Systematic reviews that considered:	
	A specific rare disease	**Multiple rare diseases**
Issues with evidence		
Natural history of disease	• Prevalence and incidence: ○ are limited by lack of harmonised disease and diagnostic codes, or data to allow for their calculation (Kwon *et al.*, 2021). ○ vary with age, gender, diagnostic criteria, and study sample. Limited by number of publications (Querol *et al.*, 2021). ○ are limited by reporting, and there is insufficient data on treatment eligible / ineligible cohorts (Raut *et al.*, 2022). ○ limited literature reporting incidence and prevalence (Jakes *et al.*, 2021). • Diagnosis: ○ misdiagnosis can occur and lead to inappropriate treatment (Querol *et al.*, 2021).	• Epidemiological uncertainties: ○ uncertainty on patient population and subpopulations is frequently cited in HTA (Faulkner *et al.*, 2019). ○ slow disease progression or chronic nature of rare diseases make clinical trials insufficient for the purpose of informing cost-effectiveness analyses (Schuller, Hollak and Biegstraaten, 2015) ○ limited understanding of epidemiology of rare diseases and impact on clinical trial designs (Pearson *et al.*, 2018). ○ prevalence is difficult to estimate due to lack of epidemiological research (Sequeira *et al.*, 2021). • Diagnosis: ○ physicians encounter few cases of rare diseases, which lead to delayed diagnosis or misdiagnosis (Kanters *et al.*, 2013).

	○ delays in diagnosis can be a result of multiple specialists involved in process (Lee *et al.*, 2022). ○ between-study heterogeneity on diagnosis (Jakes *et al.*, 2021)	○ efficacy measurement is hindered by diagnostic capacities and capabilities (Zelei *et al.*, 2016). ○ few clinical experts and specialist centers (Pearson *et al.*, 2018). ○ difficult diagnosis (Sequeira *et al.*, 2021) • Registries: ○ patient registries may suffer from challenges pertaining to data ownership, comparators, and bias (Pearson *et al.*, 2018).
Clinical effectiveness	• Clinical trials: ○ small number of clinical trials, short duration of trials, few included subjects, and missing separation of paediatric and adult data, which led to a scarcity of paediatric data. (Chen *et al.*, 2021). ○ Paucity of clinical evidence, e.g., published data from clinical trials (Lee *et al.*, 2022). • Studies on effectiveness: ○ interventions, comparators, time points and methods are heterogenous. Studies may have limited data on ability to predict or varying definitions of treatment response (Querol *et al.*, 2021).	• Clinical trials: ○ for rare disease are mostly multi-centre international studies and include smaller study populations as compared to clinical trials for ordinary diseases. There may be only one chance to get a study right, because of limited treatment naïve patients. It is unethical to set up placebo-controlled trials if this means withholding treatment, especially for children (Kanters *et al.*, 2013). ○ suffer from few subjects, limited understanding of epidemiology, limited experts and centres, patient heterogeneity, inadequate power, imbalance between trial arms, lack of control arm and comparator data, no “hard” outcomes, no validated instruments, premature

- lack of consensus on treatment regimens, especially for subsequent lines of therapy after initial treatment (Raut *et al.*, 2022).
- variation in study design, population and therapies, and failure to control for disease severity (Rice *et al.*, 2018).
- wide variation in pharmacological management, thus limiting conclusions on efficacy and safety (Dhawan *et al.*, 2019).

termination, and limited trial duration (Pearson *et al.*, 2018).

- suffer from small sample sizes, difficult to confirm added clinical value, e.g., when randomisation or control arm is missing, limited trial duration, use of surrogate outcomes, limited understanding of epidemiology and disease management, and delayed diagnosis because of diagnostic capabilities (Zelei *et al.*, 2016).
- there is large variation in clinical outcomes across diseases, in time-points for measurement, inconsistent reporting, which collectively limit conclusions on effectiveness and safety. This review rated all evidence low to moderate quality (Onakpoya *et al.*, 2015).
- concerns on the linkage between surrogate markers, patient relevance and final health outcomes, because surrogates are not adequately described compared to diseases with higher prevalence. Uncertainty on short and long-term effects, insufficient comparator data, and study designs (Faulkner *et al.*, 2019).

- Effectiveness studies:

- limited long-term data, e.g., post marketing data on safety and efficacy (Cheng *et al.*, 2012).
- small patient numbers (Sequeira *et al.*, 2021).

Costs	• Non-existent or limited cost-of-illness studies: ○ data on economic burden and healthcare resource use are scarce (Gruhn *et al.*, 2022). ○ limited evidence is published on the economic burden and consequences (Kwon *et al.*, 2021). ○ very little recent rigorous or original research on economic burden. Few papers reported direct costs, but not indirect costs (societal) (Rice *et al.*, 2018). ○ few studies on healthcare resource use and management patterns (Weidlich, Kefalas and Guest, 2016). ○ no data on economic burden was found (Dhawan *et al.*, 2019). ○ no articles reporting indirect costs (Lee *et al.*, 2022). ○ no studies reported non-medical-direct or indirect costs (Leonart *et al.*, 2018). ○ studies were generally concerned with direct treatment costs with limited information on indirect or intangible costs (Raut *et al.*, 2022). ○ limited publications impair current assessment of burden of disease (Querol *et al.*, 2021).	• Non-existent or limited cost-of-illness studies: ○ costs-of-illness data were non-existent or very limited, even for conditions with relative high prevalence. Data availability appears correlated with treatment availability rather than disease rarity or severity. Limited indirect costs (Angelis, Tordrup and Kanavos, 2015). ○ paucity of cost-of-illness studies. Data on costs in aggregated format and/or primary data are rarely available. Why patient-reported, claims, or registry data are used No studies followed recommendations to handle uncertainty in the methods used, human-capital and friction method, respectively (Garcia-Perez *et al.*, 2021). ○ limited information on direct, indirect costs and costs associated with poorly controlled disease (Pearson *et al.*, 2018). • Transferability: ○ methodological variations prevent any detailed comparison of conditions (Angelis, Tordrup and Kanavos, 2015).

	• Transferability: ○ transferability to other countries is limited, varying data sources may explain in-country variation for results (Gruhn *et al.*, 2022). ○ these findings cannot be transferred across countries because of healthcare system variability (Querol *et al.*, 2021).	○ methods vary, e.g., scope, population characteristics, and perspective. Thus, making comparisons impossible (Garcia-Perez *et al.*, 2021). ○ local costs differences across countries and discount rates were a source of variability (Katsigianni and Petrou, 2022)
Quality of life	• Non-existent or limited quality-of-life studies: ○ limited published evidence on health-related quality of life (Kwon *et al.*, 2021). ○ no data on humanistic burden, except for isolated anecdotal comments (Dhawan *et al.*, 2019). ○ studies are small, not randomised or controlled, and their conclusions limited (Raut *et al.*, 2022). ○ none of the identified studies included utility values or mapping algorithms, despite the inclusion of measures that allowed for it (Lee *et al.*, 2022). ○ methods to elicit utility weights are insufficiently explained (Thorat, Neumann and Chambers, 2018). ○ quality-of-life trajectories are often not measured across more than two time-points, which is	• Non-existent or limited quality-of-life studies: ○ missing quality of life data can result in assumptions of equal health utilities for parameters (Katsigianni and Petrou, 2022). ○ although quality-of-life data are integrated through various methods, there is a scarcity of patient-preference methods. Quantitative methods such as choice experiments or conjoint analysis may not be feasible for rare diseases (Babac, Damm and Graf von der Schulenburg, 2019). ○ limited focus on social, financial, and emotional burden for informal caregivers (Sequeira *et al.*, 2021).

	insufficient for diseases with a non-linear quality-of-life progression and for statistical mediation analysis. Mapping algorithms are used to facilitate modelling. Generic-preference-based measures are readily available and easy to administer but may miss important quality-of-life attributes for rare diseases. Disease-specific measures might suffer from validation issues (Knoble *et al.*, 2018).	
Issues with modelling		
Cost effectiveness	• Scarcity of cost-effectiveness analyses: ○ no cost-effectiveness analyses or cost-utility analyses identified (Lee *et al.*, 2022). • Assumptions ○ patient-level data might not be available. Use of assumptions on mortality, efficacy, and complications and expert opinions. In a review, a model failed to account for patient characteristics and disease severity (Weidlich, Kefalas and Guest, 2016). • Reporting: ○ cost-effectiveness studies often fail to report discount rates and sensitivity analyses. Overall,	• Scarcity of cost-effectiveness analyses: ○ There is a remarkable absence of pharmacoeconomic evidence(Kanters *et al.*, 2013). ○ cost-effectiveness analyses are not performed and considered unachievable because of limited evidence, high drug costs, and inability to measure effects. Use of patient-level simulations might be limited by low patient numbers and unreliable estimates for effects, symptoms, and complications (Schuller, Hollak and Biegstraaten, 2015). ○ scarcity of evidence makes cost-effectiveness analyses difficult or almost impossible. Uncertainty can hinder reimbursement until more evidence becomes available (Zelei *et al.*, 2016).

studies are rated as having moderate quality (Leonart *et al.*, 2018).

- Input-parameter heterogeneity:
  - comparison is difficult because of different perspective, time horizon, and patient population (Weidlich, Kefalas and Guest, 2016).
  - Lack of consistency in reported treatment patterns, country settings, discount rates, sample population, and time horizons (Thorat, Neumann and Chambers, 2018).
  - the findings were affected by different country settings (Raut *et al.*, 2022).
  - Analyses were performed in different countries and years resulting non-comparable heterogeneous scenarios (Leonart *et al.*, 2018).
- Bias:
  - Most studies are industry sponsored, which evidence suggests report more favourable results. Note that this comparison was not made in this study, because of few studies (Thorat, Neumann and Chambers, 2018).

  - data are extrapolated for long-term outcomes using modelling techniques due to clinical-data limitations. Common shortcomings are failure to report funding source, discounting, time horizons or time horizons that adequately reflect diseases (Cheng *et al.*, 2012).
- Bias:
  - no study explicitly mentioned the magnitude of bias (Park, Griggs and Suh, 2015).
  - publication bias may be prominent for orphan drugs as pharmaceutical companies have decreased incentives to publish because these drugs are rarely cost-effective (Kanters *et al.*, 2013).
  - a larger number of cost-effectiveness analyses in United Kingdom and Netherlands is explained by a post-marketing obligation to publish (Schuller, Hollak and Biegstraaten, 2015).
  - studies use data from potentially biased sources without explicit explanation of direction and magnitude of bias. The paucity of cost-effectiveness analyses may reflect publication bias or evidence limitations (Cheng *et al.*, 2012).
  - majority of studies are funded by industry. They have significant lower incremental cost-effectiveness ratios

		compared to those published by academia and foundations. This may be due to publication bias or selection bias (Picavet, Cassiman and Simoens, 2015). ○ sponsorship bias may be an issue because most papers were industry sponsored. They may have sponsored or co-authored publications in competitive markets (Woersching, Borrego and Raisch, 2016). • Transferability: ○ cost-effectiveness results are not consistent. Medication and healthcare settings need consideration when interpreting results (Park, Griggs and Suh, 2015). ○ remarkable result differences which are attributed to assumptions on quality of life and survival (Schuller, Hollak and Biegstraaten, 2015). ○ cost-effectiveness comparisons are complicated by differences in settings (Picavet, Cassiman and Simoens, 2015). ○ transferability is highly limited between countries and needs careful consideration (Zelei *et al.*, 2016).
Budget impact	No systematic reviews identified	• Scarcity: ○ budget-impact analyses are rarely published (Kanters *et al.*, 2013).

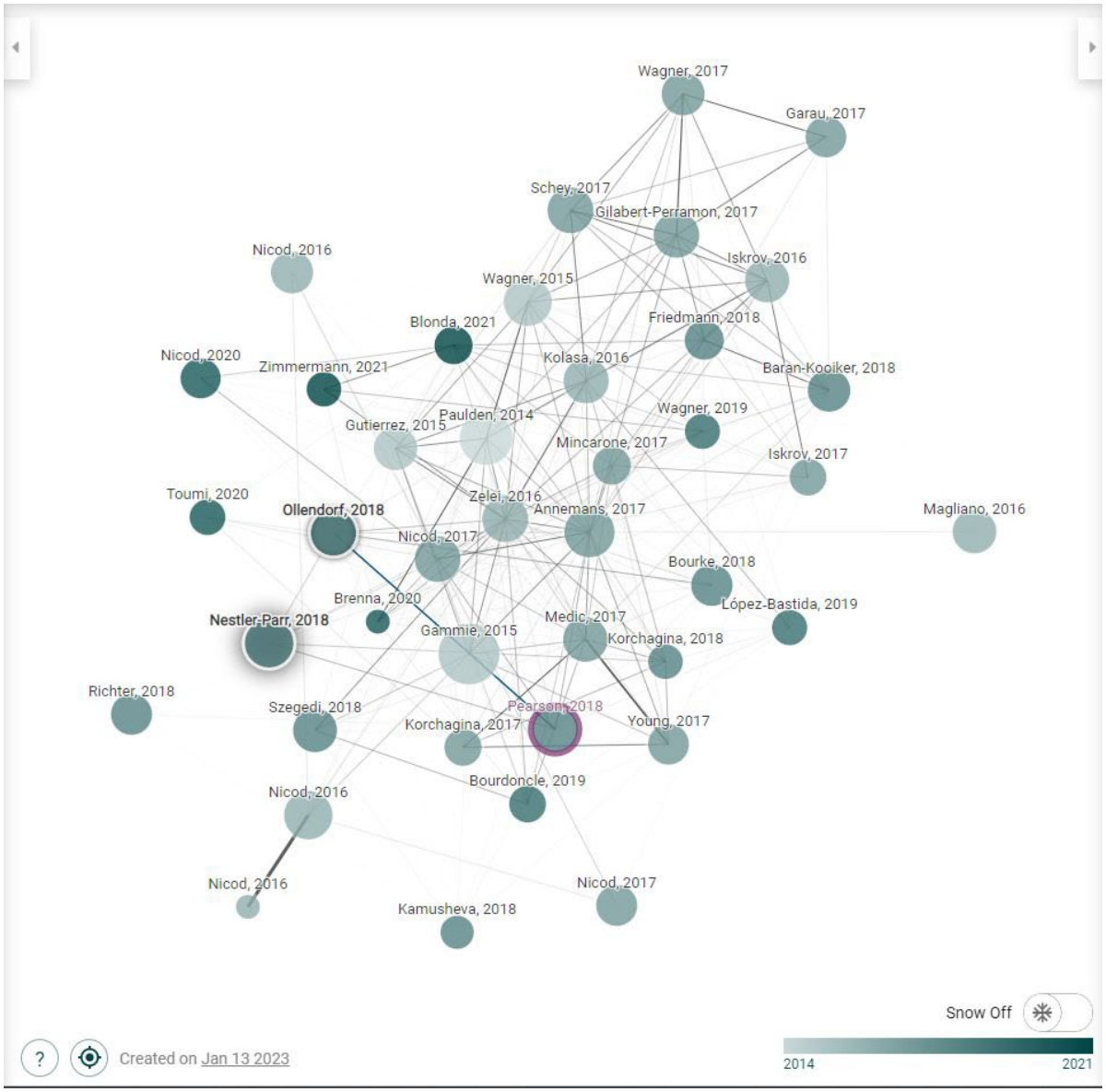
		• Assumptions ○ budget-impact analyses are mostly performed in high-income countries. They have low quality and results are rarely validated. Moreover, they show poor adherence to guidelines. Reported use of assumptions. For example, on, target population, population size, intervention, comparators, costs, and market uptake (Abdallah *et al.*, 2021). • Reporting ○ changing market dynamics. For example, market-exclusivity status for drugs may change during the time-period of analysis, thereby leading to different estimates. Budget-impact analyses do not report items such as costs of drug-related services (monitoring, tests, or counselling), treatment of side-effects, life-extension, but also savings from reduction in morbidity (Schlander, Dintsios and Gandjour, 2018).
Issues with estimating value / reimbursement		
Thresholds	No systematic reviews identified	• Transferability: ○ the variation in cost-effectiveness thresholds across countries mean that results are likely country specific (Park, Griggs and Suh, 2015).

Value frameworks	No systematic reviews identified	• Value frameworks: ○ reimbursement processes have transparency and consistency issues, thus, limited country comparability (Baran-Kooiker, Czech and Kooiker, 2018). ○ most countries require an economic evaluation and budget impact model incl. estimates on prevalence, incidence, and patient numbers. Requirements vary across countries, but common factors are clinical benefit, cost-effectiveness, and budget impact (Short, Stafinski and Menon, 2015). • Reference pricing ○ reference-pricing may prevent orphan-drug launches in low-income European countries (Zelei *et al.*, 2016).
MCDA	No systematic reviews identified	• Transparency and consistency: ○ MCDA is getting increased interest for orphan drugs, but no consistent recommendations for decision-making could be observed. More search is needed (Friedmann *et al.*, 2018). ○ few studies provide recommendations on scoring functions, which is a significant limitation, because of uncertainty and reduced transparency (Zelei *et al.*, 2021).

	○ multiple criteria for priority setting can lead to country-dependent priorities (Mohammadshahi *et al.*, 2022). ○ the use of MCDA frameworks is still emerging for orphan drugs (Lasalvia *et al.*, 2019).

I.II Appendix: Figures

Figure 18: Force-directed graph of references created through https://www.connectedpapers.com.



Description: Node size is the number of citations, node colour is publishing year, and similar papers have strong connecting lines and cluster together.

Figure 19: Screenshot of the JBI critical appraisal checklist for systematic reviews and research syntheses. Included with permission from the authors.



JBI CRITICAL APPRAISAL CHECKLIST FOR SYSTEMATIC REVIEWS AND RESEARCH SYNTHESES

Reviewer_____ Date_____

Author_____ Year_____ Record Number_____

		Yes	No	Unclear	Not applicable
1.	Is the review question clearly and explicitly stated?	☐	☐	☐	☐
2.	Were the inclusion criteria appropriate for the review question?	☐	☐	☐	☐
3.	Was the search strategy appropriate?	☐	☐	☐	☐
4.	Were the sources and resources used to search for studies adequate?	☐	☐	☐	☐
5.	Were the criteria for appraising studies appropriate?	☐	☐	☐	☐
6.	Was critical appraisal conducted by two or more reviewers independently?	☐	☐	☐	☐
7.	Were there methods to minimize errors in data extraction?	☐	☐	☐	☐
8.	Were the methods used to combine studies appropriate?	☐	☐	☐	☐
9.	Was the likelihood of publication bias assessed?	☐	☐	☐	☐
10.	Were recommendations for policy and/or practice supported by the reported data?	☐	☐	☐	☐
11.	Were the specific directives for new research appropriate?	☐	☐	☐	☐

Overall appraisal: Include ☐ Exclude ☐ Seek further info ☐

Comments (Including reason for exclusion)

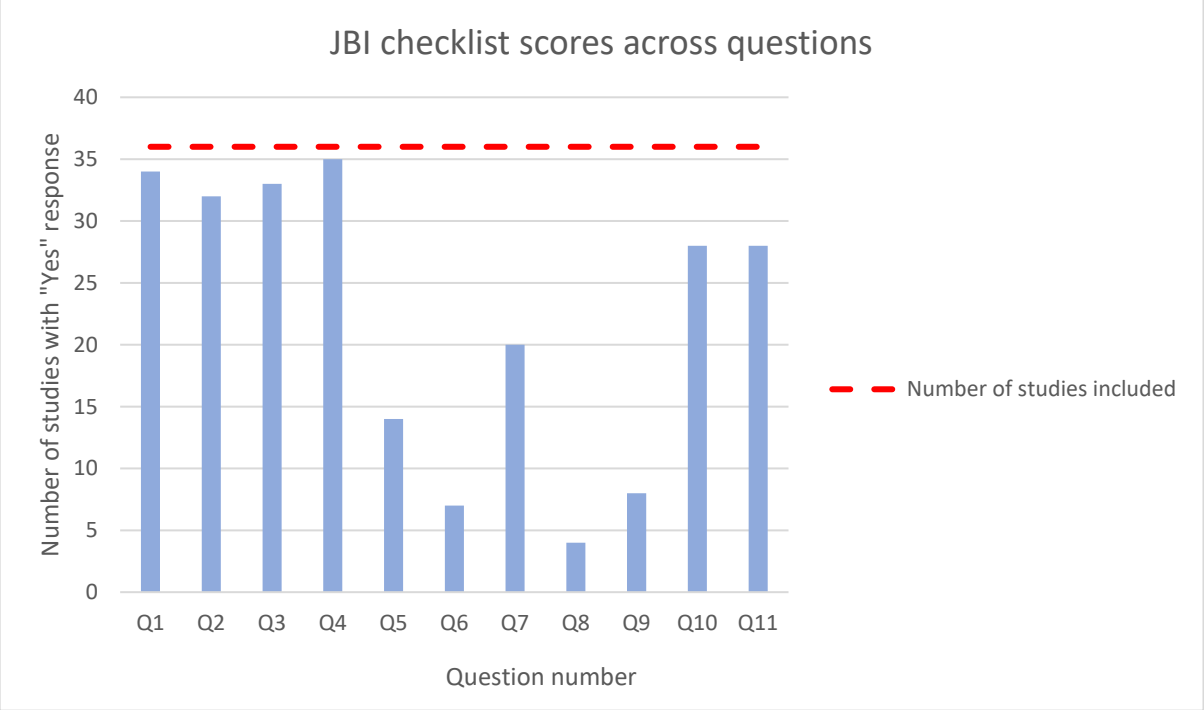
_____

_____

_____

© JBI, 2020. All rights reserved. JBI grants use of these tools for research purposes only. All other enquiries should be sent to jbisynthesis@adelaide.edu.au.

Figure 20: JBI checklist scores per question. The dotted red line represents the number of studies included.



II Appendix (Cost-effectiveness modelling early health technology assessment)

Appendix B is divided by two headings: tables and figures.

II.I Appendix: Tables

Table 35:MEDLINE (1946 to May 06, 2024) systematic searches on Ovid platform for economic models and multiple system atrophy performed on the 7th of May 2024.

Topic	Number	Search	Hits
Economic evaluations and Models (incl. cost analyses)	1	Economics/	27533
	2	exp "Costs and Cost Analysis"/	270270
	3	Economics, Nursing/	4013
	4	Economics, Medical/	9279
	5	Economics, Pharmaceutical/	3133
	6	exp Economics, Hospital/	25831
	7	Economics, Dental/	1922
	8	exp "Fees and Charges"/	31441
	9	exp Budgets/	14207
	10	budget*.ti,ab,kf.	37902

	11	(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ti,kf.	295331
	12	(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ab. /freq=2	406230
	13	(cost* adj2 (effective* or utilit* or benefit* or minimi* or analy* or outcome or outcomes)).ab,kf.	224612
	14	(value adj2 (money or monetary)).ti,ab,kf.	3202
	15	exp models, economic/	16300
	16	economic model*.ab,kf.	4440
	17	markov chains/	16134
	18	markov.ti,ab,kf.	30733
	19	monte carlo method/	32826
	20	monte carlo.ti,ab,kf.	62995
	21	exp Decision Theory/	13626
	22	(decision* adj2 (tree* or analy* or model*)).ti,ab,kf.	42583
	23	or/1-22	945427
Multiple system atrophy	24	exp Multiple System Atrophy/ or system atrophy.ti,ab,kf.	6762
Total	25	23 and 24	43

Table 36: MEDLINE (1946 to May 6, 2024) systematic searches on Ovid platform for health utilities and multiple system atrophy performed on the 7th of May 2024.

Topic	Number	Search	Hits
Health Utilities / Quality of life	1	"Value of Life"/	5824
	2	Quality of Life/	287639
	3	quality of life.ti,kf.	124851
	4	((instrument or instruments) adj3 quality of life).ab.	4053
	5	Quality-Adjusted Life Years/	16350
	6	quality adjusted life.ti,ab,kf.	18592
	7	(qaly* or qald* or qale* or qtime* or life year or life years).ti,ab,kf.	30384
	8	Disability-Adjusted Life Years/	271
	9	disability adjusted life.ti,ab,kf.	6127
	10	Healthy Life Expectancy/	79
	11	(daly* or disability free life expectanc* or haly* or health* life expectanc*).ti,ab,kf.	7269
	12	(sf36 or sf 36 or short form 36 or shortform 36 or short form36 or shortform36 or sf thirtysix or sfthirtysix or sfthirty six or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab,kf.	31921
	13	(sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or shortform6 or short form6).ti,ab,kf.	2773
	14	(sf8 or sf 8 or sf eight or sfeight or shortform 8 or shortform 8 or shortform8 or short form8 or shortform eight or short form eight).ti,ab,kf.	649

	15	(sf12 or sf 12 or short form 12 or shortform 12 or short form12 or shortform12 or sf twelve or sftwelve or shortform twelve or short form twelve).ti,ab,kf.	8111
	16	(sf16 or sf 16 or short form 16 or shortform 16 or short form16 or shortform16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).ti,ab,kf.	42
	17	(sf20 or sf 20 or short form 20 or shortform 20 or short form20 or shortform20 or sf twenty or sftwenty or shortform twenty or short form twenty).ti,ab,kf.	468
	18	(hql or hqol or h qol or hrqol or hr qol).ti,ab,kf.	25982
	19	(hye or hyes).ti,ab,kf.	78
	20	(health* adj2 year* adj2 equivalent*).ti,ab,kf.	48
	21	(pqol or qls).ti,ab,kf.	484
	22	(quality of wellbeing or quality of well being or index of wellbeing or index of well being or qwb).ti,ab,kf.	751
	23	nottingham health profile*.ti,ab,kf.	1260
	24	sickness impact profile.ti,ab,kf.	1101
	25	exp health status indicators/	346253
	26	(health adj3 (utilit* or status)).ti,ab,kf.	97496
	27	(utilit* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or weight)).ti,ab,kf.	16960
	28	(preference* adj3 (valu* or measur* or health or life or estimat* or elicit*	15350

		or disease or score* or instrument or instruments)).ti,ab,kf.	
	29	disutilit*.ti,ab,kf.	677
	30	rosser.ti,ab,kf.	112
	31	willingness to pay.ti,ab,kf.	9450
	32	standard gamble*.ti,ab,kf.	922
	33	(time trade off or time tradeoff).ti,ab,kf.	1719
	34	tto.ti,ab,kf.	1491
	35	(hui or hui1 or hui2 or hui3).ti,ab,kf.	2095
	36	(eq or euroqol or euro qol or eq5d or eq 5d or euroqual or euro qual).ti,ab,kf.	24584
	37	duke health profile.ti,ab,kf.	94
	38	functional status questionnaire.ti,ab,kf.	134
	39	dartmouth coop functional health assessment*.ti,ab,kf.	14
	40	or/1-39	787924
Multiple system atrophy	41	*Multiple System Atrophy/ or system atrophy.ti.	3333
Total	42	40 and 41	189

Table 37: EMBASE (1974 to 2024 Week 19) systematic searches on Ovid platform for economic models and multiple system atrophy performed on the 14th of May 2024.

Topic	Number	Search	Hits
Economic evaluations and Models (incl. cost analyses)	1	Economics/	245267
	2	Cost/	64432
	3	exp Health Economics/	1070886
	4	Budget/	34595
	5	budget*.ti,ab,kf.	50039
	6	(economic* or cost or costs or costly or costing or price or prices or pricing	362619

		or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ti,kf.	
	7	(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ab. /freq=2	564330
	8	(cost* adj2 (effective* or utilit* or benefit* or minimi* or analy* or outcome or outcomes)).ab,kf.	309050
	9	(value adj2 (money or monetary)).ti,ab,kf.	4298
	10	Statistical Model/	177261
	11	exp economic model/	4176
	12	economic model*.ab,kf.	6641
	13	Probability/	154211
	14	markov.ti,ab,kf.	40337
	15	monte carlo method/	53484
	16	monte carlo.ti,ab,kf.	66644
	17	Decision Theory/	1870
	18	Decision Tree/	24333
	19	(decision* adj2 (tree* or analy* or model*)).ti,ab,kf.	57182
	20	or/1-19	2099755
Multiple system atrophy	21	multiple system atrophy.ti,ab. or *Shy Drager syndrome/	8677
Total	22	20 and 21	128

Table 38: EMBASE (1974 to 2024 Week 19) systematic searches on Ovid platform for health utilities and multiple system atrophy performed on the 14th of May 2024.

Topic	Number	Search	Hits
Health Utilities / Quality of life	1	socioeconomics/	166841
	2	exp Quality of Life/	693638
	3	quality of life.ti,kf.	190790
	4	((instrument or instruments) adj3 quality of life).ab.	5609
	5	Quality-Adjusted Life Year/	37299
	6	quality adjusted life.ti,ab,kf.	28233
	7	(qaly* or qald* or qale* or qtime* or life year or life years).ti,ab,kf.	47641
	8	disability-adjusted life year/	5446
	9	disability adjusted life.ti,ab,kf.	7316
	10	healthy life expectancy/	284
	11	(daly* or disability free life expectanc* or haly* or health* life expectanc*).ti,ab,kf.	8985
	12	exp Short form 36/	52185
	13	(sf36 or sf 36 or short form 36 or shortform 36 or short form36 or shortform36 or sf thirtysix or sfthirtysix or sfthirty six or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab,kf.	51851
	14	(sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or shortform6 or short form6).ti,ab,kf.	3104
	15	(sf8 or sf 8 or sf eight or sfeight or shortform 8 or shortform 8 or shortform8 or short form8 or	1066

		shortform eight or short form eight).ti,ab,kf.	
	16	(sf12 or sf 12 or short form 12 or shortform 12 or short form12 or shortform12 or sf twelve or sftwelve or shortform twelve or short form twelve).ti,ab,kf.	12812
	17	(sf16 or sf 16 or short form 16 or shortform 16 or short form16 or shortform16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).ti,ab,kf.	72
	18	(sf20 or sf 20 or short form 20 or shortform 20 or short form20 or shortform20 or sf twenty or sftwenty or shortform twenty or short form twenty).ti,ab,kf.	544
	19	(hql or hqol or h qol or hrqol or hr qol).ti,ab,kf.	41715
	20	(hye or hyes).ti,ab,kf.	195
	21	(health* adj2 year* adj2 equivalent*).ti,ab,kf.	56
	22	(pqol or qls).ti,ab,kf.	781
	23	(quality of wellbeing or quality of well being or index of wellbeing or index of well being or qwb).ti,ab,kf.	927
	24	nottingham health profile*.ti,ab,kf.	1702
	25	nottingham health profile/	683
	26	sickness impact profile.ti,ab,kf.	1301
	27	sickness impact profile/	2407
	28	health status indicator/	3579
	29	(health adj3 (utilit* or status)).ti,ab,kf.	127707

	30	(utilit* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or weight)).ti,ab,kf.	26998
	31	(preference* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or instrument or instruments)).ti,ab,kf.	20290
	32	disutilit*.ti,ab,kf.	1349
	33	rosser.ti,ab,kf.	146
	34	Willingness To Pay/	4262
	35	willingness to pay.ti,ab,kf.	14046
	36	Standard Gamble/	93
	37	standard gamble*.ti,ab,kf.	1232
	38	time trade-off method/	562
	39	(time trade off or time tradeoff).ti,ab,kf.	2482
	40	tto.ti,ab,kf.	2353
	41	(hui or hui1 or hui2 or hui3).ti,ab,kf.	3319
	42	(eq or euroqol or euro qol or eq5d or eq 5d or euroqual or euro qual).ti,ab,kf.	40880
	43	duke health profile.ti,ab,kf.	121
	44	functional status questionnaire.ti,ab,kf.	181
	45	dartmouth coop functional health assessment*.ti,ab,kf.	14
	46	or/1-45	1059231
Multiple system atrophy	47	multiple system atrophy.ti. or *Shy Drager syndrome/	4760
Total	48	46 and 47	159

Table 39: MEDLINE (1946 to May 12, 2024) targeted searches on Ovid platform for economic evaluations and rare-disease analogues performed on the 13th of May 2024.

Topic	Number	Search	Hits

Economic evaluations – narrow search	1	*economics/	10814
	2	exp *"costs and cost analysis"/	80909
	3	(economic adj2 model*).mp.	15528
	4	(cost minimi* or cost-utilit* or health utilit* or economic evaluation* or economic review* or cost outcome* or cost analys?s or economic analys?s or budget* impact analys?s).ti,ab,kf.	45382
	5	(cost-effective* or pharmacoeconomic* or pharmaco-economic* or cost-benefit or costs).ti,kf.	92372
	6	(life year or life years or qaly* or cost-benefit analys?s or cost-effectiveness analys?s).ab,kf.	42502
	7	or/1-6	205908
Amyotrophic lateral sclerosis (ALS)	8	Amyotrophic Lateral Sclerosis.ti. or *Amyotrophic Lateral Sclerosis/	23401
Multiple sclerosis (MS)	9	multiple sclerosis.ti. or *Multiple Sclerosis/	71000
Loss of ambulation through, e.g., late-stage Parkinson disease or spinal cord injury	10	(Parkinson Disease/ or spinal cord injuries/) AND (gait disorders, neurological/ or ambulation/ or walking/ or mobility limitation/)	2179
Combining diseases	11	or/8-10	96365
Combining searches	12	7 and 11	871
Total (last 5 years)	13	12 and 2019:2024.(sa_year).	297
Total (journal articles)	14	Limit 13 to journal article	287

Table 40: EMBASE (1974 to 2024 Week 19) targeted searches on Ovid platform for economic evaluations and rare-disease analogues performed on the 14th of May 2024.

Topic	Number	Search	Hits

Economic evaluations – narrow search	1	*economics/	27819
	2	exp *"costs and cost analysis"/	93183
	3	(economic adj2 model*).mp.	10531
	4	(cost minimi* or cost-utilit* or health utilit* or economic evaluation* or economic review* or cost outcome* or cost analys?s or economic analys?s or budget* impact analys?s).ti,ab,kf.	69210
	5	(cost-effective* or pharmacoeconomic* or pharmaco-economic* or cost-benefit or costs).ti,kf.	135380
	6	(life year or life years or qaly* or cost-benefit analys?s or cost-effectiveness analys?s).ab,kf.	65039
	7	or/1-6	292448
Amyotrophic lateral sclerosis (ALS)	8	Amyotrophic Lateral Sclerosis.ti. or *amyotrophic lateral sclerosis/	28817
Multiple sclerosis (MS)	9	*multiple sclerosis/ or multiple sclerosis.ti.	110469
Loss of ambulation through, e.g., late-stage Parkinson disease or spinal cord injury	10	(Parkinson Disease/ or spinal cord injuries/) AND (gait disorders, neurological/ or ambulation/ or walking/ or mobility limitation/)	4682
Combining diseases	11	or/8-10	143574
Combining searches	12	7 and 11	1455
Total (last 5 years)	13	12 and 2019:2024.(sa_year).	473
Total (journal articles)	14	limit 13 to article	260

Table 41: Searches performed in the NHS EED database on the 7th of May 2024 for multiple system atrophy. Disease analogues only for the period 2019 – 2024 (last 5 years).

Topic	Number	Search	Hits

Multiple system atrophy	1	economic evaluation or cost-effectiveness and multiple system atrophy	0
Advanced Parkinson	2	economic evaluation or cost-effectiveness and advanced parkinson	0
Amyotrophic lateral sclerosis	3	economic evaluation or cost-effectiveness and amyotrophic lateral sclerosis	0
Multiple sclerosis	4	economic evaluation or cost-effectiveness and multiple sclerosis	0

Table 42: Cost-effectiveness analyses identified from literature and hand searches (only disease analogues)

Author: Title	**Disease**	**Model type**
Acosta et al 2021: Modelling the cost-effectiveness of prolonged release fampridine for the treatment of walking in impairment in patients with multiple sclerosis in Sweden	Multiple Sclerosis	Markov Model - with responder and non-responder health states
Auguste et al 2020: Ocrelizumab for Treating Patients with Primary	Multiple Sclerosis	Markov Model - health states based on EDSS (Expanded Disability Status Scale)

Progressive Multiple Sclerosis: An Evidence Review Group Perspective of a NICE Single Technology Appraisal		
Ayati et al 2021: Cost-effectiveness of ocrelizumab for treatment of Iranian patients with relapsing multiple sclerosis	Multiple Sclerosis	Markov Model - EDSS with on- and off-treatment cycles
Ayati et al 2021 (2nd paper): Cladribine tablets are a cost-effective strategy in high-disease activity relapsing multiple sclerosis patients in Iran	Multiple Sclerosis	Markov Model - EDSS with on- and off-treatment cycles
Baharnoori et al 2022: Cost-Effectiveness Analysis of Ofatumumab for the Treatment of	Multiple Sclerosis	Markov Model - EDSS states

Relapsing-Remitting Multiple Sclerosis in Canada		
Bolanos-Diaz et al 2019: Cost-effectiveness of interferon therapy for multiple sclerosis in Peru: impact of treatment adherence	Multiple Sclerosis	Markov Model - with types of multiple sclerosis (relapsing-remitting or progressive) and clinical isolated syndrome.
Chaudhuri et al 2020: The Cost Effectiveness of Levodopa-Carbidopa Intestinal Gel in the Treatment of Advanced Parkinson's Disease in England	Advanced Parkinson's Disease	Markov model - Hoehn and Yahr Disease states and off-time (due to motor fluctations)
Cortesi et al 2022: Cost-effectiveness and budget impact analysis of siponimod in the treatment of secondary progressive	Multiple Sclerosis	Markov Model - EDSS states

multiple sclerosis in Italy		
Daroudi et al 2023: Cost-utility analysis of multiple sclerosis rehabilitation in Iran	Multiple Sclerosis	Markov Model - EDSS states
Dimitrova et al 2023: COST-EFFECTIVENESS OF DISEASE-MODIFYING TREATMENTS FOR MULTIPLE SCLEROSIS IN BULGARIA BASED ON EVIDENCE FOR MULTIPLE SCLEROSIS IN BULGARIA BASED ON EVIDENCE	Multiple Sclerosis	Model type was unclear - but the endpoint was cost per relapse avoided.
Furneri et al 2019: Cost-effectiveness analysis of escalating to natalizumab or switching among immunomodulators in relapsing-	Multiple Sclerosis	Markov model - EDSS states with progression (relapse) states

remitting multiple sclerosis in Italy		
Gajdos et al 2021: Cost–Utility Analysis of Home Mechanical Ventilation in Patients with Amyotrophic Lateral Sclerosis	Amyotrophic Lateral Sclerosis	Two Markov Models stratified by treatments - Home Care (Ventilation) and Healthcare Facility (Ventilation). Each of the models shared the states invasive and non-invasive ventilation, but the home care model had hospitalisation as an additional health state.
Ginestal et al 2022: Cost–effectiveness of cladribine tablets and dimethyl fumarate in the treatment of relapsing remitting multiple sclerosis in Spain	Multiple Sclerosis	Markov Model with EDSS five states and death. Developed by CADTH.
Giovannoni et al 2019: Glatiramer acetate as a clinically and cost-effective treatment of relapsing multiple sclerosis over treatment of	Multiple Sclerosis	Markov Model with EDSS states. Developed for a NICE submission.

relapsing multiple sclerosis over Final results from the UK Risk Sharing Scheme		
Hua et al 2019: The Impact of Price Reductions After Loss of Exclusivity in a Cost-Effectiveness Analysis: Fingolimod Versus Interferon Beta-1a for the Treatment of Relapsing-Remitting Multiple Sclerosis	Multiple Sclerosis	Markov Model with EDSS states.
Hunter et al. 2021: Cost-Effectiveness of Repository Corticotropin Injection for the Treatment of Acute Exacerbations in Multiple Sclerosis	Multiple Sclerosis	Markov model with Exacerbation / Relapse and Remission / Response states.

Huygens and Versteegh 2021: Modeling the Cost-Utility of Treatment Sequences for Multiple Sclerosis	Multiple Sclerosis	Microsimulatio model with EDSS progression or relapse
Kalabina et al 2018: Cost-utility analysis of levodopa carbidopa intestinal gel (Duodopa) in the treatment of advanced Parkinson's disease in patients in Scotland and Wales	Advanced Parkinson 's Disease	Markov model with Hoehn and Yahr states and off-states (uncontrolled motor symptoms that are non-responsive to treatment)
Kremer et al 2020: Exploring the Cost Effectiveness of Shared Decision Making for Choosing between Disease-Modifying Drugs for Relapsing-Remitting Multiple	Multiple Sclerosis	Markov model with EDSS states further stratified by relapsing-remitting multiple sclerosis and secondary progressive multiple sclerosis.

Sclerosis in the Netherlands: A State Transition Model		
Lasalvia et al 2020: Cost-Effectiveness of Natalizumab Compared With Fingolimod for Relapsing-Remitting Multiple Sclerosis Treatment in Colombia	Multiple Sclerosis	Markov model with EDSS states and death.
Lazzaro et al 2021: Cost-utility analysis of teriflunomide in naïve vs. previously treated patients with relapsing–remitting multiple sclerosis in Italy	Multiple Sclerosis	Markov model with four states: controlled relapsing-remitting (MS), relapsing remitting (MS), relapsing remitting (MS) remission, all-cause mortality
Mantovani et al 2020: Cost-Effectiveness Analysis of Cannabinoid Oromucosal Spray Use for the Management of	Multiple Sclerosis	No model used, but rather modelled accumulated treatment effects and costs from registry data directly for Sativex vs. No treatment.

Spasticity in Subjects with Multiple Sclerosis		
Martins et al 2022: Cost-effectiveness Analysis of Ocrelizumab for the Treatment of Relapsing and Primary Progressive Multiple Sclerosis in Portugal	Multiple Sclerosis	Markov model with EDSS states stratified by on or off treatment, and further by relapsing remitting MS and secondary progressive MS.
Michel et al 2022: Cost-utility of oral methylprednisolone in the treatment of multiple sclerosis relapses: Results from the COPOUSEP trial	Multiple Sclerosis	No model was used, but a randomised trial modelled directly.
Michels et al 2019: Cost Effectiveness of Cladribine Tablets for the Treatment of Relapsing-Remitting Multiple	Multiple Sclerosis	Markov model with EDSS states, relapse / remitting MS and secondary progressive MS and Death

Sclerosis in The Netherlands		
Montgomery et al 2022: Stick or twist? Cost-effectiveness of siponimod compared with continuing existing disease modifying therapies in the treatment of active secondary progressive multiple sclerosis in the UK	Multiple Sclerosis	Markov model with EDSS states, on treatment, off treatment and death
Nakhaipour et al 2020: Cost-effectiveness of fingolimod versus interferon-β1a for the treatment of pediatric-onset multiple sclerosis in Canada	Multiple Sclerosis	Markov model with EDSS states, remitting MS with disease modifying therapy, remitting MS with best supportive care and secondary progressive MS with best supportive care, and death
Navarro et al 2023: Cost-Utility Analysis Comparing Ocrelizumab	Multiple Sclerosis	Markov model with EDSS states, relapse / remitting MS and secondary progressive MS and Death

Versus Rituximab in the Treatment of Relapsing-Remitting Multiple Sclerosis: The Colombian Perspective		
Oppe et al 2021: Cost-effectiveness of adding Sativex® spray to spasticity care in Belgium: using bootstrapping instead of Monte Carlo simulation for probabilistic sensitivity analyses	Multiple Sclerosis	Markov model for resistant MS with spasciticity with three Markov states for mild, moderate and severe spasticity and death
Palmer et al 2023: Exploring the cost-effectiveness of EBV vaccination to prevent mulitple sclerosis in an Australian setting	Multiple Sclerosis	Markov model with no, mild, moderate and severe MS.

Pinheiro et al 2020: Cost-effectiveness of cladribine tablets versus fingolimod in patients with highly active relapsing multiple sclerosis in Portugal	Multiple Sclerosis	Markov model with EDSS states, remitting MS with disease modifying therapy, remitting MS without disease modifying therapy and secondary progressive MS without disease modifying therapy, and all-cause death
Pinto et al 2023: Cost-effectiveness analysis of overground robotic training versus conventional locomotor training in people with spinal cord injury	Spinal cord injury	An Excel-based simulation was used, but the model structure was unclear, but appears to model a practice-based cohort comparing training strategies.
Poveda et al 2020: Cost-effectiveness of Cladribine Tablets and fingolimod in the treatment of relapsing multiple sclerosis with	Multiple Sclerosis	Markov model with EDSS states stratified by treatment / not treatment or from remitting MS to secondary progressive MS.

high disease activity in Spain		
Renfrew et al 2019: The clinical- and cost-effectiveness of functional electrical stimulation and ankle-foot orthoses for foot drop in Multiple Sclerosis: a multicentre randomized trial	Multiple Sclerosis	Unclear if which model structure is used, although they do mention an ANOVA model (regression) for computing outcomes.
Rezaee et al 2019: Fingolimod versus natalizumab in patients with relapsing remitting multiple sclerosis: a costeffectivenes s and cost-utility study in Iran	Multiple Sclerosis	Markov model with EDSS states
Rosato et al 2021: Living with severe multiple sclerosis: Cost-effectiveness of a palliative care	Multiple Sclerosis	Unclear, but text indicates no model. Instead, a comparison of accummulated differences across trial arms

intervention and cost of illness study		
Schur et al 2021: Cost Effectiveness and Budget Impact of Siponimod Compared to Interferon Beta-1a in the Treatment of Adult Patients with Secondary Progressive Multiple Sclerosis with Active Disease in Switzerland	Multiple Sclerosis	Markov model with EDSS states - on and off treatment.
Spelman et al 2022: Comparative Effectiveness and Cost-Effectiveness of Natalizumab and Fingolimod in Patients with Inadequate Response to Disease-Modifying Therapies in Relapsing-Remit	Multiple Sclerosis	Markov model with EDSS states, RRMS and SPMS, and death.

ting Multiple Sclerosis in the United Kingdom		
Spelman et al 2024: Comparative Effectiveness and Cost-Effectiveness of Natalizumab and Fingolimod in Patients with Inadequate Response to Disease-Modifying Therapies in Relapsing-Remitting Multiple Sclerosis in the United Kingdom	Multiple Sclerosis	Markov model with EDSS states, RRMS and SPMS, and death.
Stanisic et al 2019: The cost-effectiveness of alemtuzumab in the management of relapse-remitting multiple sclerosis in Italy	Multiple Sclerosis	Markov model with EDSS states across RRMS and SPMS, and death. Model was developed by ScHARR
Taheri et al 2019: Cost-effectiveness of alemtuzumab	Multiple Sclerosis	Markov models with EDSS states across RRMS (on - off treatment), SPMS, and death

and natalizumab for relapsing-remitting multiple sclerosis treatment in Iran: decision analysis based on an indirect comparison		
Thakore et al 2020: A Cost-Effectiveness Framework for Amyotrophic Lateral Sclerosis, Applied to Riluzole	Amyotrophic Lateral Sclerosis	Appears to be a Markov model using FT9 models with five states (ALS functional rating scale) and death. FT9: Propose an empirical amyotrophic lateral sclerosis (ALS) staging approach called Fine'til 9 (FT9) based on how many of the patient's ALS functional rating scale (ALSFRS-R) subscores are 9 or less (of normal 12). See table 1 in https://www.tandfonline.com/doi/epdf/10.1080/21678421.2018.1484925?needAccess=true for overview of state explanations.
Versteegh et al 2022: Effectiveness and Cost-Effectiveness of 360 Disease-Modifying Treatment Escalation Sequences in Multiple Sclerosis	Multiple Sclerosis	Microsimulatio model with EDSS progression or relapse
Visser et al 2022: The Potential Cost-Effectivene	Multiple Sclerosis	Markov model from ICER report with EDSS states and DMT (disease-modifying treament) and BSC (best supportive care).

ss of a Cell-Based Bioelectronic Implantable Device Delivering Interferon-β1a Therapy Versus Injectable Interferon-β1a Treatment in Relapsing–Remitting Multiple Sclerosis		
Walter et al 2019: Cost-utility analysis of alemtuzumab in comparison with interferon beta, fingolimod, and natalizumab treatment for relapsing-remitting multiple sclerosis in Austria	Multiple Sclerosis	Markov model with 10 EDSS, 4 treatments, relapse, discontinuation, switching or no treatment, adverse evets and death.
Xu et al 2019: Cost-effectiveness of Teriflunomide Compared to Interferon	Multiple Sclerosis	Markov model with EDSS states, relapse and death.

Beta-1b for Relapsing Multiple Sclerosis Patients in China		
Zhao et al 2022: Cost–effectiveness analysis of prolonged-release fampridine to treat walking disability of multiple sclerosis in China	Multiple Sclerosis	Markov model with states: responder, non-responder, best supportive care and death.
Pahwa et al 2023: Cost-Effectiveness of Carbidopa-Levodopa Enteral Suspension for Advanced Parkinson's Disease in the United States	Advanced Parkinson 's Disease	Markov model with Hoehn and Yahr states and off-states (uncontrolled motor symptoms that are non-responsive to treatment)

II.II Appendix: Figure

See Code availability to assess the online repositories that contain additional table and figures.

III Appendix: (Conceptual modelling of functional milestones)

III.I Modelling report

1 Data Analysis

1.1 Data Preparation

2 Partitioned Survival Model

3 Markov model

4 Visuals for the report

Code

Partitioned Survival Model and Markov Model comparison for Multiple System Atrophy

Tobias Sydendal Grand

2026-01-28

Please see code in collapsed sections above output by pressing "Show".

1 Data Analysis

1.1 Data Preparation

In below code chunk we set the working directory and load packages required for the analysis.

Show

1.1.1 Code example (aid-required walking) to manually digitise Kaplan-Meier curves

A separate R file DigitiseKaplanMeier.R in folder named data contains code to generate digitised Kaplan-Meier curves from:

```
Watanabe et al 2022 Progression and prognosis in multiple system atrophy (DOI: 10.1093/bra
in/awf117)
```

The file is not loaded into this report - because digitisation requires manual clicking, which would prevent the Markdown report from running. However, we will show a chunk of code below on how the csv files are loaded into the R environment.Note that the images from the original source paper used for digitisation are not uploaded in the R projects to avoid copyright infringements.

Show

```
## #Load images from Watanabe et al 2002
##   ## Aid-required walking
##
## img_TimeToAidWALK <- png::readPNG("Data/TimeToAidWalk.png")
##
## ##Follow steps in IPD KM packages
##
## str(img_TimeToAidWALK)
## plot.new()
##
## ## Position x1, x2, y1, y2
## rasterImage(img_TimeToAidWALK, 0, 0, 1, 1)
##
## ## Define scale of x and y axis
##
## df1 <- getpoints(img_TimeToAidWALK, 0, 16, 0, 1)
##
## ## Rename variables
##
## df1 <- df1 %>% rename(time = x, probability = y)
## df1 %>% plot(type = "s")
##
## ## Rename dataframe
##
## TimeToAidWalk <- df1
##
## ## Save as csv file
##
## write.csv(TimeToAidWalk, "Data/TimeToAidWalk.csv", row.names = FALSE)
```

1.1.2 Load CSV files containing manually digitised Kaplan-Meier curves and prepare datasets for analysis

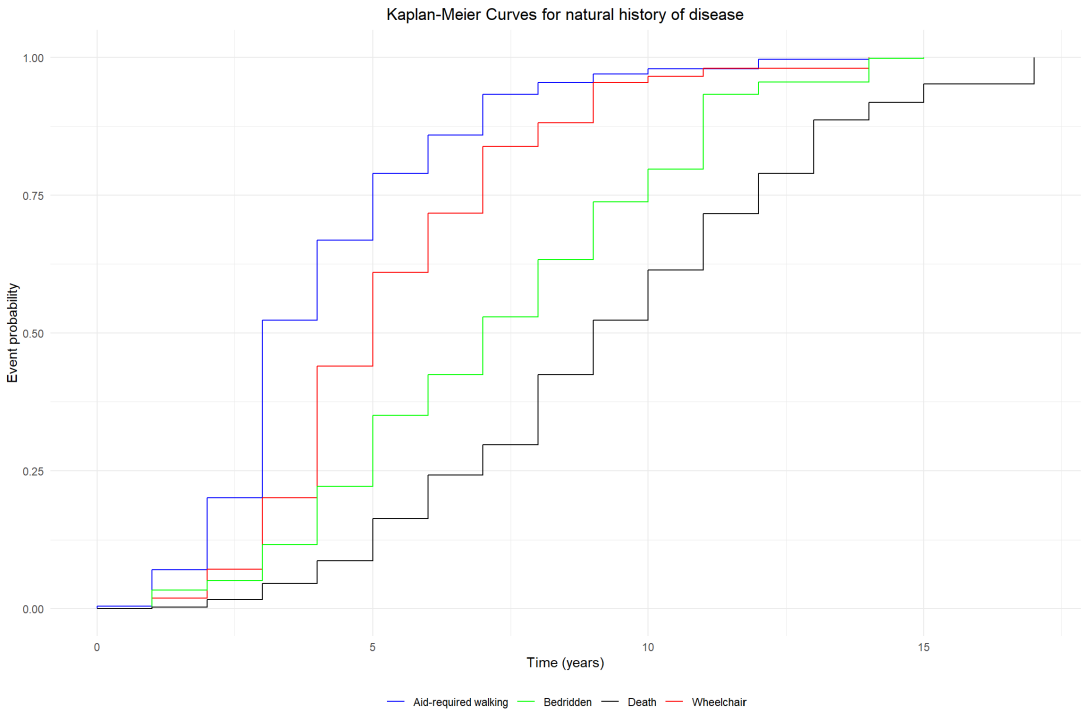
This code chunk loads the csv files that contain the time to event data. For example, the probability of an event at each recorded time point, which was extracted from the Kaplan-Meier curves. Note that we create a dataset for a hypothetical treatment arm. The hypothetical treatment arm is assumed to have a disease modifying effect, which slows the progression of the disease. The time to event is increased by 50% in the hypothetical treatment arm.

Show

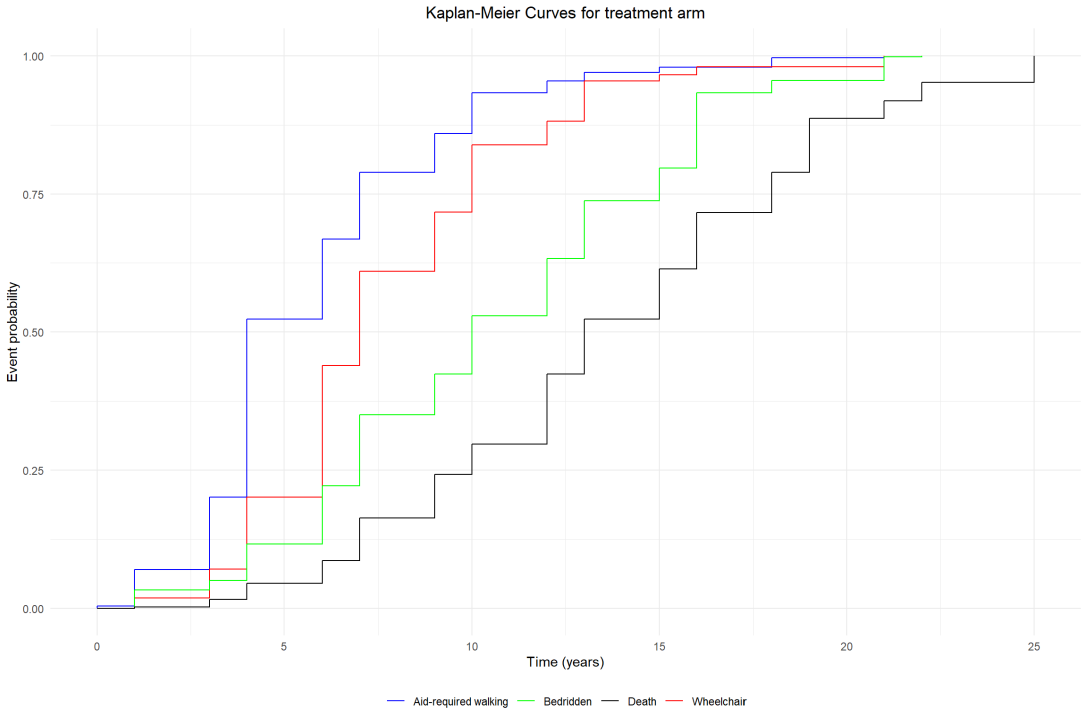
1.1.3 Illustrate the digitised Kaplan-Meier curves.

It is advantageous to compare the digitised Kaplan-Meier curves for similarity with the original Kaplan-Meier curves from the source paper, if re-using this code, to check for clicking errors and similarity. Naturally, only for the natural history of disease (ORIGINAL) data.

Show


Kaplan-Meier Curves for natural history of disease
Event probability
1.00
0.75
0.50
0.25
0.00
0
5
10
15
Time (years)
Aid-required walking
Bedridden
Death
Wheelchair

Show


Kaplan-Meier Curves for treatment arm
Event probability
1.00
0.75
0.50
0.25
0.00
0
5
10
15
20
25
Time (years)
Aid-required walking
Bedridden
Death
Wheelchair

1.1.4 Create patient level data (IPD)

In this section, we create datasets for individual patient data (IPD) based on the digitised Kaplan-Meier curves. We add information from the paper e.g., number of patients, for the natural history of disease and hypothetical treatment arm, which can later be used to fit parametric distributions.

Show

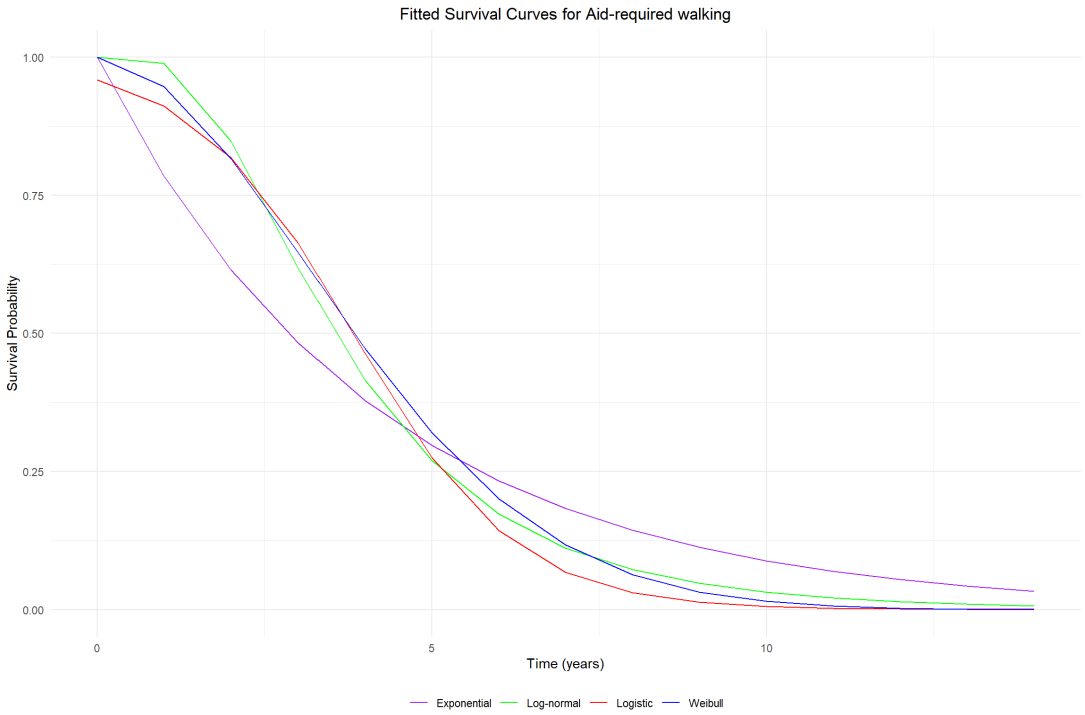
1.1.5 Fitting distributions

We fit distributions, first for the natural history of disease and for the hypothetical treatment arm.

1.1.6 Natural history of disease

We fit parametric distributions for all events for natural history of disease. It is visually clear that Weibull and logistic best fit the data. When applying Akaikes Information Criterion and Bayes Information Criterion, the Weibull distribution is the best fit with lowest AIC and BIC for all events.

Show



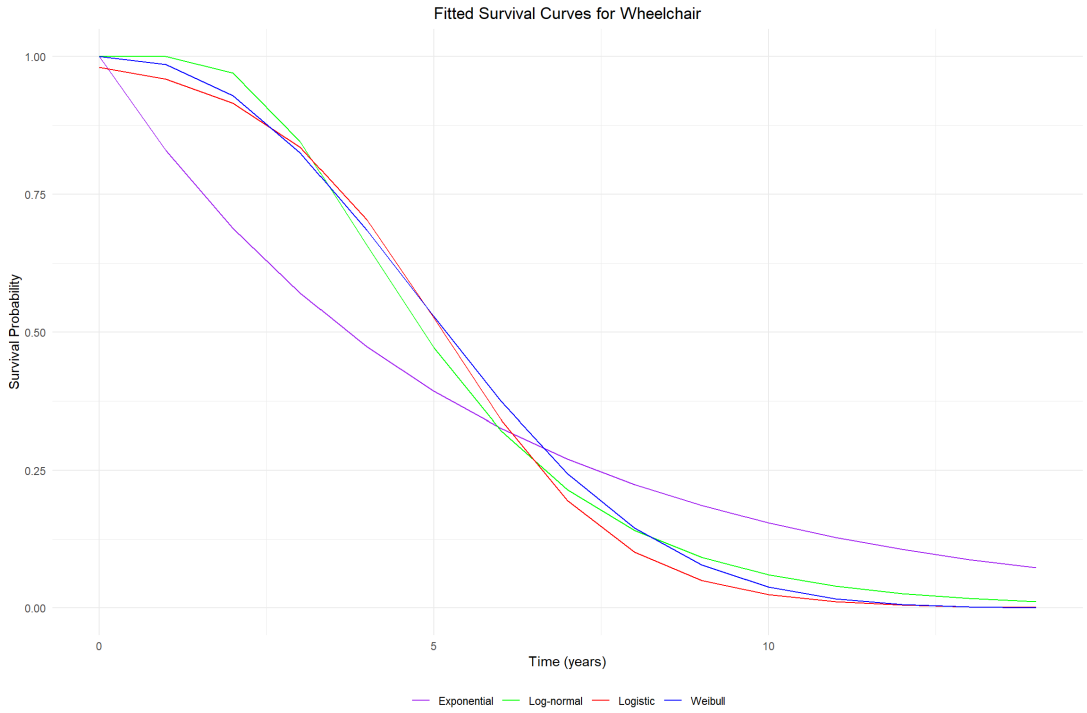
Show

Model Comparison using AIC and BIC for Aid-required walking

Distribution	AIC	BIC

Distribution	AIC	BIC
Weibull	992.3368	999.2129
Logistic	1021.4637	1028.3398
Log-normal	970.5775	977.4537
Exponential	1113.4866	1116.9247

Show

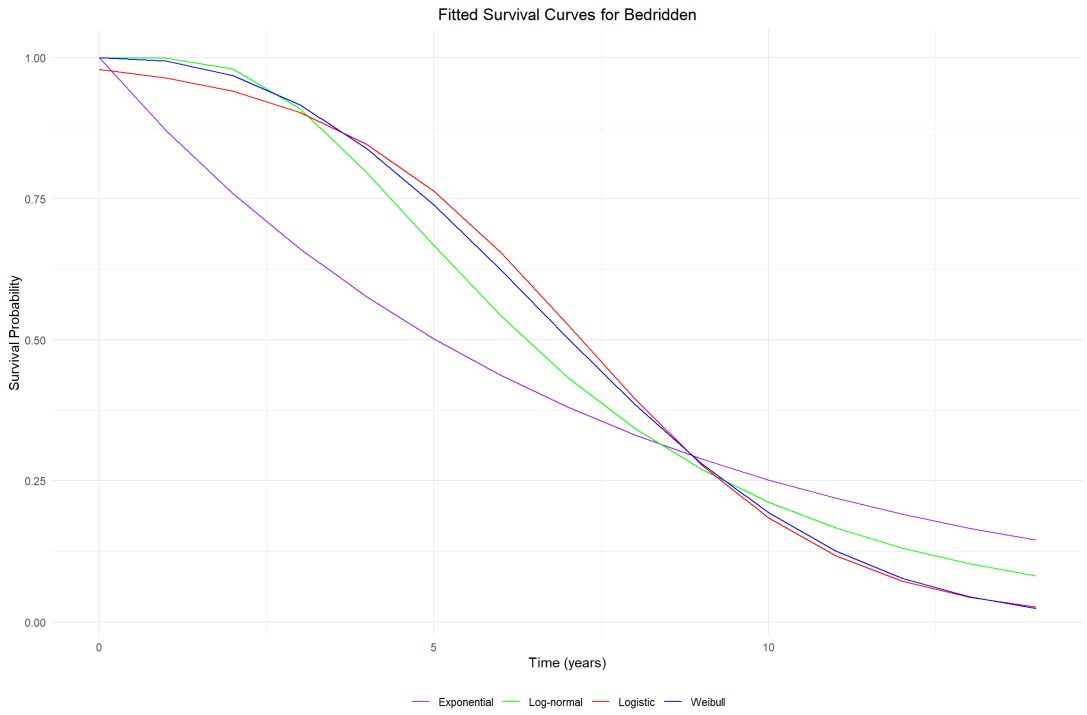


Show

Model Comparison using AIC and BIC for Wheelchair

Distribution	AIC	BIC
Weibull	1037.703	1044.580
Logistic	1049.943	1056.819
Log-normal	1032.160	1039.036
Exponential	1233.651	1237.089

Show

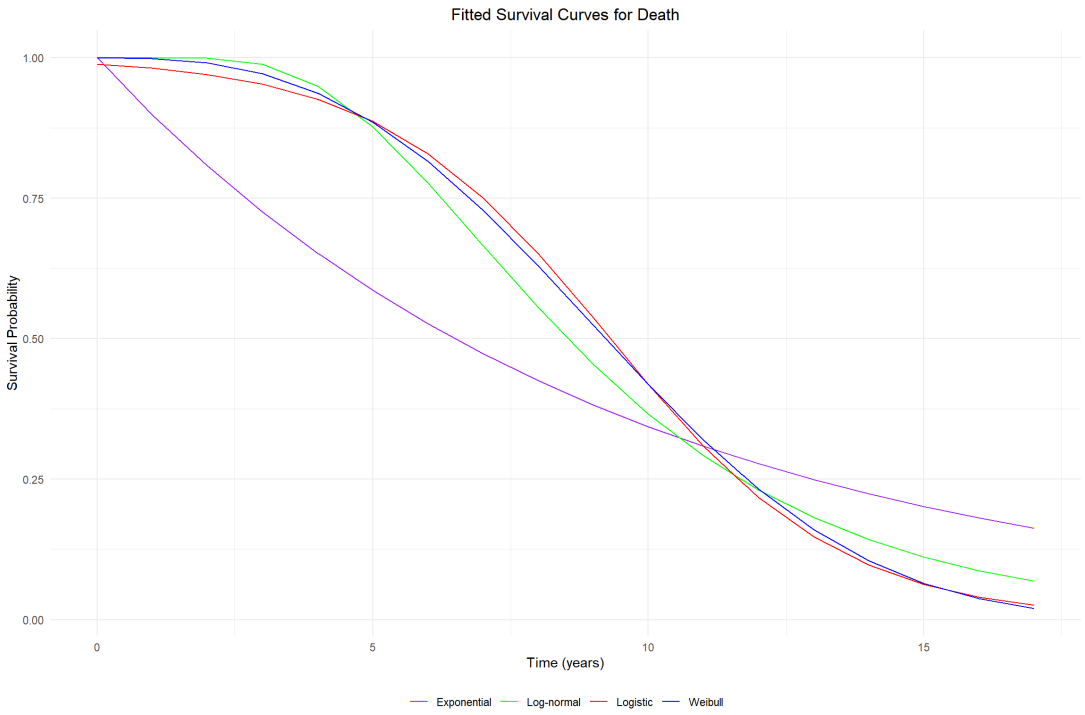


Show

Model Comparison using AIC and BIC for Bedridden

Distribution	AIC	BIC
Weibull	1183.094	1189.971
Logistic	1202.295	1209.171
Log-normal	1244.583	1251.460
Exponential	1372.847	1376.285

Show



Show

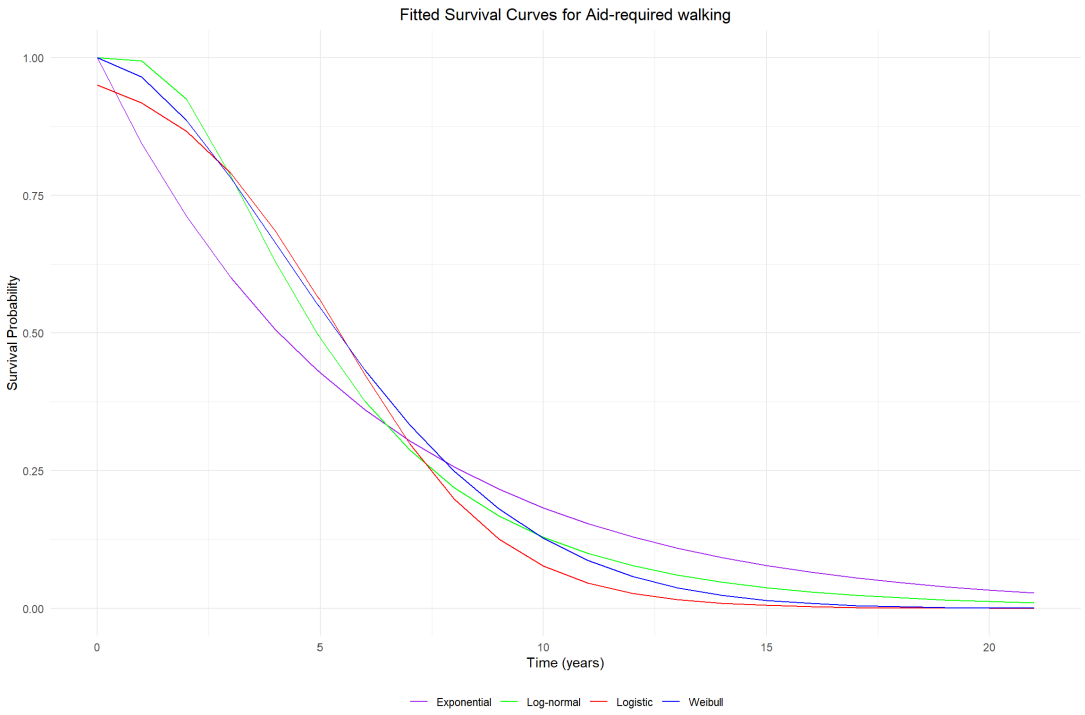
Model Comparison using AIC and BIC for Death

Distribution	AIC	BIC
Weibull	1240.810	1247.686
Logistic	1255.737	1262.613
Log-normal	1288.052	1294.928
Exponential	1490.166	1493.604

1.1.7 Treatment arm

Not surprisingly. The same distributions are the best fit for the treatment arm. Weibull and logistic distributions are the best fit for all events. The Weibull distribution is the best fit with lowest AIC and BIC for all events.

Show

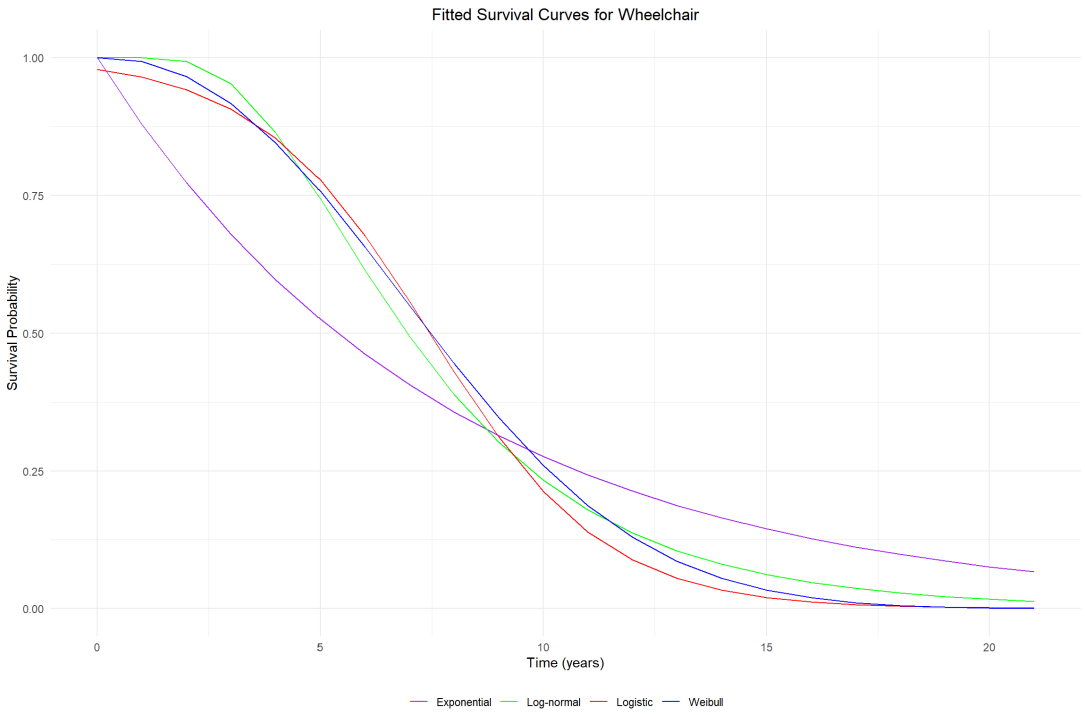


Show

Model Comparison using AIC and BIC for Aid-required walking

Distribution	AIC	BIC
Weibull	1178.481	1185.357
Logistic	1213.349	1220.225
Log-normal	1175.555	1182.431
Exponential	1277.120	1280.558

Show

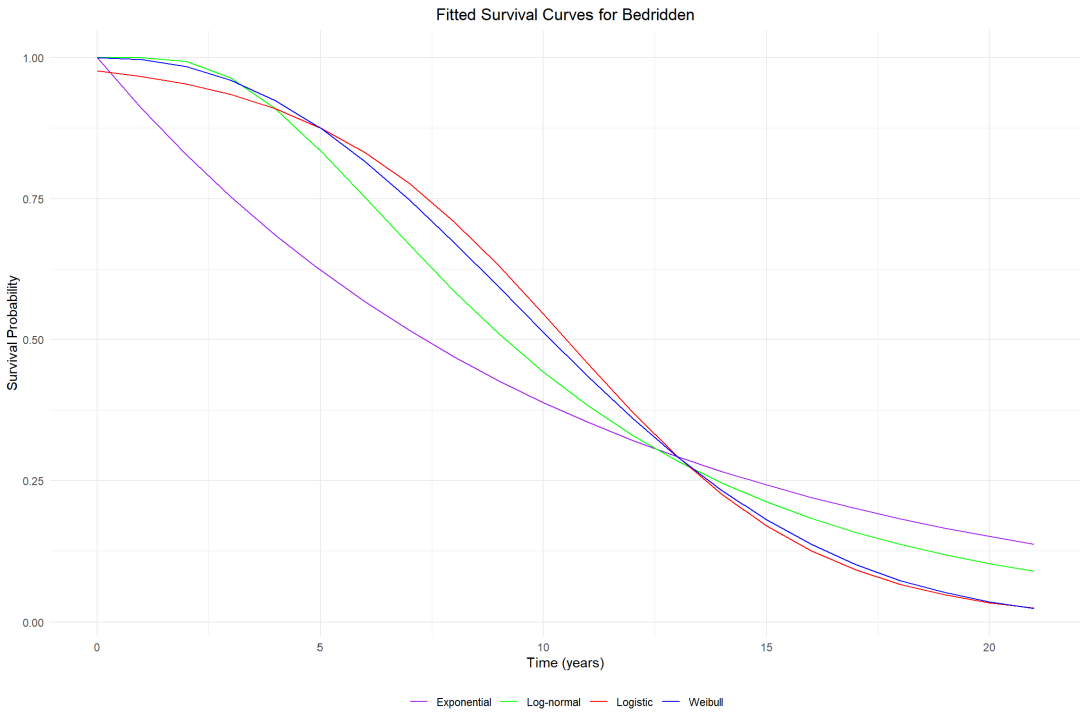


Show

Model Comparison using AIC and BIC for Wheelchair

Distribution	AIC	BIC
Weibull	1221.799	1228.675
Logistic	1232.810	1239.686
Log-normal	1229.752	1236.628
Exponential	1404.841	1408.279

Show

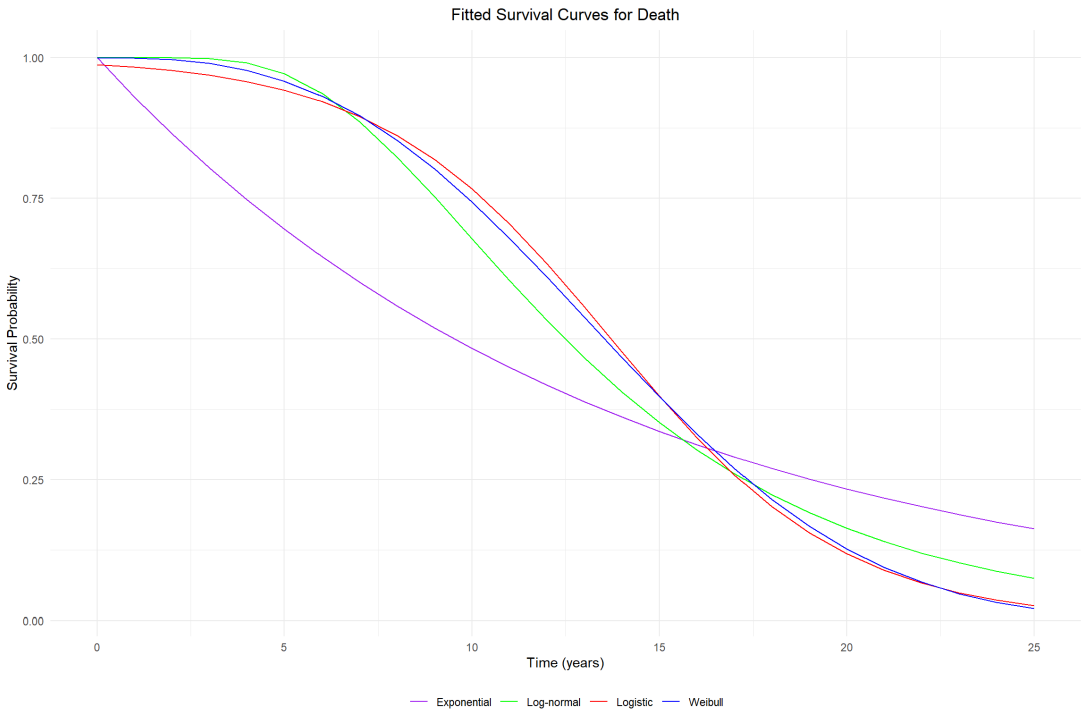


Show

Model Comparison using AIC and BIC for Bedridden

Distribution	AIC	BIC
Weibull	1375.461	1382.338
Logistic	1390.157	1397.034
Log-normal	1454.249	1461.126
Exponential	1547.048	1550.487

Show



Show

Model Comparison using AIC and BIC for Death

Distribution	AIC	BIC
Weibull	1424.894	1431.770
Logistic	1438.765	1445.641
Log-normal	1480.909	1487.785
Exponential	1667.605	1671.043

2 Partitioned Survival Model

We use the hesim package developed for modelling in R

```
Incerti D, Jansen J (2021). “hesim: Health Economic Simulation Modeling and Decision Analy
sis.” 2102.09437, https://arxiv.org/abs/2102.09437.
```

We prepare data and fit parametric distributions for the partitioned survival model. In addition, we also define input parameters for the partitioned survival model. The model has a cycle length of one year, because of the long follow-up data.

2.1 Define model parameters

So far, we have called the arms: natural history of disease and hypothetical treatment arm, but for the cost-effectiveness analysis. We make the assumption that natural history of disease is best Best Supportive Care (BSC) and the hypothetical treatment arm is a hypothetical drug with best supportive care (Hypothetical Drug & BSC). Note that a state for Death is not added, because the R package inherently knows that this will be the last health state for a partitioned survival model.

Show

Treatment strategies

strategy_id	strategy_name
1	BSC
2	Hypothetical Drug & BSC

Show

Health states

state_id	state_name
1	Able to walk
2	Aid-required walking
3	Wheelchair confinement
4	Bedridden

Show

2.2 Parameter estimation for survival models

A dataset that is compatible with the hesim package is constructed. We fit Weibull models for each endpoint, because we previously found that the Weibull distribution was the best fit for all events.

Show

2.3 Define utilities and costs parameters for drug and health states

For the state able to walk, we assume from triangulation purposes that this state correspond to UMSARS IV item 1 + 2 (mostly or not completely independent). Similarly, we use the aggregated costs for UMSARS IV item 1 + 2 from

```
Winter et al 2011 Health-related quality of life in multiple system atrophy and progressiv
e supranuclear palsy' (DOI: 10.1007/s00415-011-6028-7).
```

The utility data is generated from individual patient level data, which cannot be shared, but is stored as a csv file. Moreover, we add disutilities directly in the code from literature to account for short-term events that may not have been captured by the six-monthly eq-5d measurements in the patient level data. However, for costs, we provide an R file that shows the generation of cost estimates.

Show

Utilities for health states

strategy_id	state_id	mean	sd	dist
1	1	0.4250447	0.2161165	beta
1	2	0.2736029	0.2682017	beta
1	3	-0.0101367	0.2737805	fixed
1	4	-0.2791564	0.2282108	fixed
2	1	0.5018729	0.2161165	beta
2	2	0.3504311	0.2682017	beta
2	3	0.0666915	0.2737805	fixed
2	4	-0.2791564	0.2282108	fixed

Show

Gamma Distribution Parameters for Health State Costs

State	Mean	SE	Distribution	Shape	Scale
Able to walk	11853	1748	gamma	46	257.6737
Aid-required walking	24002	3539	gamma	46	521.7893
Wheelchair confinement	44152	6510	gamma	46	959.8347
Bedridden	44152	6510	gamma	46	959.8347

Show

Treatment costs

strategy_id	state_id	est
1	1	0
1	2	0

strategy_id	state_id	est
1	3	0
1	4	0
2	1	12000
2	2	12000
2	3	12000
2	4	0

2.4 Constructing individual components of the cost-effectiveness model

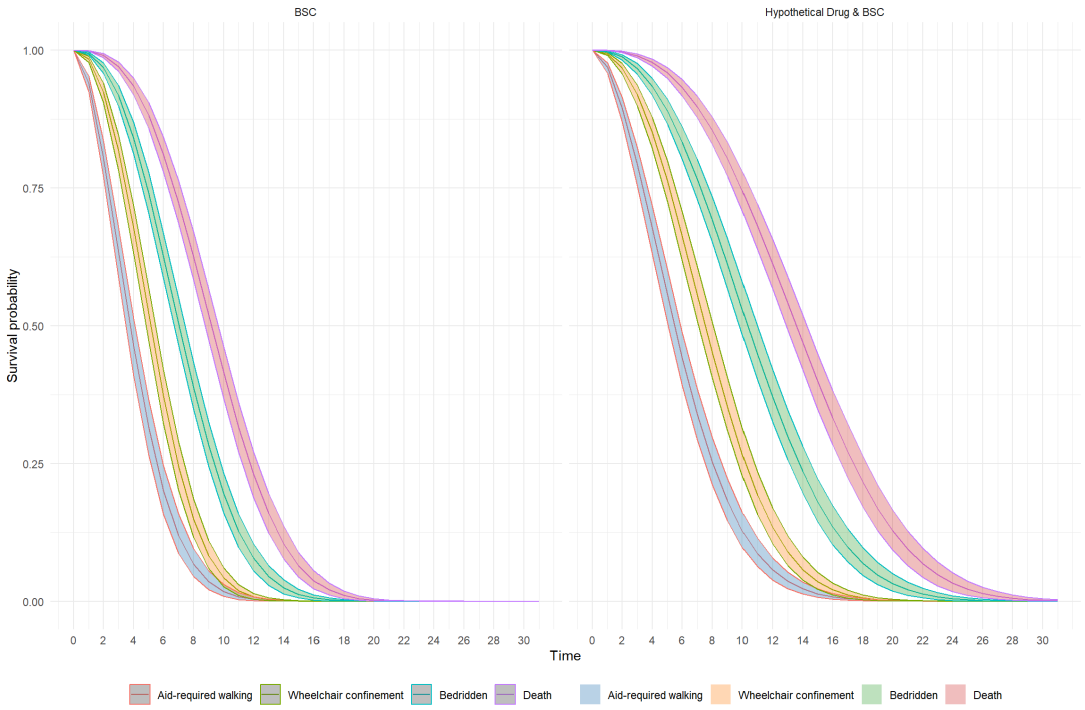
In accordance with the hesim package - models are constructed for sets of parameters and lastly bound together as a partitioned survival model.

Show

2.5 Simulate outcomes

We simulate survival curves generating a plot of survival curves averaged across patients by treatment strategy

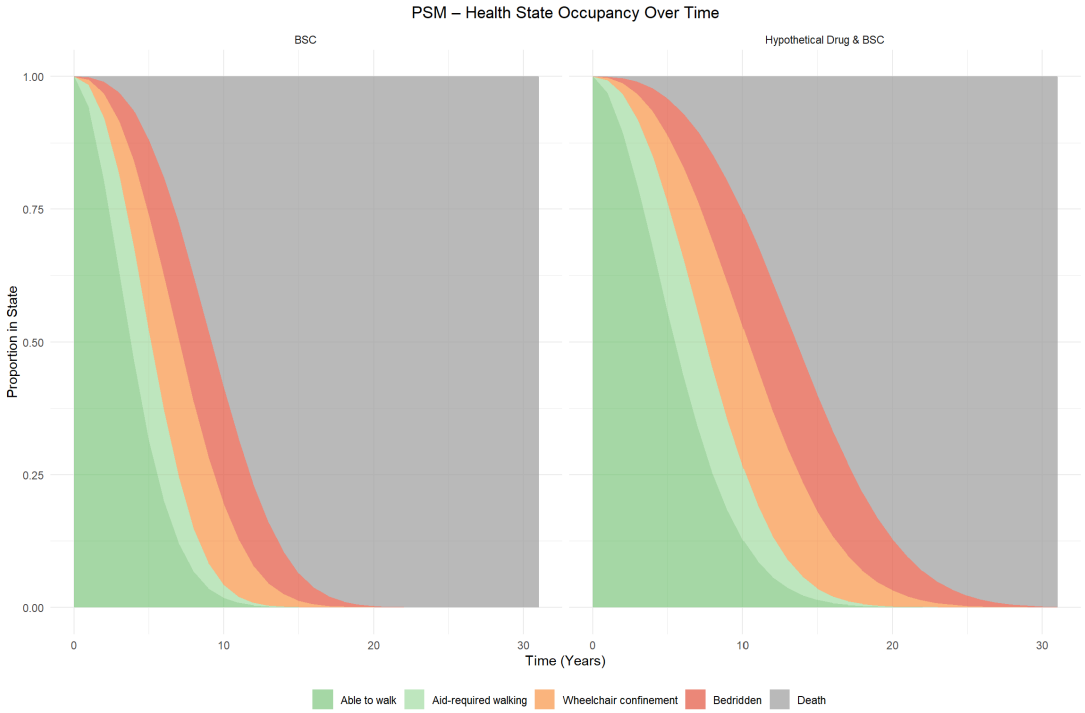
Show



2.6 Health state occupancies

We simulate state occupancy here, which is important for assigning utilities and costs at various time points.

Show



2.7 Probabilistic results

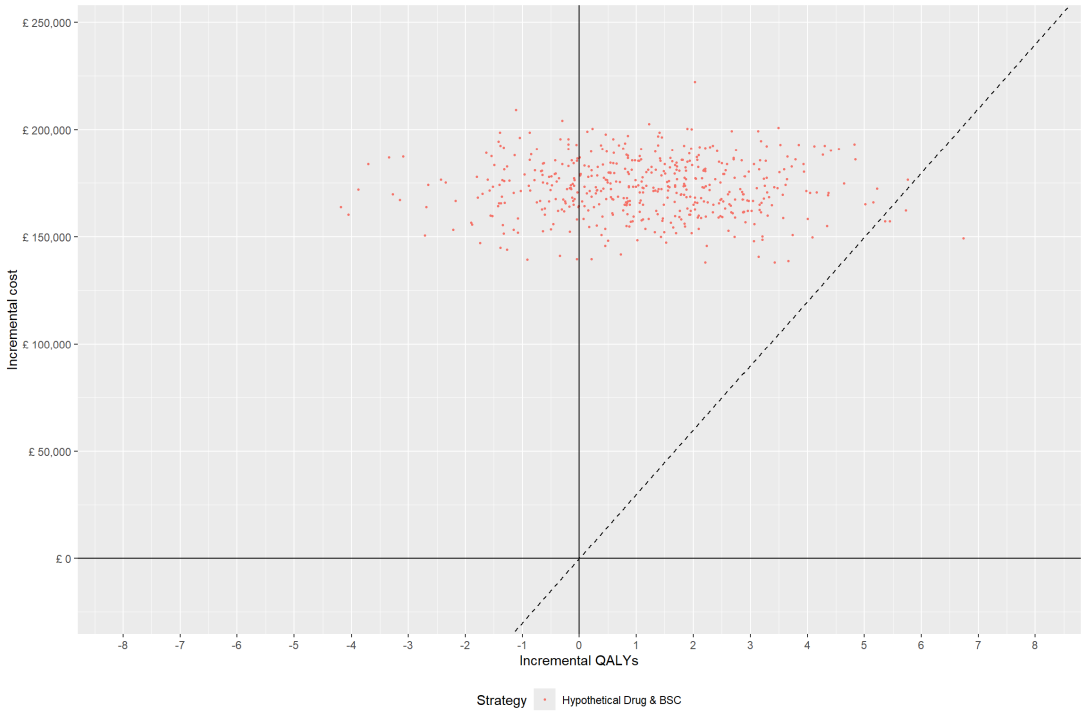
Results for the probabilistic cost-effectiveness analysis.

Show

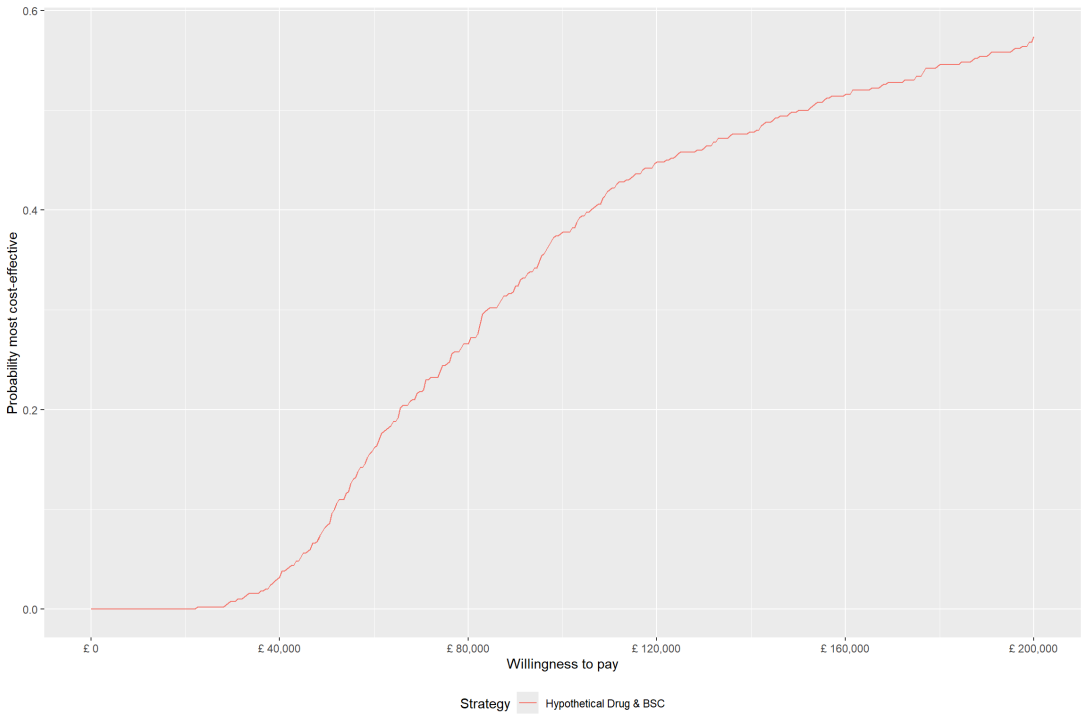
ICER(£ per QALY) = 151,944

Strategy	Costs	QALYs	ICER
Hypothetical Drug & BSC	374840	2.66	NA
BSC	201966	1.53	151944

Show


£ 250,000
£ 200,000
£ 150,000
£ 100,000
£ 50,000
£ 0
Incremental cost
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Incremental QALYs
Strategy
Hypothetical Drug & BSC

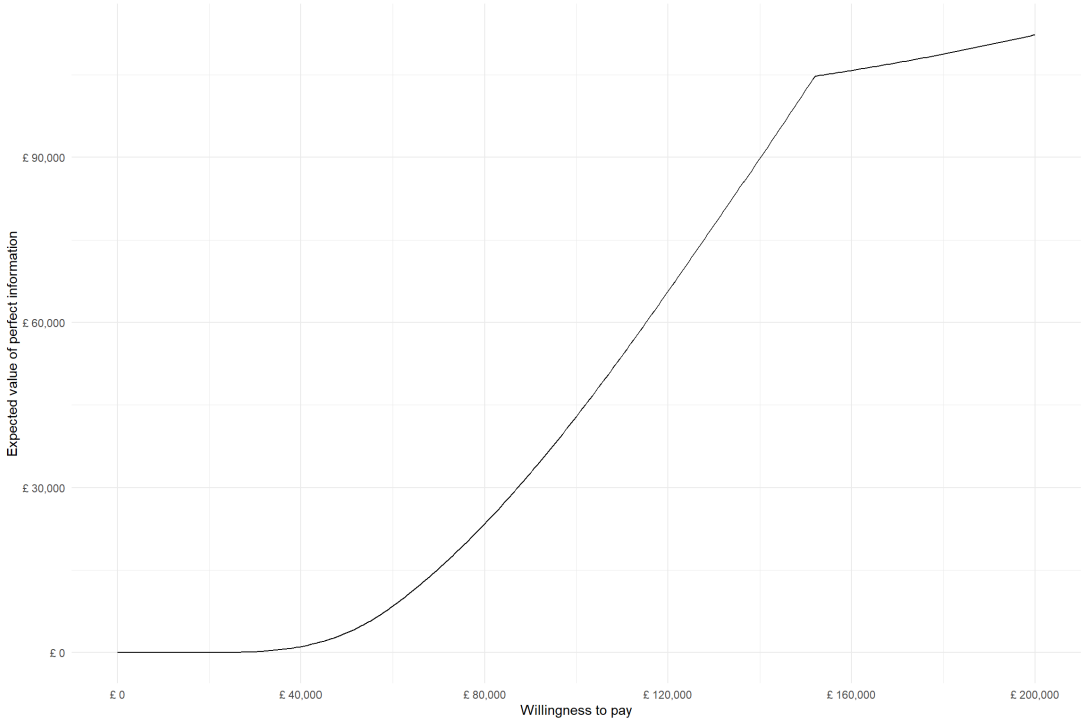
Show


0.6
0.4
0.2
0.0
Probability most cost-effective
£ 0
£ 40,000
£ 80,000
£ 120,000
£ 160,000
£ 200,000
Willingness to pay
Strategy
Hypothetical Drug & BSC

2.7.1 Expected Value of perfect information

In this section we conduct an analysis of the expected value of perfect information on the individual level.

Show



2.8 Deterministic results

We can estimate a deterministic result by simulating the model with a single sample without sampling uncertainty "none", which means that the model parameters are kept constant and no distribution for sampling is assigned.

Show

ICER(£ per QALY) = 131,380

Strategy	Costs	QALYs	ICER
Hypothetical Drug & BSC	357937	2.42	NA
BSC	189607	1.14	131380

3 Markov model

We code a Markov model for comparison, but first we need to prepare the data. The Markov model will need transition probabilities for each event rather than survival equations, but they will be calculated from the same data that was originally digitised. We will use the msm package to calculate transition probabilities. Other health-economic parameters are kept identical for comparison purposes. Similarly, for psa sampling we try to keep consistency where possible.

3.1 Data preparation

Data is modified to be used in the msm package. We need to ensure that the data is in long format, and that the time of each event is sequential.

Show

3.2 Calculate transitions matrices

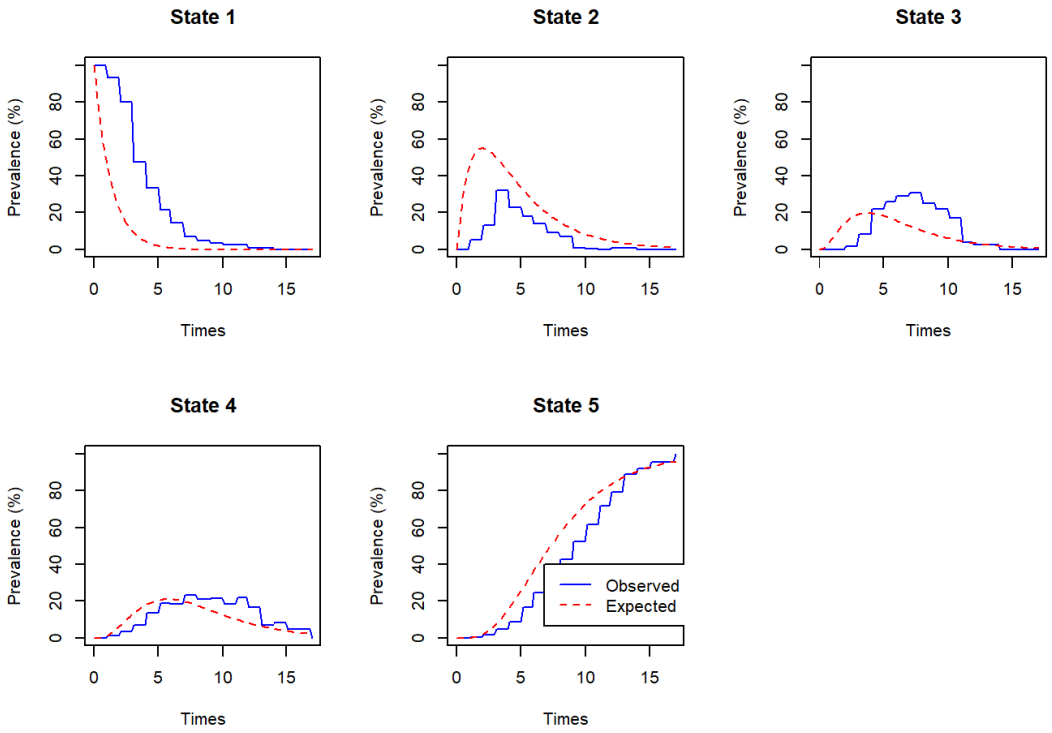
We calculate transition probabilities using multistate modelling or the msm package. From the Figures it should be obvious that the msm package estimates predicted prevalence by time, which means that one can specify time t, for any interval. The transitions are governed by the Q matrices, which is a snapshot of the Markov process, and despite having a sequential progression assumption. We will observe transitions that skip health states, because of cycle length is too long, which means that these transitions are estimated. These transitions would be much closer to zero if transitions were estimated with a cycle length of one month, but we are consistent with the assumption to compare across model types. So despite not allowing certain transitions in the Q matrix, they will nonetheless be illustrated in the model structure and discussed in the manuscript.

Show

Transition probabilities for BSC

	Able to walk	Aid-required walking	Wheelchair confinement	Bedridden	Death
Able to walk	0.4613327	0.4576580	0.0644013	0.0144915	0.0021165
Aid-required walking	0.0000000	0.7442716	0.1836230	0.0602442	0.0118612
Wheelchair confinement	0.0000000	0.0000000	0.5134126	0.3686135	0.1179739
Bedridden	0.0000000	0.0000000	0.0000000	0.5943891	0.4056109
Death	0.0000000	0.0000000	0.0000000	0.0000000	1.0000000

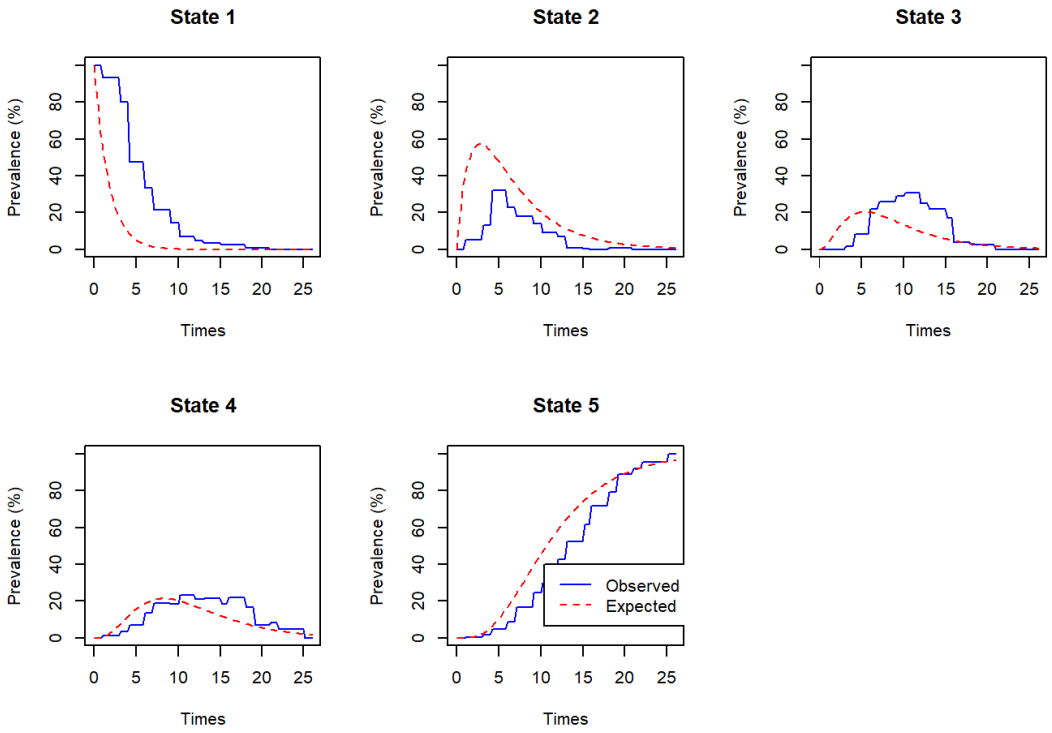
Show



Show

Transition probabilities for Hypothetical Drug & BSC

	Able to walk	Aid-required walking	Wheelchair confinement	Bedridden	Death
Able to walk	0.5566611	0.3982430	0.0387457	0.0058024	0.0005479
Aid-required walking	0.0000000	0.8199496	0.1443956	0.0316545	0.0040004
Wheelchair confinement	0.0000000	0.0000000	0.6420696	0.2983491	0.0595814
Bedridden	0.0000000	0.0000000	0.0000000	0.7057026	0.2942974
Death	0.0000000	0.0000000	0.0000000	0.0000000	1.0000000



3.3 Define model parameters

We define model parameters, and fetch those already defined in PSM.

Show

3.4 Markov traces

We need to set up an empty Markov trace and set initial distribution of patients.

Show

3.5 Run Markov traces

We make the calculations of state occupancy.

Show



3.6 We calculate costs and QALYs

We run the deterministic analysis for both strategies and calculate costs and QALYs.

Show

ICER(£ per QALY) = 179,362

Strategy	Costs	QALYs	ICER
Hypothetical Drug & BSC	375882	2.08	NA
BSC	208437	1.14	179362

3.7 Generate PSA inputs

Assign distributions to the parameters varied in PSA.

Show

3.8 We run probabilistic sensitivity analysis

We run the loop the sampled parameters through the Markov model.

Show

3.9 PSA results

We provide an ICER table of results, which is the average of our simulations.

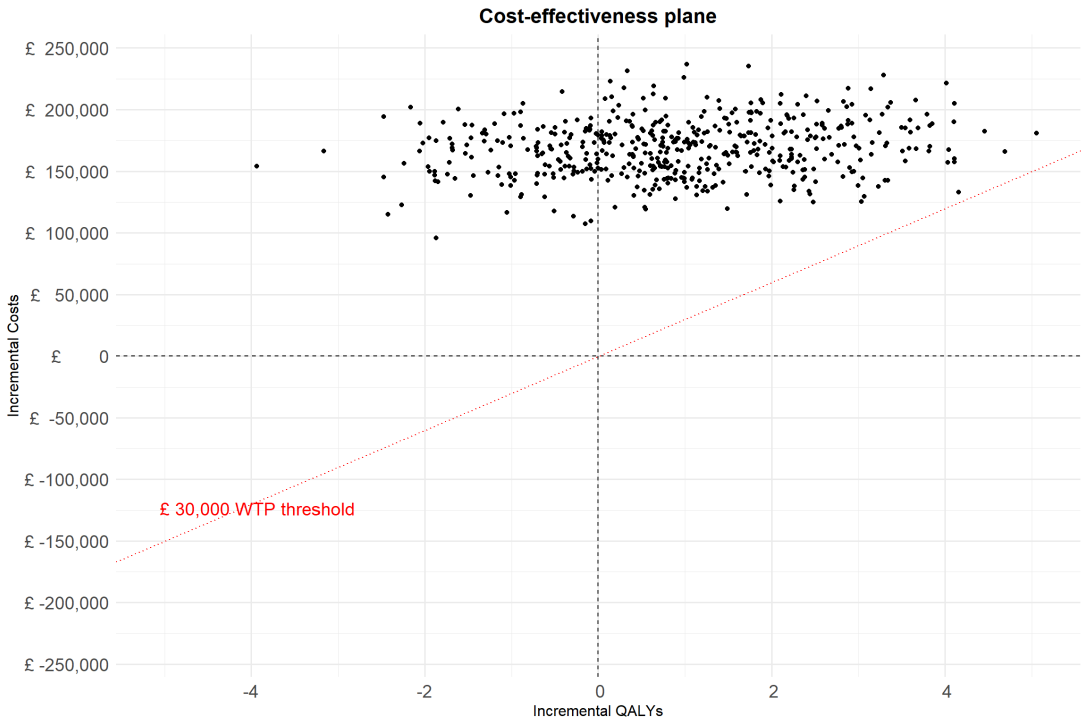
Show

ICER(£ per QALY) = 183,808

Strategy	Costs	QALYs	ICER
Hypothetical Drug & BSC	377759	2.10	NA
BSC	209328	1.18	183808

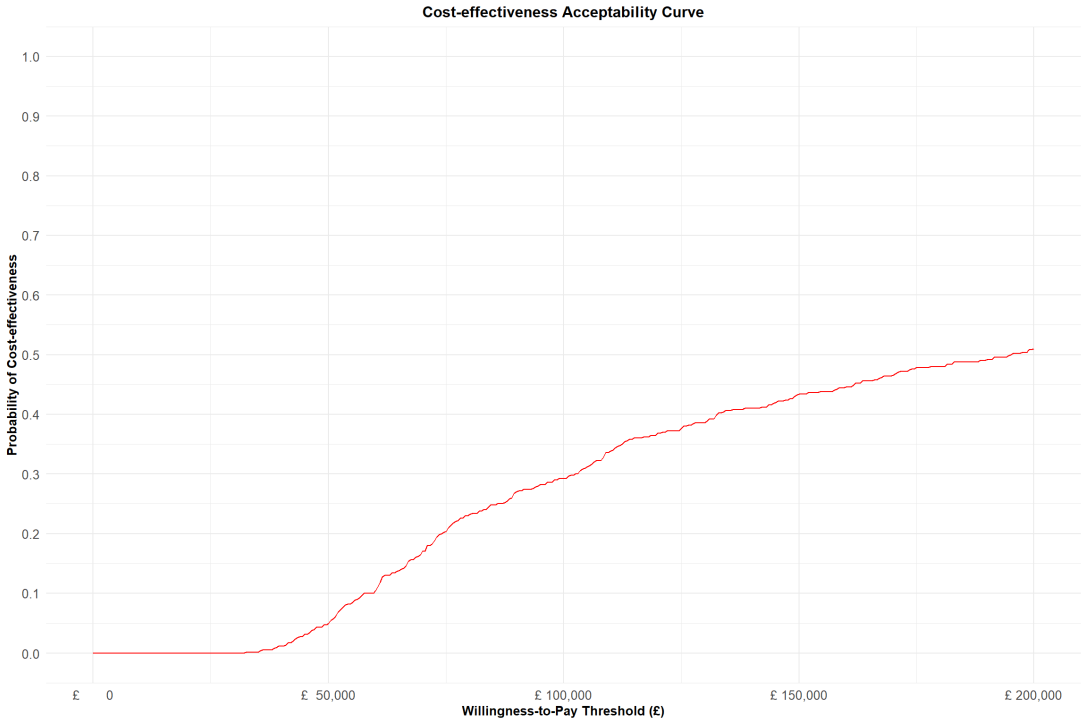
3.10 Cost-effectiveness plane

Show



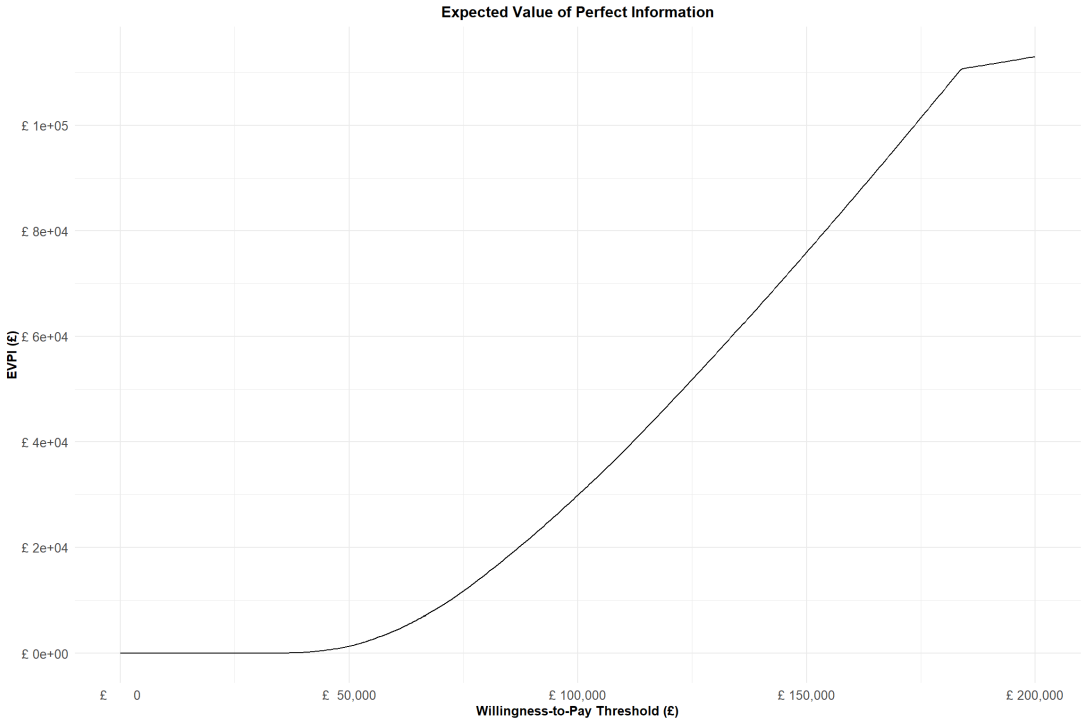
3.11 Cost-effectiveness acceptability curve

Show



3.12 Expected value of perfect information on the individual level

Show

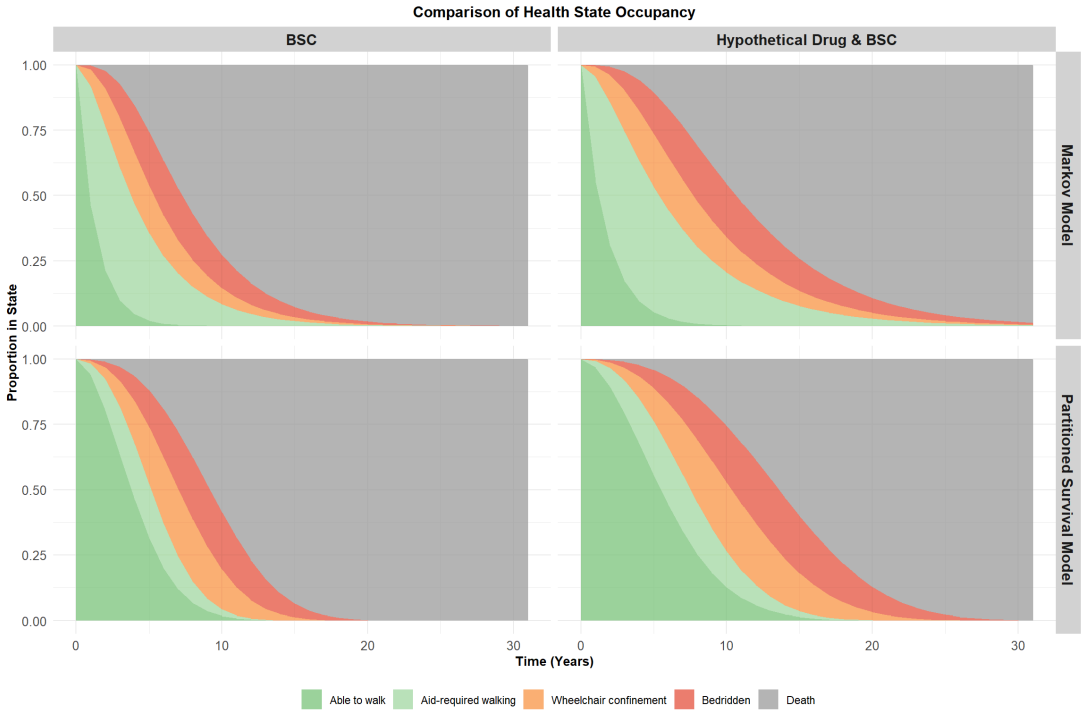


4 Visuals for the report

In this section we seek to combine the plots for illustration purposes.

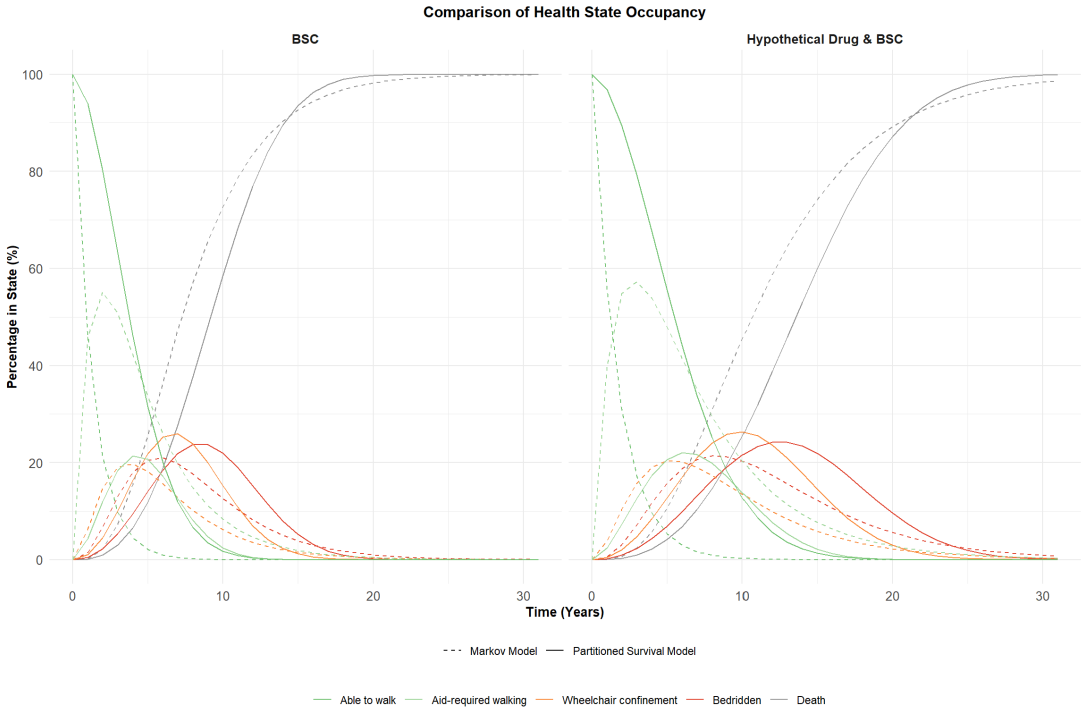
4.1 State Occupancy

Show



4.2 Alternative plot for Markov traces

Show



4.3 ICER tables

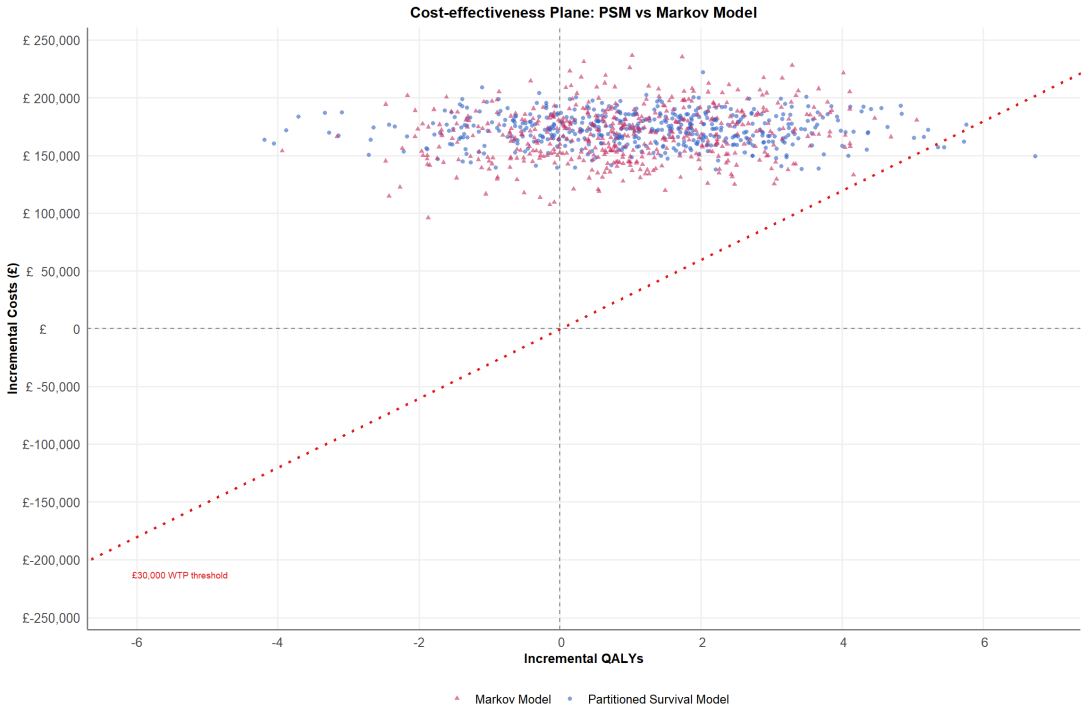
Show

Comparison of Cost-Effectiveness Results Between Models

Model	Analysis	Strategy	Costs	QALYs	ICER
Partitioned Survival Model	Deterministic	Hypothetical Drug & BSC	£357,937	2.42	NA
		BSC	£189,607	1.14	£131,380
	Probabilistic	Hypothetical Drug & BSC	£374,840	2.66	NA
		BSC	£201,966	1.53	£151,944
Markov Model	Deterministic	Hypothetical Drug & BSC	£375,882	2.08	NA
		BSC	£208,437	1.14	£179,362
	Probabilistic	Hypothetical Drug & BSC	£377,759	2.10	NA
		BSC	£209,328	1.18	£183,808

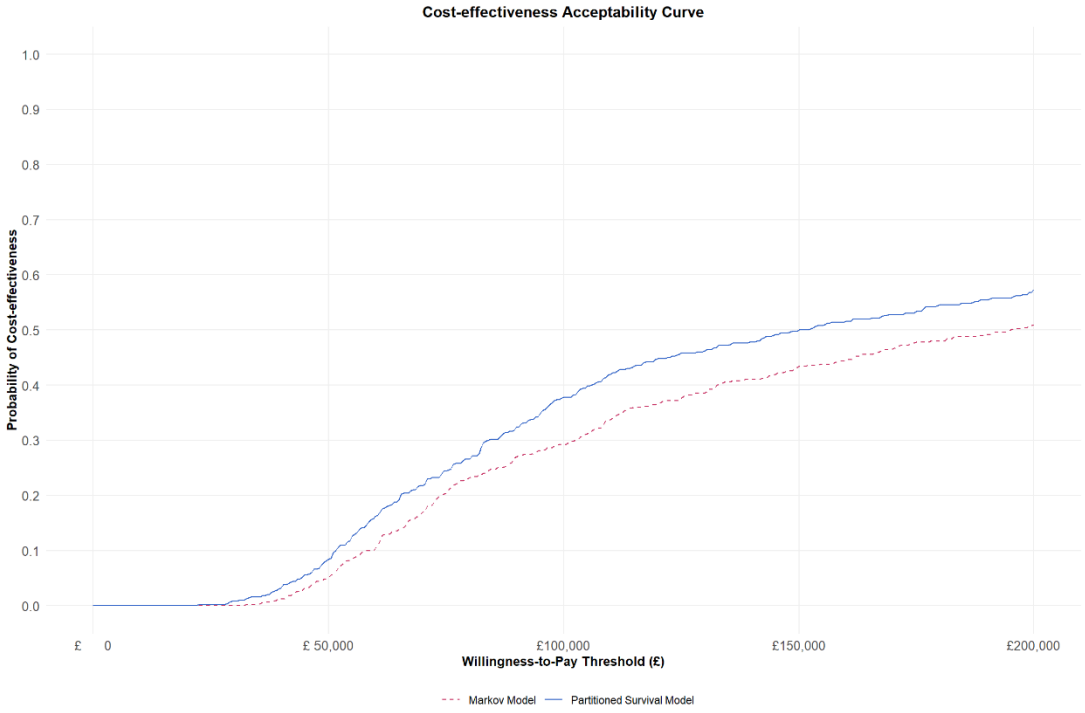
4.4 Cost-effectiveness plane

Show



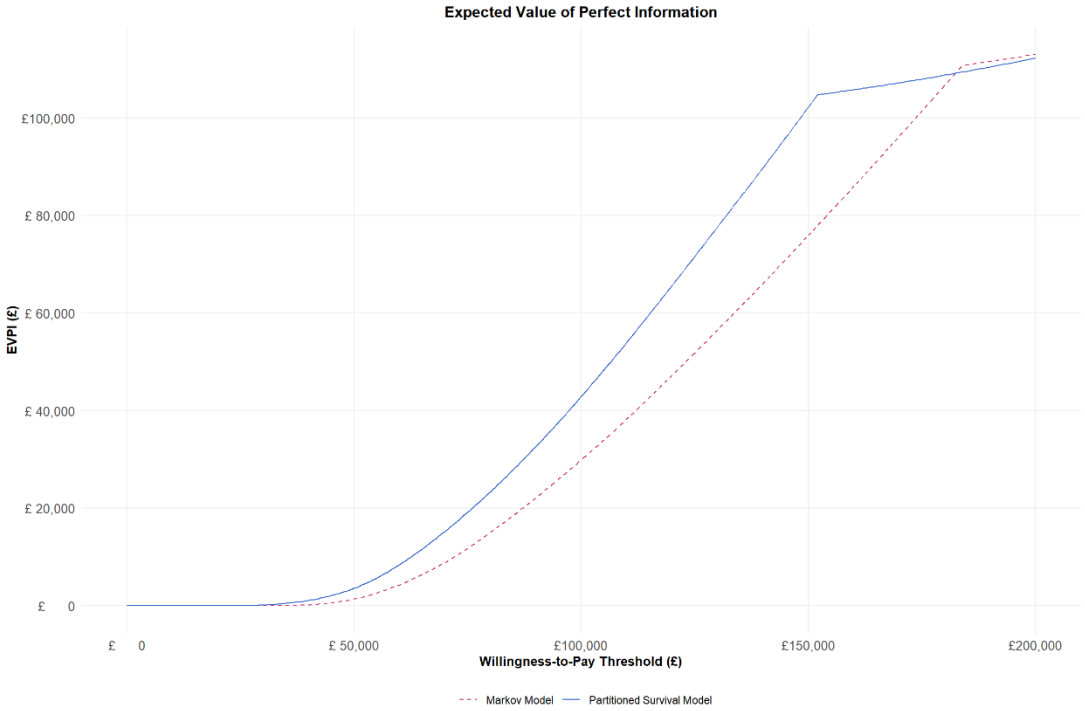
4.5 Cost-effectiveness acceptability curve

Show



4.6 Expected value of perfect information on the individual level

Show


Expected Value of Perfect Information
£100,000
£ 80,000
£ 60,000
£ 40,000
£ 20,000
£ 0
EVPI (£)
£ 0
£ 50,000
£100,000
£150,000
£200,000
Willingness-to-Pay Threshold (£)
Markov Model
Partitioned Survival Model

IV Appendix: Publications

It is our plan to publish three papers from this PhD. Two of the research projects are already published and the third is ready for submission. All three papers are included in the following sections.

IV.I Issues, Challenges and Opportunities for Economic Evaluation of Orphan Drugs in Rare Diseases: An Umbrella Review

The formatting is not high quality, which is why I suggest readers visit the online version: https://link.springer.com/article/10.1007/s40273-024-01370-2.

PharmacoEconomics
https://doi.org/10.1007/s40273-024-01370-2

REVIEW ARTICLE

Issues, Challenges and Opportunities for Economic Evaluations of Orphan Drugs in Rare Diseases: An Umbrella Review

Tobias Sydendal Grand[1,3] · Shijie Ren[1] · James Hall[2] · Daniel Oudin Åström[3] · Stephane Regnier[3] · Praveen Thokala[1]

Accepted: 3 March 2024
© The Author(s) 2024

Abstract

Background and Objectives There are significant challenges when obtaining clinical and economic evidence for health technology assessments of rare diseases. Many of them have been highlighted in previous systematic reviews but they have not been summarised in a comprehensive manner. For all stakeholders working with rare diseases, it is important to be aware and understand these issues. The objective of this review is to identify the main challenges for the economic evaluation of orphan drugs in rare diseases.

Methods An umbrella review of systematic reviews of economic studies concerned with orphan and ultra-orphan drugs was conducted. Studies that were not systematic reviews, or on advanced therapeutic medicinal products, personalised medicines or other interventions that were not considered orphan drugs were excluded. The database searches included publications from 2010 to 2023, and were conducted in MEDLINE, EMBASE and the Cochrane library using filters for systematic reviews, and economic evaluations and models. These filters were combined with search terms for rare diseases and orphan drugs. A hand search supplemented the literature searches. The findings were reported by a compliant Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

Results Two hundred and eighty-two records were identified from the literature searches, of which 64 were duplicates, whereas five reviews were identified from the hand search. A total of 36 reviews were included after screening against inclusion/exclusion criteria, 35 from literature searches and one from hand searching. Of those studies 1, 27 and 8 were low, moderate and high quality, respectively. The reviews highlight the scarcity of evidence for health economic parameters, for example, clinical effectiveness, costs, quality of life and the natural history of disease. Health economic evaluations such as cost-effectiveness and budget-impact analyses were scarce, and generally low-to-moderate quality. The causes were limited health economic parameters, together with publications bias, especially for cost-effectiveness analyses.

Conclusions The results highlighted issues around a considerable paucity of evidence for economic evaluations and few cost-effectiveness analyses, supporting the notion that a paucity of evidence makes economic evaluations of rare diseases more challenging compared with more prevalent diseases. Furthermore, we provide recommendations for more sustainable approaches in economic evaluations of rare diseases.

1 Background

The term orphan drug is recommended by the International Society for Pharmacoeconomics and Outcomes Research when a drug is indicated for the treatment of rare diseases with a prevalence threshold of 40–50 patients per 100,000 people [2]. The US Orphan Drug Act of 1983 and the European Orphan Regulation (No. 141/2000) have provided drug manufacturers with research incentives for rare diseases [3, 4]. They are widely regarded as successful and have led to an increase in orphan-drug designations [5–7]. For example, in the USA, the number of orphan designations more than quadrupled from the 1990s to 2010s [8].

Before the introduction of incentives, there was a widely held view that manufacturers should be rewarded for orphan-drug development, which in exchange, meant that they could claim prices that ensured profitability. Although drug prices were high, the impact on healthcare budgets was negligible because of few marketed orphan drugs, and patients to benefit from them [9]. The situation has now changed because of the policy-induced surge in orphan drugs, and

Extended author information available on the last page of the article

Published online: 14 April 2024

Key Points for Decision Makers

This umbrella review provides a comprehensive overview of current issues for economic evaluations of orphan drugs in rare diseases.

For economic evaluations of rare diseases, there is a paucity of evidence and a pronounced publication bias, as a result, few cost-effectiveness analyses exist for orphan drugs.

Stakeholders working with rare diseases can improve their work by following recommendations outlined in this umbrella review, for example, using comprehensive and flexible cost-effectiveness models.

both policymakers and researchers are attempting to find sustainable solutions to the issue of reimbursement [9, 10].

The fundamental aim of clinical trials is regulatory approval, which involves a risk-benefit evaluation that should answer whether the benefit of an intervention outweighs the risk [11]. It is difficult to obtain high-quality trial data when investigating rare diseases. For example, it may be hard to recruit enough trial participants, hard endpoints may be missing and including placebo arms in clinical trials may be unethical [12, 13]. However, these challenges are often magnified when it comes to health technology assessments (HTAs), where the aim is a systematic assessment of both clinical and cost effectiveness [14]. These challenges lead to high uncertainty for cost-effectiveness analyses and along with their high prices result in many orphan drugs not being recommended for reimbursement [12, 13].

Multiple authors have described economic evaluation challenges for rare diseases, focusing on various aspects such as the decision analytic modelling component of economic evaluations. Some of the most influential papers, based on a number of citations, are from 2018 [9, 12, 13]. However, the literature is diverse, with researchers and policymakers looking for ways to alleviate the challenges for economic evaluations of orphan drugs [15, 16]. Recent events include the introduction of the Innovative Medicines Fund in the UK that facilitates the collection of additional data for promising orphan drugs or a living HTA, which is the concept of continuous updating of economic models [17, 18]. The existing reviews are limited in terms of their ability to synthesise the most recent policy, economic and clinical developments because they have been superseded by recent developments. Consequently, the issues, challenges and opportunities associated with the economic evaluation of orphan drugs have not been summarised comprehensively. As a result, an umbrella review that focuses on the challenges for economic evaluations of rare diseases is warranted.

2 Methods

Scoping searches helped inform the literature searches [1]. They confirmed that the surge in orphan drugs had resulted in a growing and disparate field of literature. Ultimately, the decision to conduct an umbrella review was made, which in this case, was deemed as an appropriate solution. Umbrella reviews aim to synthesise systematic reviews, with or without meta-analyses, and have been described as a natural option to handle increases in systematic reviews to provide a summary of broad topic areas [19]. Previously, this approach proved useful in similar situations, where fields of research expanded rapidly, and consequently, resulted in a diffuse body of literature [20–22].

2.1 Research Objectives

This research was informed by a modified version of the Setting, Perspective, Interest, Comparison, Evaluation (SPICE) framework [23]. The perspective component was omitted because all perspectives were considered relevant. When applying the framework with its parameters in brackets, for example, [Setting]. The research question became: in health-economic-research settings [Setting] are there any issues and challenges [Evaluation] for the economic evaluation of orphan drugs in rare diseases [Interest], which apply less to other drugs [Comparison]?

2.2 Literature Searches

The most relevant databases for the umbrella review were MEDLINE, Cochrane and EMBASE. Thus, during January 2023, MEDLINE and EMBASE were accessed through the Ovid platform and Cochrane independently through its website. For both MEDLINE and EMBASE, search filters for economic evaluations and models, and systematic reviews were sourced from the Canadian Agency for Drugs and Technologies in Health and Scottish Intercollegiate Guidelines Network databases, respectively [24–26]. These filters were combined with search terms for orphan drugs and rare diseases. Eligibility criteria, scoping and literature searches are available from the Electronic Supplementary Material (ESM).

As recommended by Booth and colleagues, a hand search of references and bibliographies of papers from the review was conducted [27]. This was followed by a verification process where it was checked if any known and relevant papers were missing from the review.

2.3 Data Collection Process

Titles and abstracts were screened by two independent researchers against the inclusion and exclusion criteria. Discrepancies were discussed until a consensus was reached.

The papers that met the inclusion criteria, after screening of the title and abstract, were further subjected to full screening. Papers were also excluded at full screening if they were deemed as containing insufficient information to allow for meaningful data collection, for example, abstracts. The data collection process was divided into three steps: summary of characteristics, critical appraisal and data extraction.

2.3.1 Summary of Characteristics

An extraction table captured summary characteristics recommended for umbrella reviews: citations details, type of review, objectives, date range of database searching, number of studies, rating by the Joanna Briggs Institute checklist and themes [19].

2.3.2 Critical Appraisal

For the critical appraisal, the Joanna Briggs Institute critical appraisal checklist was used. This checklist is recommended by the umbrella review methodology working group for a critical appraisal of systematic reviews [19]. The checklist contains 11 questions that were used to critically appraise the reviews [28]. For this tool, there is high degree of freedom for deciding on a scoring system for the inclusion or exclusion of papers. To avoid missing any information, it was decided not to exclude any papers based on their scores. The reviews were divided into three levels according to their quality scores: 8–11 (high quality), 4–7 (moderate quality) and 0–3 (low quality) [29, 30].

2.3.3 Data Extraction

The included reviews were carefully assessed with the aim of identifying broader themes that pertain to economic evaluations of orphan drugs. Challenges were extracted and tabulated according to their themes, based on an approach previously used to extract modelling challenges for rare diseases [12].

3 Results

3.1 Literature Search and Study Selection

The study selection is illustrated by a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram in Fig. 1. The number of identified records was 282. They were retrieved from the following databases: EMBASE ($n = 211$), MEDLINE ($n = 67$) and Cochrane Library ($n = 4$). A total of 64 duplicate records were removed. Moreover, 172 records were excluded during screening of the abstract and title, which left 46 studies for full screening. Of those, 11 were excluded because of not containing components for economic evaluations ($n = 5$) or systematic reviews ($n = 4$), or because they were abstracts ($n = 2$). Overall, 35 reviews from the database searches were deemed eligible for inclusion as listed in the ESM under literature search results.

The hand search yielded five papers, of which four were excluded for the following reasons: no component of economic evaluations ($n = 1$) or systematic reviews ($n = 2$), and for not being concerned with orphan drugs ($n = 1$). It meant that one paper was carried forward from the hand search, which brought the total number of eligible reviews to 36. The ESM lists papers included for full screening.

3.2 Study Characteristics

A two-step approach was used to determine if studies could qualify as systematic reviews. First, a Scottish Intercollegiate Guidelines Network search filter for systematic reviews was used, which is a pre-tested search strategy that identifies the higher quality evidence from vast amounts of literature indexed in a medical database. Second, eligibility was assessed, and a consensus obtained between the first and second reviewer on their inclusion. Using this approach, two scoping reviews were included because the methods were sufficiently systematic [31, 32]. Similarly, a study described their approach as a series of targeted literature reviews, which was also sufficiently systematic for inclusion [12]. The number of records included in the systematic reviews varied between 2 and 338. The ESM provides a summary of study characteristics.

3.3 Critical Appraisal

One study had low quality [31]. The highest frequency was found in the category of moderate quality, which comprised 27 studies [12, 32–57], whereas eight studies were rated as having high quality [58–65]. The ESM includes scores for each individual Joanna Briggs Institute checklist question across all studies, which showed that most studies ($n = 35$) obtained points from question 4, which was: Were the sources and resources used to search for studies adequate? Question 8 was not widely applicable and was fulfilled by the least studies ($n = 4$). Question 8 was: Were the methods used to combine studies appropriate? Critical appraisal methods used in individual systematic reviews were assessed by question 5: Were the criteria for appraising studies appropriate? Fourteen studies included appropriate criteria for critical appraisal, whereas in 13 studies it was unclear whether they did, seven studies did not, and for two studies the question was not applicable.

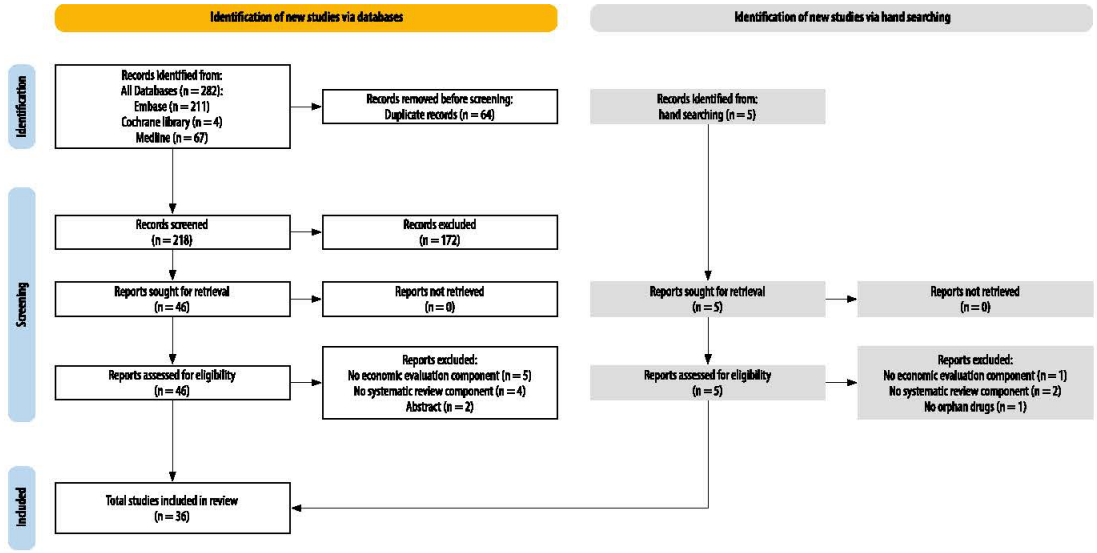


Fig. 1 Preferred reporting items for systematic reviews and meta-analyses (PRISMA) flow diagram

3.4 Data Extraction

The systematic reviews were divided into two categories: those that considered a specific rare disease (13 studies) and those that considered multiple rare diseases (23 studies). As shown in Fig. 2, three broad themes were identified: issues with health economic parameters, issues with health economic evaluations and issues with estimating value/reimbursement, with subtopics further developed for each theme. For issues with health economic parameters, the subtopics were the natural history of disease, clinical effectiveness, costs and quality of life. For issues with health economic evaluations, the subtopics were cost effectiveness and budget impact. For issues with estimating value or reimbursement, the subtopics were thresholds, value frameworks and multiple criteria decision analyses (MCDA). A repository of all extracted data on issues for economic evaluations of rare diseases is available in the ESM.

3.5 Issues with Health Economic Parameters

3.5.1 Natural History of Disease

Rare diseases often progress slowly or are chronic by nature, which make clinical trials insufficient as they tend to have short durations [12, 62]. The non-existence or limited number of studies that include data on prevalence and incidence further magnify issues [44, 48, 51]. Moreover, clinical experts are few and private practitioners may only encounter few rare disease cases, which make them difficult to diagnose, and expert advice on rare diseases might not be easy to find [12, 44, 45, 57]. Delayed diagnosis and misdiagnosis make it difficult to define treatment-eligible cohorts [45, 50, 51, 58].

To summarise, an economic evaluation is challenging, for example for long-term modelling, because of missing data on the natural history of the disease or unknown rare disease trajectories [40]. Although registries can alleviate data issues, they may suffer from challenges such as diverging disease and diagnostic codes, data ownership and missing comparator data [12, 48, 50].

3.5.2 Clinical Effectiveness

Whilst clinical trials are common sources for effectiveness data in economic evaluations, appropriate clinical evidence is not always available for this purpose [57, 58]. Moreover, clinical trials may suffer from short durations, small sample sizes, premature termination, inadequate power, missing data or missing control arms, for example, for ethical reasons [12, 37, 45, 47, 57]. In addition, published long-term studies providing post-marketing data on safety and efficacy are rarely available [37, 38].

Other challenges are missing treatment guidelines, data to predict treatment responses, concerns on the patient relevance and the use of surrogate endpoints [40, 50, 52, 60, 63]. Comparator data are essential for economic evaluations,

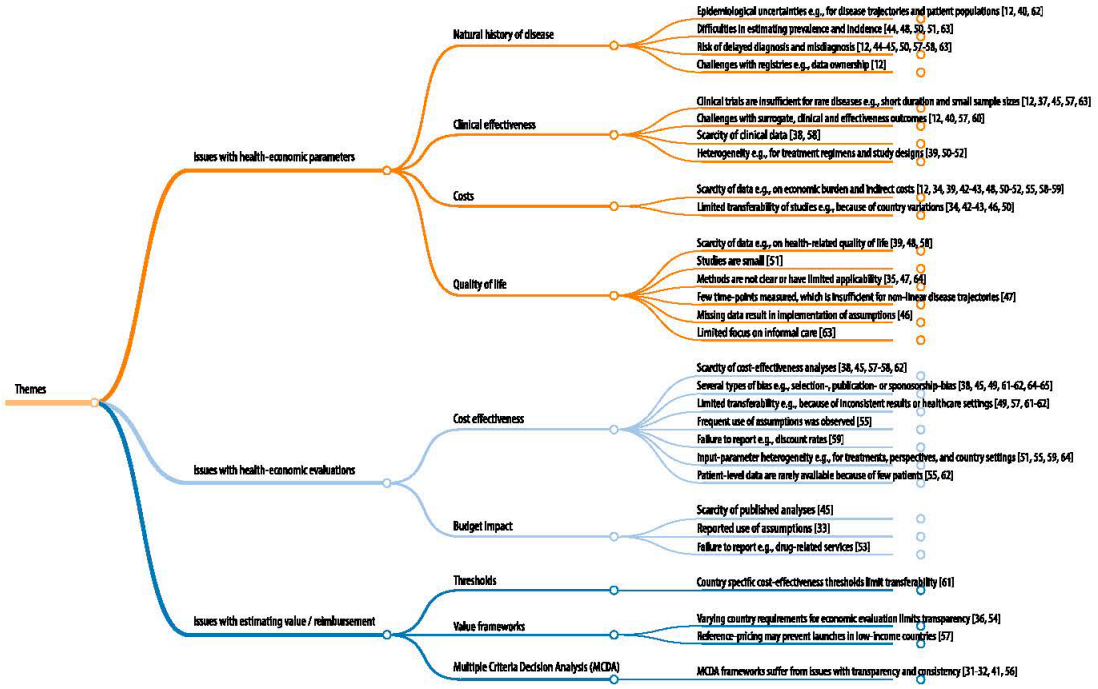


Fig. 2 Data extraction themes, sub-topics and findings

but might be missing for rare diseases, and if they are available, there might not be a consensus on the use of treatment regimes or treatment eligibility of patients, which result in heterogeneity across studies [39, 48, 50, 51]. A review found that studies reporting clinical evidence for orphan drugs had low-to-moderate quality, and none of them had high quality [60].

3.5.3 Costs

Cost-of-illness or burden-of-disease studies are scarce in rare diseases [12, 34, 39, 42, 43, 48, 52, 55]. Of those studies available, most are retrospective and only a small proportion of studies report indirect, non-medical or informal-care costs. [12, 34, 51, 58, 59]. Aggregated primary data are rarely available, hence, studies tend to report patient-reported claims or registry data [42].

It is complicated to transfer cost-of-illness results between different rare disease settings because of differences in study designs, methods and results. For example, one study estimated lost productivity without following recommendations for handling uncertainty [42]. A multitude of factors influence transferability such as data sources, geographical perspective, nomenclature, assumptions, discount rates, unit costs, treatment guidelines and value frameworks [34, 43, 46, 50].

3.5.4 Quality of Life

Quality-of-life studies in rare diseases are limited, but availability depends on the rare disease of interest [35, 39, 47]. For example, a review found two studies that included utility values for Cushing's syndrome, whereas another review concerned with Crigler–Najjar syndrome found no data on the humanistic burden, apart from anecdotes on treatment challenges [39, 47]. In addition, there are data limitations on the quality of life of caregivers [63]. A probable explanation for the scarcity is the limited applicability of quantitative methods such as choice experiments or conjoint analysis in rare diseases, for example because of small sample sizes [35]. Furthermore, studies tend to be small, not randomised or controlled, which decreases the reliability of conclusions [51]. This scarcity of evidence may lead to the use of assumptions, for example, assumption of equal utility values across treatment arms or a linearity assumption of utilities between different timepoints [46, 47]. Moreover, the reviews highlight shortcomings in methods and reporting, for example, the failure to include utility values or mapping

algorithms, and insufficiently describing the elicitation of utility weights [58, 64].

3.6 Issues with Health Economic Evaluations

3.6.1 Cost Effectiveness

Health economic evaluations for rare diseases are scarce. For example, a systematic review failed to identify any studies, whereas another noted a remarkable absence of pharmacoeconomic evidence [45, 58]. A notable opinion on the cause of scarcity is that limited information on input parameters simply deter people from attempting to construct cost-effectiveness analyses because it is presumed unachievable [62]. In brief, causes are missing patient-level data, high drug costs and the inability to measure effects for clinical or quality-of-life outcomes [55, 57, 62].

The difficulties for economic evaluations are driving factors for the use of assumptions to overcome challenges for cost-effectiveness modelling. For example, assumptions on mortality, efficacy, treatment and complications [55]. It is commonplace to use modelling techniques such as mapping algorithms or long-term extrapolation for outcomes because of data limitations [38, 47]. Moreover, limited patient numbers coupled with unreliable estimates of effects, symptoms and complications suggest that methods such as patient-level simulation modelling may have limited applicability in rare diseases [62].

Additionally, publication bias in relation to positive results or industry-sponsorship bias seems to be prominent in rare diseases [66, 67]. It may occur when manufacturers decide to publish only if they have favourable cost-effectiveness results, a post-marketing obligation, or an opportunity to adopt favourable input parameters and an advantageous interpretation of results [45, 62, 65]. Numerous reviews suggest issues of publication bias [38, 45, 49, 61, 65]. For example, Schuller et al. indicated a higher frequency of analyses in countries with post-marketing obligations [62]. Others found that studies failed to discuss the direction and magnitude of bias, despite using data from potentially biased sources [38, 61, 65]. Another review highlighted selection bias to explain conflicting cost-effectiveness results for a particular drug [49]. Additionally, it was highlighted that most studies were industry funded in a systematic review of cost-utility analyses for haemophilia [64]. Furthermore, incremental cost-utility ratios were significantly lower when published by industry compared with foundations and academia [49].

Most economic evaluations have moderate quality, and the failure to reach high quality may be partly attributed to a lack of good-quality model inputs (e.g. utility values that do not account for patient characteristics and disease severity) or because they omit lifetime horizons for chronic rare diseases [55, 59, 61]. Moreover, problems with reporting are frequently highlighted as another factor that may contribute to insufficient quality. For example, not adequately reporting discount rates, sensitivity analyses, utility weights, patient characteristics, funding sources and time horizons [38, 59, 64].

Transferability is another issue for cost-effectiveness results [57]. Cost-effectiveness analyses are heterogenous because of modelling variations in treatments, patient populations, time horizons, countries, cost-effectiveness thresholds, settings, year of analysis, comparators and assumptions [49, 51, 55, 59, 61, 62, 64]. Thus, a high degree of carefulness is advised when assessing the transferability of results across different healthcare settings [61].

3.6.2 Budget Impact

Studies on budget impact modelling are few, mostly from high-income or native English-speaking countries. If Kanters and colleagues' suggestion is accurate, it is not possible to rule out publication bias as a cause for the scarcity of studies on budget-impact modelling [45]. Furthermore, they are low quality and show poor adherence to guidelines [33, 45]. A proportion of budget-impact studies fail to report side effects, drug-related services, life-extension costs, savings from mortality reductions and validation methods [33, 53]. The importance of assumptions should not be overlooked, which are frequently incorporated for target populations, population sizes, interventions, comparators, costs and market uptake [33].

3.7 Issues with Estimating Value and Reimbursement

3.7.1 Value Frameworks and Thresholds

Most countries require budget-impact and cost-effectiveness models as part of HTAs, but the appraisal process (e.g. cost-effectiveness thresholds) may vary across countries, thus making comparisons difficult. As mentioned, whilst evidence may be scarce, input parameters on prevalence, incidence, number of treatment-eligible patients, and clinical benefits are nonetheless needed when estimating the budget impact and cost effectiveness for rare diseases [54]. For Europe, reference pricing further adds to the complexity and may prevent launches of orphan drugs in low-income countries [57]. Overall, value frameworks may suffer from transparency and consistency issues. This largely makes budget-impact and cost-effectiveness analyses country specific [36, 61].

3.7.2 Multiple Criteria Decision Analysis

A multiple criteria decision analysis (MCDA) is an emerging value framework for orphan drugs because it offers an opportunity to include a broad range of value criteria, for example, societal, disease or treatment criteria [31, 41]. Critics highlight variations in scoring functions for value criteria as a significant limitation and for decision making it is difficult to observe consistent recommendations [41, 56]. Interestingly, by meticulous examination of value criteria weights and scores in MCDAs, Friedmann and colleagues suggested that traditional value aspects used in HTAs (budget impact and cost effectiveness) were considered unimportant by stakeholders involved in orphan drug appraisal processes. The most cited value criterion was disease severity ($n = 10$), cost effectiveness ($n = 7$) and budget impact ($n = 3$) were cited ten times, collectively [41]. By contrast, Mohammadshahi and colleagues found in their review an equal citation frequency for the value criteria: disease severity ($n = 8$), cost effectiveness ($n = 8$) and budget effect ($n = 8$) [32].

4 Discussion

This section discusses the umbrella review findings, which indicated multiple issues for the economic evaluation of orphan drugs in rare diseases. However, it was not possible, with confidence, to assert whether all issues for orphan drugs applied less to other drugs, which was part of the original research objective [1]. Many papers focused on the evidence for a specific disease or multiple diseases, rather than how it compares to other drugs. For example, a systematic review of available evidence on 11 high-priced inpatient orphan drugs found that study populations were significantly smaller in randomised trials for orphan drugs as compared with non-orphan drugs [45]. Other systematic reviews in rare diseases confirmed that study populations were small but did not compare to other drugs [12, 37, 57]. The magnitude of issues varies, and this is the case for orphan drugs and other drugs. Thus, some of these issues may also be applicable to other drugs; however, these issues are critical in the case of orphan drugs as the issues tend to be amplified. In acknowledgement of this inability to consistently compare to other drugs, the ESM provides an indication of commonality for issues with economic evaluations of orphan drugs.

4.1 Issues with Health Economic Parameters

Scarcity of evidence was reported for natural history of the disease, clinical effectiveness, costs and quality of life [12, 34, 39, 42–45, 47, 48, 51, 52, 55, 57, 58]. It was previously pointed out that there were simply no easy answers to the problem of assessing evidence for orphan drugs [9]. In this review, this was exemplified by analysts who expressed a hope, rather than an actionable plan, for better availability of clinical trials with longer time horizons to conduct a thorough analysis of cost effectiveness, for example, for paediatric pulmonary arterial hypertension [37]. Others have suggested that high drug prices and the inability to measure effects would discourage people from even attempting to construct cost-effectiveness analyses [62]. This interpretation contrasts with that of Picavet and colleagues who conclude that orphan drugs can meet traditional cost-effectiveness thresholds [49]. It is an option to use expert opinion if few data are available, although it may be difficult to obtain [68, 69].

Some strategies may help improve evidence sources, but most do require extensive resources. For example, registries have the potential to inform modelling on the natural history of disease or can help construct a replacement for the standard of care, which may be relevant for trials without a control arm [12, 62, 63]. In addition, surrogate markers can play a vital role when clinical trials have short durations, they may, however, be difficult to validate without long-term data [57]. Analysts have drawn attention to this matter and highlighted the importance of consulting experts and to source data from other similar diseases to fill data gaps, for example, quality of life associated with wheelchair confinement between multiple sclerosis (more prevalent) and Duchenne's disease (less prevalent) [12]. Last, authors suggest investigating the geographical variation in treatment patterns, reporting of side effects, long-term trials in disease areas with little evidence and a Cochrane review group dedicated to systematic reviews that reduce evidence gaps for orphan drugs [37, 48, 60].

For cost-of-illness studies in rare diseases, first, the studies should be clear on their perspective; second, they should report indirect costs separately from direct costs, for example, lost productivity; third, they should report costs associated with prevented comorbidities; and fourth, they should provide clarity on applied discount rates [34, 42, 59, 63]. The importance of future research for informal care, in terms of costs and quality of life, was highlighted by multiple authors because rare diseases may have severe implications for the closest providers of care, for example, family and friends [34, 55, 63].

4.2 Issues with Health Economic Evaluations

Systematic reviews reported a scarcity of cost-effectiveness modelling studies [45, 58]. As alluded to earlier, it could suggest a strong link between evidence issues, publication bias and the observed paucity of cost-effectiveness analyses [62]. Researchers want economic evaluations with higher quality and extended time horizons [61]. To achieve this aim, without conducting a clinical trial, one could evaluate:

entry-level agreements and registries for data collection, patient surveys to assess the burden of disease, Delphi techniques for validation, expert opinion for estimation, population-adjusted indirect comparisons to account for patient characteristics and rare events with high costs [12, 64].

The explanations for the paucity of budget-impact models may be in terms of input parameters, for example, issues around a lack of data for prevalence or incidence estimation could contribute to their paucity [48, 51]. Budget-impact models were low quality and rarely validated. Summarising recommendations for improvement, they simply were that researchers should adhere to guidelines [33, 70]. Furthermore, publication bias for budget-impact models cannot not be ruled out [45, 54]. Health technology assessment bodies often require them, but for manufacturers, being the cause of increased healthcare costs might not be a message worth communicating, thus providing an explanation for potential publication bias. It is plausible that the budget impact is less of a concern for rare diseases because a low prevalence can translate to a lower impact on budgets for payers, thus providing another explanation for the scarcity of publications.

4.3 Issues with Estimating Value and Reimbursement

The appropriateness of value frameworks in the context of rare diseases is debated. For traditional value frameworks, examples of proposed solutions are: weighting of quality-adjusted life-years according to disease severity and prevalence, categorising quality-adjusted life-years based on disease states, implementing higher cost-effectiveness thresholds and special rules for those that exceed thresholds, for example, managed entry-level agreements and stopping rules for cost containment [12, 57]. The UK is an example where some of these measures have been incorporated through the Innovative Medicines Fund for medicines that are promising but associated with high uncertainty or decision modifiers through highly specialised technology appraisals [17, 71].

As highlighted throughout this review, criticism of traditional value frameworks has partly been related to their limited transparency and transferability of results. Critics have suggested policymakers explore other frameworks, for example, MCDA. So far, this method has only seen sporadic implementation, but it is clearly emerging [31, 41]. The benefit of MCDA is the ability to include a range of value criteria, for example, the burden on caregivers [36, 41]. However, like traditional frameworks, transferability and transparency for MCDA are areas that warrant further research [41, 56]. However, it should be noted that using a different value framework will not solve the problem of evidence scarcity.

5 Recommendations

Challenges are abundant and solutions are not plentiful and rarely forthcoming. Stakeholders, however, must recognise that certain types of research are costly, and demanding these could further eliminate company incentives to research rare diseases [57]. For example, clinical trials with extended time horizons. Thus, there is a need for recommendations that are more sustainable. As a first step towards these, we provide practical recommendations that may help alleviate challenges identified in this umbrella review.

5.1 Comprehensive and Flexible Cost-Effectiveness Models

Data availability is critical at the time of economic evaluations for rare diseases, this is why economic models should be transparent, and uncertainty rigorously explored through sensitivity analyses and set up for continuous updating as data become available over time [59]. Continuous updating of cost-effectiveness models with new data is an unexplored opportunity, especially considering the necessity of post-launch monitoring or real-world data [12, 60]. Such a framework has been referred to as a living HTA [18, 72].

Furthermore, transparency may increase for other stakeholders who are not trained researchers because user-friendly interfaces, for example, Shiny apps in the software R, allow them to "safely" explore model scenarios without having to face backend code [73]. For risk-sharing agreements, rather than focussing purely on clinical endpoints, for example, survival, they could potentially allow for fully updated cost-effectiveness models.

Consequently, for economic evaluations of rare diseases, there is untapped potential for using living HTAs. What is more, it has been recommended to use cost-effectiveness models in rare diseases to facilitate an expected value of information analysis using inputs from, for example, phase II or registry data [12]. It provides researchers with an opportunity to address the root causes of uncertainty by reprioritising or initiating data collection efforts, for example, before initiation of a phase III trial or an HTA [74].

In summary, we recommend using comprehensive and flexible cost-effectiveness models, which report value of information as initially suggested by Pearson and colleagues, which should as a minimum include both expected value of perfect information and expected value of perfect parameter information [12].

5.2 Publication Bias and Ability to Meet Cost-Effectiveness Thresholds

In the case of bias, one unanticipated finding was the extent to which publication bias seemed to be an issue [38, 45, 61, 62, 65]. Unfortunately, failure to account for bias can result in overambitious claims, for example, that cost-effectiveness analyses for rare diseases can indeed meet traditional cost-effectiveness thresholds. In this example, most studies were industry funded, which made the authors speculate and wary of a potential publication bias [49]. Their sample of studies was not fully representative for economic evaluations of rare diseases because they mainly came from the literature, and if the hypothesis of publication bias is correct, there must be a higher likelihood that these studies were published, simply because they showed that cost-effective thresholds were reached.

Unfortunately, biased conclusions may disrupt ongoing efforts to improve reimbursement conditions for orphan drugs, and the momentum could be lost if policymakers take their conclusion at face value. The overall conclusion that cost-effectiveness analyses can meet common cost-effectiveness thresholds seems strongly contested by the findings of this review. In this example, the research would have been more convincing if the authors had considered cost-effectiveness analyses submitted to HTA bodies as compared to those available in the literature. We recommend further research to determine the effect of publication bias on the ability to meet cost-effectiveness thresholds and caution when interpreting results.

5.3 Other Opportunities

Researchers need to identify data gap years before economic evaluations to allow for sufficient time to generate the data needed. We have already described the potential for registries, but we recommend in addition to conduct an early economic evaluation of phase II data, which may provide timely knowledge on pricing and reimbursement [75]. Furthermore, patient organisations may be able to support reimbursement efforts, as there should be a mutual interest to bring orphan drugs to the market.

Another opportunity is risk-sharing agreements. Decision makers have implemented alternative methods of financing in response to high uncertainty for interventions, for example, future clinical and economic outcomes for orphan drugs [76, 77]. In short, they are in place to facilitate risk sharing between those supplying (manufacturers) and paying (healthcare providers) for health interventions, which is why they have broadly been referred to as risk sharing, pay-for-performance or managed-entry agreements. Although the nomenclature is not consistent, they can generally be divided into two categories: health outcome-based or non-outcome-based agreements [78, 79].

6 Limitations

Our review has some limitations. First, two researchers conducted the screening of titles and abstracts, but only one reviewer conducted the full screening and quality assessment. For this reason, the reliability could have been higher. To make up for this, we transparently report the full screening and quality assessment in the ESM. Second, exclusion of studies that did not qualify as systematic reviews meant that there was a chance of missing valuable information. Such an example was a narrative review of orphan drugs, which could have supported our findings [9]. Moreover, the search only included studies from 2010. However, the literature searches were partly based on search filters, which balanced sensitivity and specificity. Third, we included all studies, no matter their quality rating, to maximise inputs into the study. This resulted in the inclusion of one study with a low-quality rating [31]. Fourth, advanced therapeutic medicinal products (ATMPs) were excluded from this umbrella review, even if they were considered orphan drugs. It has been much debated whether they should qualify as drugs because the production process typically involves modifying cells or genes. There are challenges for economic evaluations of ATMPs such as high prices and sparse supportive evidence, for example, small sample sizes, single-arm studies and insufficient follow-up [80]. Thus, the identified opportunities for orphan drugs could apply equally to them. However, there are likely differences, ATMPs are frequently curative with a one-off cost, which is why major challenges are affordability and long-term uncertainty [81–83]. Furthermore, it was previously suggested to consider economic aspects for curative and non-curative treatments differently [57]. Finally, cross-referencing in the included papers was most prominent in recent papers, and in those with a broader scope. For example, a review concerned with methods for assessment of orphan drugs included six references, whereas another review of economic evaluations for enzyme replacement therapy in lysosomal storage disease included none [32, 46].

7 Conclusions

This umbrella review set out to determine issues for the economic evaluation of orphan drugs. The most obvious finding to emerge from this study was scarcity of evidence for clinical effectiveness, costs, quality of life and natural history of the disease. Scarcity of evidence and publication bias emerged as possible causes for the limited quantity of

economic evaluations from the literature. The results support the notion that an economic evaluation of rare diseases is challenging.

We recommend that researchers focus on sustainable initiatives and explore flexible cost-effectiveness models, for example, using living HTAs. We highlight that further research is required to determine the effect of publication bias on the ability to meet cost-effectiveness thresholds.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40273-024-01370-2.

Acknowledgements The authors acknowledge Anthea Tucker, librarian at the University of Sheffield, for valuable inputs on search strategies. The PROSPERO administration team confirmed that this review was not eligible for registration. Instead, the protocol for this umbrella review was made available in preprint before data extraction commenced [1]. It is available from University of Sheffield's online data repository. License: CC BY 4.0; https://doi.org/10.15131/shef.data.23390060; URL: Issues, challenges and opportunities for economic evaluation of orphan drugs: an umbrella review protocol (shef.ac.uk).

Declarations

Funding Tobias Sydendal Grand is an industrial PhD candidate at the University of Sheffield sponsored by Innovation Fund Denmark (case number: 2040-00017B) and Lundbeck A/S.

Conflicts of Interest/Competing Interests Tobias Sydendal Grand, Shijie Ren, James Hall, Daniel Oudin Åström, Stephane Regnier and Praveen Thokala have no conflicts of interest that are directly relevant to the content of this article.

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Availability of Data and Material All data generated or analysed during this study are included in this published article and its supplementary information files.

Code Availability The code used to create a PRISMA compliant flow diagram in Fig. 1 is available from: https://doi.org/https://doi.org/10.5281/zenodo.8232536. The code used to create an interactive version of Fig. 2 data extraction themes, sub-topics and findings is available from: https://tobiasgrand.github.io/-data-extraction-themes.github.io/ and files are available from: https://zenodo.org/doi/https://doi.org/10.5281/zenodo.10201217.

Authors' Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by TSG. JH was a second reviewer, hence assisted with the screening of papers. The first draft of the manuscript was written by TSG, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc/4.0/.

References

1. Grand T, Ren S, Thokala P, Oudin Åström D, Regnier S, Hall J. Issues, challenges and opportunities for economic evaluation of orphan drugs: an umbrella review protocol. 2023. The University of Sheffield's research data repository. 2023. https://doi.org/10.15131/shef.data.23390060.v1.
2. Richter T, Nestler-Parr S, Babela R, Khan ZM, Tesoro T, Molsen E, Hughes DA. Rare disease terminology and definitions: a systematic global review: report of the ISPOR Rare Disease Special Interest Group. Value Health. 2015;18(6):906–14. https://doi.org/10.1016/j.jval.2015.05.008.
3. The European Parliament. Regulation (EC) No 141/2000 of the European Parliament and of the Council of 16 December 1999 on orphan medicinal products. 2000. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32000R0141&from=EN. Accessed 20 Oct 2022.
4. Federal Food, Drug and Cosmetic Act. Orphan Drug Act. In: Senate and house of representatives of the United States of America in Congress, editor. 97th Congress: Public Law 97–114; 1983.
5. Thomas S, Caplan A. The orphan drug act revisited. JAMA. 2019;321(9):833–4. https://doi.org/10.1001/jama.2019.0290.
6. Miller KL, Lanthier M. Investigating the landscape of US orphan product approvals. Orphanet J Rare Dis. 2018;13(1):183. https://doi.org/10.1186/s13023-018-0930-3.
7. Pulsinelli GA. The Orphan Drug Act: what's right with It. Santa Clara High Technol Law J. 1999;15:2.
8. Miller KL, Fermaglich LJ, Maynard J. Using four decades of FDA orphan drug designations to describe trends in rare disease drug development: substantial growth seen in development of drugs for rare oncologic, neurologic, and pediatric-onset diseases. Orphanet J Rare Dis. 2021;16(1):265. https://doi.org/10.1186/s13023-021-01901-6.
9. Ollendorf DA, Chapman RH, Pearson SD. Evaluating and valuing drugs for rare conditions: no easy answers. Value Health. 2018;21(5):547–52. https://doi.org/10.1016/j.jval.2018.01.008.
10. Department of Health. The UK strategy for rare diseases. 2023. https://assets.publishing.service.gov.uk/media/5a7c61d540f0b626628abaaa/UK_Strategy_for_Rare_Diseases.pdf. Accessed 10 Oct 2023.
11. European Medicines Agency. Benefit-risk methodology. 2020. https://www.ema.europa.eu/en/about-us/what-we-do/regulatory-science-research/benefit-risk-methodology. Accessed 25 Oct 2022.
12. Pearson I, Rothwell B, Olaye A, Knight C. Economic modeling considerations for rare diseases. Value Health. 2018;21(5):515–24. https://doi.org/10.1016/j.jval.2018.02.008.
13. Nestler-Parr S, Korchagina D, Toumi M, Pashos CL, Blanchette C, Molsen E, et al. Challenges in research and health

technology assessment of rare disease technologies: report of the ISPOR Rare Disease Special Interest Group. Value Health. 2018;21(5):493–500. https://doi.org/10.1016/j.jval.2018.03.004.
14. York Health Economics Consortium. Health technology assessment. 2016. http://www.yhec.co.uk/glossary/health-technology-assessment/. Accessed 25 Oct 2022.
15. Dharssi S, Wong-Rieger D, Harold M, Terry S. Review of 11 national policies for rare diseases in the context of key patient needs. Orphanet J Rare Dis. 2017;12(1):63. https://doi.org/10.1186/s13023-017-0618-0.
16. Forman J, Taruscio D, Llera VA, Barrera LA, Coté TR, Edfjäll C, et al. The need for worldwide policy and action plans for rare diseases. Acta Paediatr. 2012;101(8):805–7. https://doi.org/10.1111/j.1651-2227.2012.02705.x.
17. National Institute for Health and Care Excellence. The Innovatives Medicines Fund principles. 2022. https://www.england.nhs.uk/wp-content/uploads/2022/06/B1686-the-innovate-medicines-fund-principles-june-2022.pdf. Accessed 27 Oct 2023.
18. Smith RA, Schneider PP, Mohammed W. Living HTA: automating health technology assessment with R. Wellcome Open Res. 2022;7:194. https://doi.org/10.12688/wellcomeopenres.17933.1.
19. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. Int J Evid Based Healthc. 2015;13(3):132–40. https://doi.org/10.1097/XEB.0000000000000055.
20. Salleh S, Thokala P, Brennan A, Hughes R, Booth A. Simulation modelling in healthcare: an umbrella review of systematic literature reviews. Pharmacoeconomics. 2017;35(9):937–49. https://doi.org/10.1007/s40273-017-0523-3.
21. Alsulamy N, Lee A, Thokala P, Alessa T. What influences the implementation of shared decision making: an umbrella review. Patient Educ Couns. 2020;103(12):2400–7. https://doi.org/10.1016/j.pec.2020.08.009.
22. De Freitas L, Goodacre S, O'Hara R, Thokala P, Hariharan S. Interventions to improve patient flow in emergency departments: an umbrella review. Emerg Med J. 2018;35(10):626–37. https://doi.org/10.1136/emermed-2017-207263.
23. Booth A. Clear and present questions: formulating questions for evidence based practice. Library Hi Tech. 2006;24(3):355–68. https://doi.org/10.1108/07378830610692127.
24. CADTH. Economic evaluations & models: MEDLINE. In: CADTH Search Filters Database. 2022. https://searchfilters.cadth.ca/link/16. Accessed 17 Nov 2022.
25. CADTH. Economic evaluations & models: Embase. In: CADTH Search Filters Database. 2022. https://searchfilters.cadth.ca/link/15. Accessed 23 Nov 2022.
26. SIGN. Systematic reviews. 2021. https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.sign.ac.uk%2Fassets%2Fsearch-filters-systematic-reviews.docx&wdOrigin=BROWSELINK. Accessed 10 Jan 2023.
27. Booth A, Sutton A, Papaioannou D. Searching the literature: systematic approaches to a successful literature review. New York: SAGE Publications; 2016. p. 110.
28. Joanna Briggs Institute. JBI critical appraisal checklist for systematic reviews and research syntheses. 2017. https://jbi.global/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Systematic_Reviews2017_0.pdf. Accessed 19 Jan 2024.
29. Sharif M, Sharif F, Ali H, Ahmed F. Systematic reviews explained: AMSTAR: how to tell the good from the bad and the ugly. Oral Health Dental Manage. 2013;12:9–16.
30. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol. 2007;7(1):10. https://doi.org/10.1186/1471-2288-7-10.
31. Lasalvia P, Prieto-Pinto L, Moreno M, Castrillon J, Romano G, Garzon-Orjuela N, Rosselli D. International experiences in multicriteria decision analysis (MCDA) for evaluating orphan drugs: a scoping review. Expert Rev Pharmacoecon Outcomes Res. 2019;19(4):40920. https://doi.org/10.1080/14737167.2019.1633918.
32. Mohammadshahi M, Olyaeemanesh A, Ehsani-Chimeh E, Mobinizadeh M, Fakoorfard Z, Akbari Sari A, Aghighi M. Methods and criteria for the assessment of orphan drugs: a scoping review. Int J Technol Assess Health Care. 2022;38(1):e59. https://doi.org/10.1017/S0266462322000393.
33. Abdallah K, Huys I, Claes K, Simoens S. Methodological quality assessment of budget impact analyses for orphan drugs: a systematic review. Front Pharmacol. 2021;2021:12. https://doi.org/10.3389/fphar.2021.630949.
34. Angelis A, Tordrup D, Kanavos P. Socio-economic burden of rare diseases: a systematic review of cost of illness evidence. Health Policy. 2015;119(7):964–79.
35. Babac A, Damm K, Graf von der Schulenburg JM. Patient-reported data informing early benefit assessment of rare diseases in Germany: a systematic review. Health Econ Rev. 2019;9(1):34. https://doi.org/10.1186/s13561-019-0251-9.
36. Baran-Kooiker A, Czech M, Kooiker C. Multi-criteria decision analysis (MCDA) models in health technology assessment of orphan drugs: a systematic literature review. Next steps in methodology development? Front Public Health. 2018;6:287. https://doi.org/10.3389/fpubh.2018.00287.
37. Chen T, Chen J, Chen C, Zheng H, Chen Y, Liu M, Zheng B. Systematic review and cost-effectiveness of bosentan and sildenafil as therapeutic drugs for pediatric pulmonary arterial hypertension. Pediatr Pulmonol. 2021;56(7):2250–8. https://doi.org/10.1002/ppul.25427.
38. Cheng MM, Ramsey SD, Devine EB, Garrison LP, Bresnahan BW, Veenstra DL. Systematic review of comparative effectiveness data for oncology orphan drugs. Am J Manage Care. 2012;18(1):47–62.
39. Dhawan A, Lawlor MW, Mazariegos GV, McKiernan P, Squires JE, Strauss KA, et al. Disease burden of Crigler-Najjar syndrome: systematic review and future perspectives. J Gastroenterol Hepatol. 2019;35(4):530–43. https://doi.org/10.1111/jgh.14853.
40. Faulkner E, Spinner DS, Ringo M, Carroll M. Are global health systems ready for transformative therapies? Value Health. 2019;22(6):627–41. https://doi.org/10.1016/j.jval.2019.04.1911.
41. Friedmann C, Levy P, Hensel P, Hiligsmann M. Using multi-criteria decision analysis to appraise orphan drugs: a systematic review. Expert Rev Pharmacoecon Outcomes Res. 2018;18(2):135–46. https://doi.org/10.1080/14737167.2018.1414603.
42. Garcia-Perez L, Linertova R, Valcarcel-Nazco C, Posada M, Gorostiza I, Serrano-Aguilar P. Cost-of-illness studies in rare diseases: a scoping review. Orphanet J Rare Dis. 2021;16(1):178. https://doi.org/10.1186/s13023-021-01815-3.
43. Gruhn S, Witte J, Greiner W, Damm O, Dietzsch M, Kramer R, Knuf M. Epidemiology and economic burden of meningococcal disease in Germany: a systematic review. Vaccine. 2022;40(13):1932–47. https://doi.org/10.1016/j.vaccine.2022.02.043.
44. Jakes RW, Kwon N, Nordstrom B, Goulding R, Fahrbach K, Tarpey J, Van Dyke MK. Burden of illness associated with eosinophilic granulomatosis with polyangiitis: a systematic literature review and meta-analysis. Clin Rheumatol. 2021;40(12):4829–36. https://doi.org/10.1007/s10067-021-05783-8.
45. Kanters TA, de Sonneville-Koedoot C, Redekop WK, Hakkaart L. Systematic review of available evidence on 11 high-priced inpatient orphan drugs. Orphanet J Rare Dis. 2013;8:124. https://doi.org/10.1186/1750-1172-8-124.

46. Katsigianni EI, Petrou P. A systematic review of economic evaluations of enzyme replacement therapy in Lysosomal storage diseases. Cost Eff Resour Alloc. 2022;20(1):51. https://doi.org/10.1186/s12962-022-00369-w.
47. Knoble N, Nayroles G, Cheng C, Arnould B. Illustration of patient-reported outcome challenges and solutions in rare diseases: a systematic review in Cushing's syndrome. Orphanet J Rare Dis. 2018;13(1):228. https://doi.org/10.1186/s13023-018-0958-4.
48. Kwon CS, Daniele P, Forsythe A, Ngai C. A systematic literature review of the epidemiology, health-related quality of life impact, and economic burden of immunoglobulin A nephropathy. J Health Econ Outcomes Res. 2021;8(2):36–45. https://doi.org/10.36469/001c.26129.
49. Picavet E, Cassiman D, Simoens S. What is known about the cost-effectiveness of orphan drugs? Evidence from cost-utility analyses. J Clin Pharm Ther. 2015;40(3):304–7. https://doi.org/10.1111/jcpt.12271.
50. Querol L, Crabtree M, Herepath M, Priedane E, Viejo Viejo I, Agush S, Sommerer P. Systematic literature review of burden of illness in chronic inflammatory demyelinating polyneuropathy (CIDP). J Neurol. 2021;268(10):3706–16. https://doi.org/10.1007/s00415-020-09998-8.
51. Raut M, Singh G, Hiscock I, Sharma S, Pilkhwal N. A systematic literature review of the epidemiology, quality of life, and economic burden, including disease pathways and treatment patterns of relapsed/refractory classical Hodgkin lymphoma. Expert Rev Hematol. 2022;15(7):607–17. https://doi.org/10.1080/17474086.2022.2080050.
52. Rice J, White A, Scarpati L, Philbin M, Wan G, Nelson W. Burden of noninfectious inflammatory eye diseases: a systematic literature review. J Manage Care Spec Pharm. 2018;23(3):S67. https://doi.org/10.1080/03007995.2018.1512961.
53. Schlander M, Dintsios CM, Gandjour A. Budgetary impact and cost drivers of drugs for rare and ultrarare diseases. Value Health. 2018;21(5):525–31. https://doi.org/10.1016/j.jval.2017.10.015.
54. Short H, Stafinski T, Menon D. A national approach to reimbursement decision-making on drugs for rare diseases in Canada? Insights from across the ponds. Healthc Policy. 2015;10(4):24–46.
55. Weidlich D, Kefalas P, Guest JF. Healthcare costs and outcomes of managing beta-thalassemia major over 50 years in the United Kingdom. Transfusion. 2016;56(5):1038–45. https://doi.org/10.1111/trf.13513.
56. Zelei T, Mendola ND, Elezbawy B, Nemeth B, Campbell JD. Criteria and scoring functions used in multi-criteria decision analysis and value frameworks for the assessment of rare disease therapies: a systematic literature review. Pharmacoecon Open. 2021;5(4):605–12. https://doi.org/10.1007/s41669-021-00271-w.
57. Zelei T, Molnar MJ, Szegedi M, Kalo Z. Systematic review on the evaluation criteria of orphan medicines in central and eastern European countries. Orphanet J Rare Dis. 2016;11(1):72. https://doi.org/10.1186/s13023-016-0455-6.
58. Lee C, Lam A, Kangappaden T, Olver P, Kane S, Tran D, Ammann E. Systematic literature review of evidence in amyloid light-chain amyloidosis. J Comp Eff Res. 2022;11(6):451–72. https://doi.org/10.2217/cer-2021-0261.
59. Leonart LP, Borba HHL, Ferreira VL, Riveros BS, Pontarolo R. Cost-effectiveness of acromegaly treatments: a systematic review. Pituitary. 2018;21(6):642–52. https://doi.org/10.1007/s11102-018-0908-0.
60. Onakpoya IJ, Spencer EA, Thompson MJ, Heneghan CJ. Effectiveness, safety and costs of orphan drugs: an evidence-based review. BMJ Open. 2015;5(6):e007199. https://doi.org/10.1136/bmjopen-2014-007199.
61. Park T, Griggs SK, Suh DC. Cost Effectiveness of monoclonal antibody therapy for rare diseases: a systematic review. BioDrugs. 2015;29(4):259–74. https://doi.org/10.1007/s40259-015-0135-4.
62. Schuller Y, Hollak CE, Biegstraaten M. The quality of economic evaluations of ultra-orphan drugs in Europe: a systematic review. Orphanet J Rare Dis. 2015;10:92. https://doi.org/10.1186/s13023-015-0305-y.
63. Sequeira AR, Mentzakis E, Archangelidi O, Paolucci F. The economic and health impact of rare diseases: a meta-analysis. Health Policy Technol. 2021;10(1):32–44. https://doi.org/10.1016/j.hlpt.2021.02.002.
64. Thorat T, Neumann PJ, Chambers JD. Hemophilia Burden of Disease: A systematic review of the cost-utility literature for hemophilia. J Manag Care Spec Pharm. 2018;24(7):632–42. https://doi.org/10.18553/jmcp.2018.24.7.632.
65. Woersching AL, Borrego ME, Raisch DW. Assessing the quality of economic evaluations of FDA novel drug approvals: a systematic review. Ann Pharmacother. 2016;50(12):1028–40. https://doi.org/10.1177/1060028016662893.
66. Plüddemann A, Banerjee A, O'Sullivan J. Positive results bias. 2017. https://catalogofbias.org/biases/positive-results-bias/. Accessed 19 Jun 2023.
67. Holman B, Bero L, Mintzes B. Industry sponsorship bias. 2019. https://catalogofbias.org/biases/industry-sponsorship-bias/. Accessed 24 Jun 2023.
68. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices–overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. Med Decis Making. 2012;32(5):667–77. https://doi.org/10.1177/0272989x12454577.
69. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. Value Health. 2012;15(6):835–42. https://doi.org/10.1016/j.jval.2012.04.014.
70. Mauskopf JA, Sullivan SD, Annemans L, Caro J, Mullins CD, Nuijten M, et al. Principles of good practice for budget impact analysis: report of the ISPOR Task Force on good research practices–budget impact analysis. Value Health. 2007;10(5):336–47. https://doi.org/10.1111/j.1524-4733.2007.00187.x.
71. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual 2022. 2022. https://www.nice.org.uk/process/pmg36/chapter/committee-recommendations. Accessed 11 Nov 2023.
72. Thokala P, Srivastava T, Smith R, Ren S, Whittington MD, Elvidge J, et al. Living health technology assessment: issues, challenges and opportunities. Pharmacoeconomics. 2023;41(3):227–37. https://doi.org/10.1007/s40273-022-01229-4.
73. Smith R, Schneider P. Making health economic models Shiny: a tutorial. Wellcome Open Res. 2020;5:69. https://doi.org/10.12688/wellcomeopenres.15807.2.
74. Kunst N, Burger EA, Coupé VMH, Kuntz KM, Aas E. A guide to an iterative approach to model-based decision making in health and medicine: an iterative decision-making framework. Pharmacoeconomics. 2024;42(4):363–71. https://doi.org/10.1007/s40273-023-01341-z.
75. Ramsey SD, Willke RJ, Glick H, Reed SD, Augustovski F, Jonsson B, et al. Cost-effectiveness analysis alongside clinical trials II: an ISPOR Good Research Practices Task Force Report. Value Health. 2015;18(2):161–72. https://doi.org/10.1016/j.jval.2015.02.001.
76. Drummond M. When do performance-based risk-sharing arrangements make sense? Eur J Health Econ. 2015;16(6):569–71. https://doi.org/10.1007/s10198-015-0683-z.
77. Facey KM, Espin J, Kent E, Link A, Nicod E, O'Leary A, et al. Implementing outcomes-based managed entry agreements for rare disease treatments: nusinersen and tisagenlecleucel.

Pharmacoeconomics. 2021;39(9):1021–44. https://doi.org/10.1007/s40273-021-01050-5.
78. Ferrario A, Kanavos P. Dealing with uncertainty and high prices of new medicines: a comparative analysis of the use of managed entry agreements in Belgium, England, the Netherlands and Sweden. Soc Sci Med. 2015;124:39–47. https://doi.org/10.1016/j.socscimed.2014.11.003.
79. Carlson JJ, Sullivan SD, Garrison LP, Neumann PJ, Veenstra DL. Linking payment to health outcomes: a taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. Health Policy. 2010;96(3):179–90. https://doi.org/10.1016/j.healthpol.2010.02.005.
80. Gladwell D, Ciani O, Parnaby A, Palmer S. Surrogacy and the valuation of ATMPs: taking our place in the evidence generation/assessment continuum. Pharmacoeconomics. 2024;42(2):137–44. https://doi.org/10.1007/s40273-023-01334-y.
81. Angelis A, Naci H, Hackshaw A. Recalibrating health technology assessment methods for cell and gene therapies. Pharmacoeconomics. 2020;38(12):1297–308. https://doi.org/10.1007/s40273-020-00956-w.
82. Coyle D, Durand-Zaleski I, Farrington J, Garrison L, Graf-von-der-Schulenburg JM, Greiner W, et al. HTA methodology and value frameworks for evaluation and policy making for cell and gene therapies. Eur J Health Econ. 2020;21(9):1421–37. https://doi.org/10.1007/s10198-020-01212-w.
83. Fiorenza S, Ritchie DS, Ramsey SD, Turtle CJ, Roth JA. Value and affordability of CAR T-cell therapy in the United States. Bone Marrow Transplant. 2020;55(9):1706–15. https://doi.org/10.1038/s41409-020-0956-8.

Authors and Affiliations

Tobias Sydendal Grand[1,3] · Shijie Ren[1] · James Hall[2] · Daniel Oudin Åström[3] · Stephane Regnier[3] · Praveen Thokala[1]

✉ Tobias Sydendal Grand
tsgrand1@sheffield.ac.uk

Shijie Ren
s.ren@sheffield.ac.uk

James Hall
j.a.hall@bham.ac.uk

Daniel Oudin Åström
AAST@lundbeck.com

Stephane Regnier
STRR@lundbeck.com

Praveen Thokala
p.thokala@sheffield.ac.uk

1 Sheffield Centre for Health and Related Research (SCHARR), University of Sheffield, 30 Regent Street, Sheffield S1 4DA, UK

2 Institute of Applied Health Research, Health Economics Unit, University of Birmingham, Edgbaston, Birmingham, UK

3 Lundbeck A/S, Copenhagen, Denmark

IV.II Cost-Effectiveness Modelling of Multiple System Atrophy for Early Health Technology Assessment

The formatting is not high quality, which is why I suggest readers visit the online version: https://link.springer.com/article/10.1007/s41669-025-00609-8.

PharmacoEconomics - Open
https://doi.org/10.1007/s41669-025-00609-8

ORIGINAL RESEARCH ARTICLE

Cost-Effectiveness Modelling of Multiple System Atrophy for Early Health Technology Assessment

Tobias Sydendal Grand[1,2] · Shijie Ren[1] · Praveen Thokala[1,7] · Stefano Zanigni[2] · Daniel Oudin Åström[2] · Aroussi Bidani[2] · Günter Höglinger[3,4,5] · Werner Poewe[6] · Florian Krismer[6] · Stephane Regnier[2] · the EMSA-SG Natural History Study Investigators

Received: 20 August 2025 / Accepted: 23 September 2025
© The Author(s) 2025

Abstract

Background and Objectives Multiple system atrophy is a neurodegenerative and rapidly progressing disease. Although symptoms often overlap, diagnostic criteria define two motor phenotypes: cerebellar and Parkinsonian. Little is known about available health-economic parameters for multiple system atrophy, and it remains unclear if the current available data are sufficient to construct an early health technology assessment. This assessment considered (i) if data gaps can be minimised or cost-effectiveness modelling can be facilitated by conducting literature searches for disease analogues, (ii) if an early health technology assessment is feasible with the currently available data, and (iii) the potential for developing user-friendly model interfaces in rare diseases.

Methods Literature searches were conducted for economic evaluations and health-economic parameters (including costs, utilities, natural history of disease) for multiple system atrophy. The searches for disease analogues focussed on economic evaluations and showed a widespread use of functional or disability scales to inform model structures; we therefore used the unified multiple system atrophy rating scale part IV (global disability scale) to inform the model structure. Natural history of disease data were used to inform utilities and transition probabilities. Analyses of a hypothetical intervention were conducted from the perspective of the United Kingdom National Health Service with a cost-effectiveness threshold of £30,000 per quality-adjusted life year. The cost-effectiveness analysis was presented using a user-friendly interface in R Shiny, which provides stakeholders with an opportunity to explore results.

Results Cost-effectiveness analyses were not identified for multiple system atrophy; however, two cost-of-illness, 20 quality-of-life, and nine natural history of disease studies were identified. Health-economic parameters were estimated from registry data, such as utilities and transition probabilities. A cost-effectiveness analysis was constructed for a hypothetical intervention despite evidence shortcomings, for example, aggregated cost data. An R Shiny app with a user-friendly interface was developed for stakeholders to change inputs and evaluate results.

Discussion Literature searches for disease analogues proved useful for early modelling in multiple system atrophy, for example, to inform model structures, but surrogate data could only be used when sufficient granularity was available (e.g., micro costing). These findings suggest that early conceptual modelling can help identify data gaps and guide future evidence generation, such as selecting appropriate endpoints for natural history of disease studies. Cost-effectiveness results can be highly uncertain for rare diseases at these early stages and should be considered as part of an iterative modelling framework.

1 Background

Multiple system atrophy (MSA) is a debilitating, devastating, and fatal disease with a prevalence of five cases per 100,000 people in the United Kingdom (UK), which is less than an estimated average country-specific threshold for a rare disease of 40 cases per 100,000 people, but prevalence differs by region, ranging from 0.5 (Spain) to 17 (Japan) cases per 100,000 people [1–3]. MSA is neurodegenerative and rapidly progressing with a reported survival of 6.2–9.8 years from onset of symptoms [4–7]. Although, symptoms can overlap, diagnostic criteria define two phenotypes: MSA-C (cerebellar) and MSA-P (Parkinsonian). MSA-C is associated with primarily cerebellar symptoms such as gait, limb ataxia, dysarthria, and cerebellar oculomotor disturbances. For MSA-P, the predominant symptoms include bradykinesia, rigidity, and gait difficulties. A poor response

Extended author information available on the last page of the article

Published online: 31 October 2025

Key Points for Decision Makers

This paper is the first to conduct a full evidence synthesis to inform an early cost-effectiveness analysis for multiple system atrophy.

Early conceptual modelling is recommendable as it provides important insights on data gaps and required evidence generation, but early cost-effectiveness results may be uncertain. Thus, an iterative modelling framework should be considered.

Analysts should consider rare disease analogues to inform modelling and to assess potential surrogate data.

to levodopa might be an early sign of MSA-P rather than Parkinson's disease [8–10].

The mean age at disease onset is 50–59 years and functional disease-deterioration states are aid-required walking, wheelchair confinement, bedridden, with a median event time of 3, 5, and 8 years, respectively [8, 9, 11]. An effective treatment that targets the underlying aetiology does not exist. Consequently, treatment aims at managing symptoms such as orthostatic hypotension, constipation, genitourinary, and breathing disorders, which require multiple medical specialities. Social services for end-of-life planning and caregiver support are also required [8, 12–14].

Demonstrating the value of orphan drugs in rare diseases is notoriously challenging as compared with more common diseases [15–17]. Grand and colleagues found that the main challenges were lack of evidence for clinical effectiveness, costs, quality of life, and natural history of disease data, and because health-economic input parameters are needed for economic evaluation, they may partly explain the scarcity of economic evaluations for rare diseases [18]. Recent research has suggested timely or early conceptual modelling and sourcing data from disease analogues as potential solutions that are attainable with modest resources and capabilities [16, 18–21].

There is a need for more research to evaluate the potential solutions for economic evaluations of rare diseases, but also to further corroborate challenges. Early health technology assessment (HTA) was regarded as an appropriate study type to evaluate methodological initiatives to improve economic evaluations for rare diseases, whilst acknowledging that these initiatives are unlikely to completely address all challenges. MSA is an excellent research candidate for this purpose due to its archetypal rare-disease attributes, for example, delayed diagnosis, misdiagnosis, multiple phenotypes, and low prevalence [8–10, 18, 22–28]. To our knowledge, no economic evaluation exists, little is known about available health-economic parameters for MSA, and it is not clear if current available data are sufficient to construct an early conceptual cost-effectiveness model for early HTA [29].

The aims of our assessment were as follows:

1. To explore if data gaps can be minimised or model development can be facilitated by conducting literature searches for disease analogues. The criteria to define a disease analogue were overlapping symptoms with MSA to increase applicability of the surrogate data despite not having the same underlying biology. Moreover, disease analogues should have higher prevalence than MSA, for efficiency purposes, because of an expected positive correlation between disease prevalence and evidence availability [19, 30].
2. To develop a conceptual cost-effectiveness model for early HTA comparing best supportive care (BSC) against BSC plus a hypothetical intervention, for example, physiotherapy, radiation, surgery, drug, or other treatments, and to use currently available evidence to highlight data gaps for cost-effectiveness modelling. Hypothetical scenarios in early HTA have previously facilitated early modelling in the absence of direct comparators [31].
3. To assess the potential for improving the user-friendliness of decision-analytic model interfaces, using MSA as a proxy for rare diseases.

2 Methods

2.1 Literature Searches

Literature searches were conducted from the 7th to 14th of May 2024 in EMBASE, MEDLINE, and the National Health Service (NHS) Economic Evaluation Database (EED). Search filters were used to identify cost-effectiveness analyses, cost analyses (direct and indirect costs), and quality-of-life analyses (health-state utilities) for MSA which yielded 519 papers, of which 111 were duplicates [32–35]. Overall, 408 papers were screened for inclusion. Hand searching was used to expand literature searches when critical data for input parameters were missing.

Four clinicians in neurology shortlisted suitable disease analogues in verbal and written dialogue for exploration that were progressive diseases of the central nervous system featuring symptoms overlapping with MSA, for example, with significant impact on functioning, quality of life, and survival (multiple sclerosis, advanced Parkinson's disease, and amyotrophic lateral sclerosis). A narrow search filter and a search limit of 5 years for economic evaluations were applied for disease analogues which provided 287 and 260 economic evaluations across MEDLINE and EMBASE [36]. Of these, 212 were duplicates, which left 335 papers for

potential exploration of supportive surrogate data. No studies were identified from the NHS EED databases for both MSA and disease analogues. Literature searches are available from the electronic supplementary material (ESM).

Cost-effectiveness analyses were not identified for MSA, but searches identified two cost papers [37, 38]. Despite an overlap, 20 unique MSA papers reported on quality of life ($n = 20$) or natural history of disease ($n = 9$) [6, 7, 11, 39–55]. By contrast, literature and hand searches for rare-disease analogues resulted in 47 economic evaluations: 42 for cost-effectiveness analyses across multiple sclerosis, three for advanced Parkinson's disease, and two for amyotrophic lateral sclerosis [56–102]. Markov models were primarily used in multiple sclerosis and health states were informed by an Expanded Disability Status Scale (EDSS) [77–85]. A proportion included on- and off-treatment periods or disease progression, such as remitting/relapsing or secondary progressive multiple sclerosis [56–76]. There were also examples of responder/non-responder, exacerbation, and severity (mild, moderate, and severe) health states [86–90]. For advanced Parkinson's disease, all models were Markovian and were informed by Hoehn and Yahr health stages, which is a measure of clinical disability, together with on- and off-states to reflect periods of either controlled or uncontrollable motor symptoms [98–100]. Two models were found for amyotrophic lateral sclerosis, one of which was based on a staging approach called Fine'til 9, which is a functional rating scale [101, 102].

2.2 Model Structure, Analysis and Perspective

The four clinicians' opinion and findings from the literature searches were evaluated in meetings with selected members from the author group (e.g., health economists and statisticians) to decide on the model structure, which was informed by the Unified Multiple System Atrophy Rating Scale (UMSARS) part IV, a disease-specific functional rating scale [103]. Items 1 and 2 were merged, as current available data indicated comparable utilities and a previous study reported aggregated costs [38]. The model structure has the following health states: completely or not completely independent (walking), more dependent (aid-required walking), very dependent (wheelchair confinement), totally dependent (bedridden), and death. In addition, transition arrows indicated that reversal of disease was not possible, which is aligned with the current lack of a cure. In Fig. 1, clinical milestones were illustrated alongside health states, first, to align with natural history of disease studies, and second, for easier disease comprehension by non-disease experts as compared with a full item description and clinical explanations [11, 44].

The modelling analysis plan (MAP) was endorsed by the author group and is available from the ESM or R Shiny app [104]. The modelling followed published tutorials and coding framework by Decision Analysis in R for Technologies in Health (DARTH) [105–107]. The study population were people living with MSA, regardless of whether clinically possible or probable diagnostic criteria applied [10]. The inclusion and exclusion criteria are aligned with the European MSA Study Group (EMSA-SG) work package III. The work package is data on file (not published); however, published literature describe the inclusion and exclusion criteria [6].

The comparator was a hypothetical intervention, which was a necessary but appropriate comparator in the absence of a disease-modifying treatment, and because such an intervention is likely to become an add-on to existing symptomatic treatment or BSC. The treatments considered in the economic evaluation are BSC versus BSC plus a hypothetical intervention, the latter with a modelled ability to delay disease progression and reduce clinical events.

A Markov model, or specifically, a time-dependent cohort state-transition model was developed in R software [108]. The mean age at diagnosis for MSA is 56 years, which was our cohort age at model entry [6]. The time horizon was 44 years and cycle length was 1 month, which resulted in 528 model cycles. The time horizon reflects that survival distributions are highly skewed in natural history of disease studies; for example, few subjects remain alive even at 12 or 25 years, depending on whether survival is estimated from probable diagnosis or symptom onset, respectively [6, 11]. In addition, the time horizon needed to capture the extended survival caused by a hypothetical intervention. The choice of cycle length was based on two criteria: first, the methods for estimating transition probabilities were flexible and the use of a monthly cycle length produced state occupancy plots with improved aesthetics. Second, observed transition probabilities in transition matrices were more intuitively aligned with the permitted health-state transitions that were illustrated in the model structure if cycle lengths were shorter. For example, longer cycle lengths result in implausible transitions being observed from multistate modelling, such as from *completely or not completely independent (Ci)* to *very dependent (Vd)*, but they simply reflect a sequence of unobserved transitions from *completely or not completely independent (Ci)* to *more dependent (Md)* to *very dependent (Vd)*.

The analysis used the UK NHS perspective. The cost-effectiveness threshold for the analysis was £30,000 per quality-adjusted life year (QALY), because £20,000–£30,000 is regarded as good value for money by the National Institute for Health and Care Excellence (NICE) [109]. The Highly Specialised Technology (HST) appraisal process was excluded as the MSA prevalence ranged from two to five cases per 100,000 people as compared with the HST routing criterion of lower than one case per 50,000 people [2,

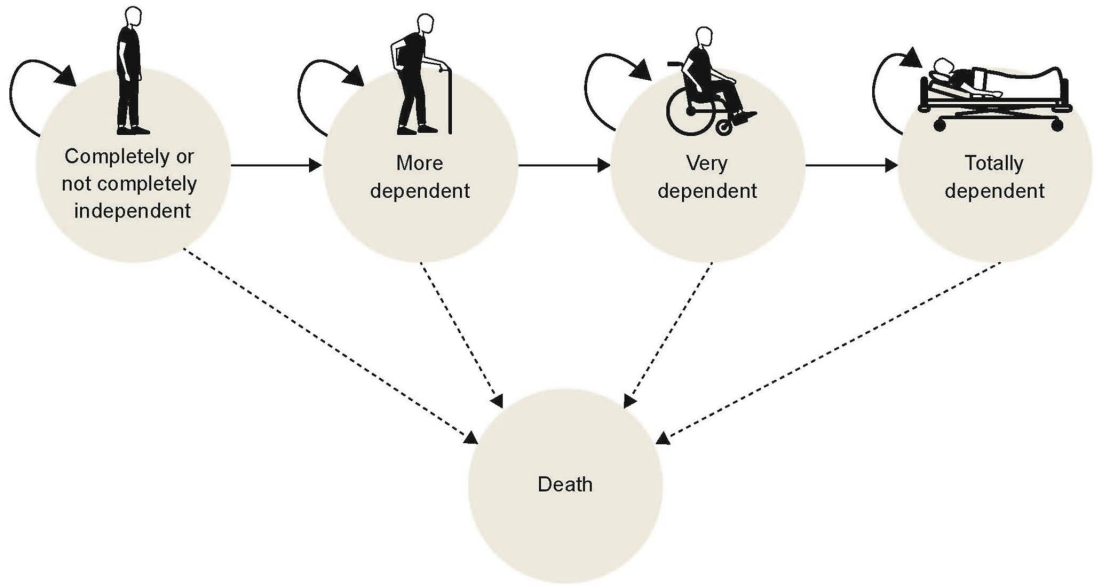


Fig. 1 Model structure for multiple system atrophy. The probabilities for urinary, bowel, and hypotension disorders are applied to health states. The first health state *completely or not completely independent* represents UMSARS part IV item 1 (completely independent) and item 2 (not completely independent). The second health state *more dependent* represents item 3 (more dependent), the third health state *very dependent* represents item 4 (very dependent), and fourth health state *totally dependent* represents item 5 (totally dependent). The fifth health state is *death* (absorbing state)

110–112]. Discount rates were 3.5% for costs and effects as per NICE guidelines [113].

2.3 Model Inputs

Derived health-economic parameters were sourced from data analysis of registry data. The data were European natural history of disease, for example, survival, UMSARS, and EuroQol-5Domains-3Levels (EQ-5D-3L) with a 2-year follow-up, which has been previously published by EMSA-SG [6].

2.3.1 Estimating Transition Probabilities

Transition probabilities for UMSARS part IV were calculated from patient-level data provided by the EMSA-SG, captured at baseline, 6, 12, 18, and 24 months. Exploration with non-parametric count methods from ESM showed variation in transition probabilities over 6-monthly intervals, possibly due to few subjects or insufficient follow-up. Consequently, multistate modelling was used, which relies on maximum likelihood estimation to endogenously estimate transition probabilities [114]. The msm R package was used. It requires an assumption for discrete time of a continuous process underlying the data. Practical requirements are data with observation time, observed state, and preferably subject id, and lastly, specification of a transition intensity matrix Q in accordance with the model structure [115].

In addition, Table 1 illustrates how a user-specified treatment modifier, *rr_trt_mod_p*, was multiplied with transition probabilities for the hypothetical intervention to reflect a disease mechanism that slows progression of disease, for example, monthly transition of *p_CiMd* * *rr_trt_mod_p*. In addition to disease-specific mortality (e.g., *p_CiD*), all-cause mortality *v_p_Dage* was introduced as a vector of probabilities to account for general population mortality. The use of a treatment modifier was inspired by a cost-effectiveness analysis in human immunodeficiency virus (HIV) [116, 117]. However, the treatment modification was not applied to the health state 'totally dependent', as prolonging a state with a negative utility that by definition is perceived worse than death could be considered unethical [118]. Our clinical experts challenged this definition and argued that this was rather a stopping rule than a state worse than death.

2.3.2 Estimating Utilities

Studies from literature searches that reported on quality of life all included UMSARS rating scales [6, 7, 11, 39–46]. The Composite Autonomic Symptom Score (COMPASS)

was used in two studies [7, 45]. Moreover, two studies included the MSA quality-of-life scale [44, 45]. One study focussed on sleep-related symptoms [52]. Importantly, for cost-effectiveness modelling, the Item Short Form Survey (SF-36) was included in four studies [7, 42, 43, 46]. Three studies reported mean EQ-5D index values with standard deviations in parenthesis, and they were between 0.2 (0.3) and 0.56 (0.28) [47, 49, 50]. Others reported on the distribution of MSA patients across EQ-5D dimensions [48–51]. The lack of stratification by model health states meant that data from literature was insufficient for estimation of utilities.

For this reason, mean utility index scores were calculated from the patient-level data provided by the EMSA-SG using the eq5d R package with UK EQ-5D-3L preference weights [119]. For simplicity, the utility value was kept constant during sensitivity analysis for the health state *totally dependent (Td)*. Moreover, disutilities were included in the analysis if they were considered to represent potentially treatable short-term (acute) clinical events and expected to result in a temporary or rapid reduction in quality of life, which might not be captured by 6-monthly EQ-5D-3L responses from patient-level data.

2.3.3 Other Parameters

2.3.3.1 Cost Data Only limited and antiquated cost data were available. The two studies claimed in their respective papers that no other cost-of-illness studies existed at their time of publication (2011) [37, 38]. McCrone and colleagues examined cost data from France, Germany, and the United Kingdom. People living with MSA required a range of services; informal care was the primary cost driver. Moreover, costs were correlated with disease severity [37]. Winter and colleagues investigated German patients and are the only ones to stratify costs by UMSARS part IV [103]. They showed that the significant cost drivers were age groups, marital status, age at disease onset, depression, and UMSARS part IV [38]. Subsequently, costs reported by UMSARS part IV were adjusted for inflation from 2011 to 2024, converted to British pounds, and assumed to be transferable to a UK setting. However, costs were not modelled for clinical events to avoid potential double-counting stemming from the use of aggregated costs despite a potential reduction in event-driven costs for the BSC and hypothetical intervention. The intervention is hypothetical, which is why the intervention costs were set at a £1000 per month, a completely arbitrary value. Users can modify the treatment costs in the Shiny app.

2.3.3.2 Clinical Events and Disutilities Five studies reported survival data by either using Kaplan-Meier curves or Cox-regression. The median survival from onset of symptoms to

Table 1 Transition matrix for hypothetical intervention with treatment modifier

	Completely or not completely independent (Ci)	More dependent (Md)	Very dependent (Vd)	Totally dependent (Td)	Death
Completely or not completely independent (Ci)	(1 − v_p_Dage[i]) * (1 - (p_CiMd + p_CiVd + p_CiTd + p_CiD) * rr_trt_mod_p)	(1 − v_p_Dage[i]) * p_CiMd * rr_trt_mod_p	(1 − v_p_Dage[i]) * p_CiVd * rr_trt_mod_p	(1 − v_p_Dage[i]) * p_CiTd * rr_trt_mod_p	(1 − v_p_Dage[i]) * p_CiD * rr_trt_mod_p + v_p_Dage[i]
More dependent (Md)	N/A	(1 − v_p_Dage[i]) * (1 − (p_MdVd + p_MdTd + p_MdD) * rr_trt_mod_p)	(1 − v_p_Dage[i]) * p_MdVd * rr_trt_mod_p	(1 − v_p_Dage[i]) * p_MdTd * rr_trt_mod_p	(1 − v_p_Dage[i]) * p_MdD * rr_trt_mod_p + v_p_Dage[i]
Very dependent (Vd)	N/A	N/A	(1 − v_p_Dage[i]) * (1 − (p_VdTd + p_VdD) * rr_trt_mod_p)	(1 − v_p_Dage[i]) * p_VdTd * rr_trt_mod_p	(1 − v_p_Dage[i]) * p_VdD * rr_trt_mod_p + v_p_Dage[i]
Totally dependent (Td)	N/A	N/A	N/A	(1 − v_p_Dage[i]) * (1 − p_TdD)	(1 − v_p_Dage[i]) * p_TdD + v_p_Dage[i]
Death (D)	N/A	N/A	N/A	N/A	p_DD = 1

N/A not applicable

*Explanatory notes: v_p_Dage (all-cause mortality vector for probabilities of death by age bands, rr_trt_mod_p (treatment modifier), and transition probabilities constructed by p = probability, from state = Ci, to state = Md resulting in p_CiMd

death were 6.4–9.8 years [6, 7, 11, 39, 40]. In addition, studies considered various disease aspects such as stridor, falls, orthostatic hypotension, depression, apathy, delusions, and urinary incontinence [39–44]. Time-to-event or probability-of-event analyses were reported for falls per day, unintelligible speech, and time to aid-required walking, wheelchair confinement, and bedridden state [6, 11].

The selection of clinical events was based on frequency of reporting, commonality of inclusion in economic evaluations for disease analogues, and clinical opinion. For example, urinary tract infections were commonly included in economic evaluations for rare-disease analogues [57, 62, 86, 87]. See ESM for additional details on clinical events. Clinical events were grouped by disorders and the most conservative event probabilities were sourced for hypotension (56.7%), urinary or bladder (51.1%), and bowel disorders (58.2%) [6]. A user-specified treatment modifier, *rr_trt_mod_e*, was multiplied with event probabilities for the hypothetical intervention, but similarly to the *rr_trt_mod_p*, it was not applied to the totally dependent health state. We assume probabilities remain constant over time.

The EQ-5D-5L value for MSA with and without orthostatic hypotension were 0.57 and 0.64, which were used to calculate a utility decrement of − 0.07 [50]. Similarly, a utility value (0.64) for functional intestinal disorders, such as functional diarrhoea and other constipation (ICD-10 Code K59.09 -K59.1), was subtracted from the UK population average (0.83), resulting in a disutility of − 0.19 [120]. A disutility value (− 0.01) for urinary tract disorder (ICD-9 599) was sourced from a UK EQ-5D catalogue [121].

2.4 Scenario Analyses

Table 2 includes a selection of relevant scenario analyses for exploration, which allows for a practical explanation of parameter interactions and reiteration of data limitations (e.g., for costs). However, readers are encouraged to explore the Shiny app by varying single or multiple parameters and construct their own scenarios. For example, in scenario 4, *rr_trt_mod_p* was set at the value 0.25 to modify transition probabilities for a hypothetical intervention as illustrated in Table 1. Likewise, *rr_trt_mod_e* with value 0.25 was multiplied with probabilities for clinical events, thus, reducing the overall disutility associated with these events. Lastly, costs were differentiated for the health states *very dependent (Vd)* at £3000 and *totally dependent (Td)* at £25,000 to circumvent one of the major data limitations, namely, that published costs are aggregated, which is moreover an issue for the *completely or not completely independent* state where costs were reduced to £0.

2.5 R Shiny App

A Shiny app was developed for the early HTA. The app consisted of three main coding components: code for a Markov model, time-dependency (three-dimensional array), and Shiny interface.

3 Results

3.1 Estimating Transition Probabilities

Table 3 contains estimated transition probabilities from multistate modelling. See ESM or R Shiny app (https://sheffield-university.shinyapps.io/Markov-Model-MSA/) for additional information, such as assigned distributions for probabilistic sensitivity analysis (PSA).

3.2 Estimating Utilities

Table 4 shows estimated utilities using the eq5d R package using UK EQ-5D-3L preference weights [119]. The health state utilities are correlated with disease severity. See ESM for additional information.

3.3 Other Parameters

3.3.1 Costs

Table 5 presents the estimated cost data, values are rounded to 2 decimals, after inflation adjustments and conversion from euro (€) to British pounds sterling (£).

Table 2 Scenario analyses for exploration

Scenario	Slowing of disease[a]	Reduce clinical events[b]	Monthly cost of Ci (£)	Monthly cost of Vd (£)	Monthly cost of Td (£)
Base case	0.5	0.5	988	3679	3679
1	0.7	0.7	988	3679	3679
2	0.5	0.5	988	3679	8000
3	0.45	0.45	100	2500	10,000
4	0.25	0.25	0	3000	25,000

Ci completely or not completely independent, *Td* totally dependent, *Vd* very dependent

[a]Note that slowing of disease is the value for rr_trt_mod_p applied to transition probabilities

[b]Reduce clinical events is the value for rr_trt_mod_e applied to event probabilities

The base-case settings correspond to those in the Shiny app

Table 3 Transition probabilities by multistate modelling

	Completely or not completely independent (Ci)	More dependent (Md)	Very dependent (Vd)	Totally dependent (Td)	Death (D)
Completely or not completely independent (Ci)	0.91[a]	0.09	0.00	0.00	0.00
More dependent (Md)	N/A	0.90	0.10	0.00	0.00
Very dependent (Vd)	N/A	N/A	0.95	0.03	0.02
Totally dependent (Td)	N/A	N/A	N/A	0.97	0.03
Death (D)	N/A	N/A	N/A	N/A	1

[a]Point estimate of transition probability. All values are rounded to 2 decimal points

Table 4 EQ-5D-3L (utilities) for UMSARS part IV health states

UMSARS part IV	Mean Index Score
Completely or not completely independent (Ci)	0.58[a]
More dependent (Md)	0.43
Very dependent (Vd)	0.14
Totally dependent (Td)	−0.13

UMSARS Unified Multiple System Atrophy Rating Scale (part IV: disability scale)

[a]Point estimate for mean EQ-5D-3L index score. All values are rounded to 2 decimals

3.4 Cost-Effectiveness Results for a Hypothetical Intervention

The deterministic cost-effectiveness analysis provided an incremental cost-effectiveness ratio (ICER) of £142,036 per QALY in Table 6. A breakdown of costs and QALYs by health states is included in the ESM.

The Markov trace in Fig. 2 illustrates the delay in disease progression from a hypothetical intervention and BSC compared with BSC alone.

The cost effectiveness in the PSA was £141,092 per QALY, which is slightly lower but close to the deterministic analysis, but higher QALYs and costs were generated as illustrated in Table 7.

The cost-effectiveness plane in Fig. 3 showed that the PSA simulations were above the cost-effectiveness threshold for the base-case settings.

The results of the scenario analyses are presented in Table 8. The ICERs varied, but as expected when considering each of the scenario specifications. Scenario 4 was within the cost-effectiveness threshold of £30,000 per QALY.

3.5 R Shiny App

A Shiny app was developed and can be accessed via this link: https://sheffield-university.shinyapps.io/Markov-Model-MSA/. Figure 4 shows an image of the user interface, which has a left side panel and upper navigation tabs; the panel allows the user to change model parameters and update the sensitivity analysis by pressing *'RUN MODEL'*. The navigation tabs contain model structure, deterministic analysis (fixed), sensitivity analyses (changeable), modelling analysis plan, introduction for non-technical users, other resources. The command *set.seed()* in the code, together

Table 6 Deterministic cost-effectiveness results for a hypothetical intervention

Deterministic analysis		Costs	QALYs	ICER
Comparators	Hypothetical intervention and BSC	£279,782	1.72	
	BSC	£147,651	0.79	142,036

BSC best supportive care, *ICER* incremental cost-effectiveness ratio, *QALY* quality-adjusted life year

Table 5 Converted monthly cost data

Converted costs for analysis (monthly)	Inpatient care	Drugs (BSC)[a]	Special equipment	Home care	Other direct costs	Total costs
Completely or not completely independent	£296.32	£197.55	£197.55		£296.32	£987.75
More dependent	£740.81	£296.32	£296.32	£49.39	£617.34	£2000.20
Very dependent	£987.75	321.02	£1037.14	£395.10	£938.36	£3679.37
Totally dependent	£987.75	£321.02	£1037.14	£395.10	£938.36	£3679.37

[a]Note that Drugs (BSC) is symptomatic treatment as part of best supportive care

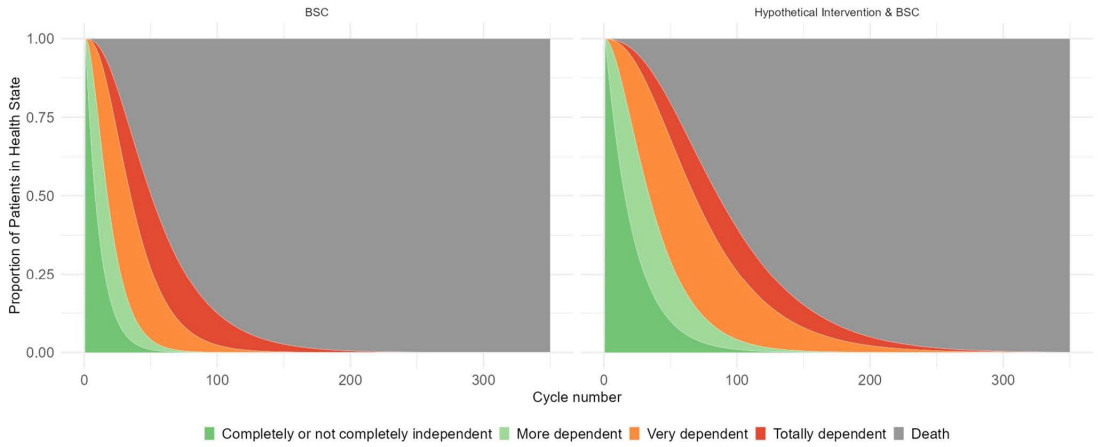


Fig. 2 State occupation by model cycles

Table 7 PSA results for a hypothetical intervention

PSA		Costs	QALYs	ICER
Comparators	Hypothetical Intervention and BSC	£304,234	1.81	
	BSC	£165,846	0.83	141,092

BSC best supportive care, *ICER* incremental cost-effectiveness ratio, *PSA* probabilistic sensitivity analysis, *QALY* quality-adjusted life year

with a download button for simulated PSA data, were implemented to facilitate reproduceable scenarios and, ultimately, to compare and scrutinise them.

4 Discussion

4.1 How Can Researchers Better Capture the Health Economic Value of MSA?

The literature searches illustrated a lack of cost-effectiveness analyses, and a scarcity of cost-of-illness or

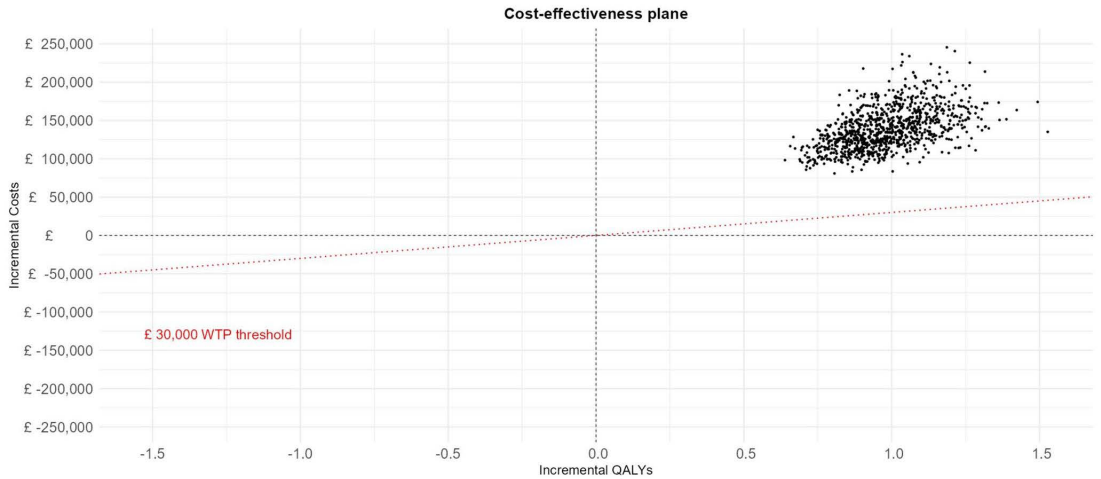


Fig. 3 Cost-effectiveness plane for probabilistic sensitivity analysis. *QALY* quality-adjusted life year, *WTP* willingness to pay

Table 8 Results from scenario analyses

Scenario		Costs	QALYs	ICER
1	Hypothetical intervention and BSC	£251,315	1.26	
	BSC	£165,846	0.83	194,865
2	Hypothetical intervention and BSC	£369,847	1.81	
	BSC	£242,458	0.83	129,827
3	Hypothetical intervention and BSC	£351,473	2.01	
	BSC	£244,508	0.83	90,683
4	Hypothetical intervention and BSC	£589,718	3.38	
	BSC	£519,271	0.83	27,646

BSC best supportive care, *ICER* incremental cost-effectiveness ratio, *QALY* quality-adjusted life year

burden-of-disease studies, which is not unusual for rare diseases such as MSA [16, 18]. Surprisingly, more research was available on quality of life where UMSARS was most often reported [6, 7, 11, 39–46]. Studies included SF-36, which allows for mapping to EQ-5D, or they directly reported EQ-5D index values for use in cost-effectiveness analyses [7, 42, 43, 46, 47, 49, 50, 122]. However, EQ-5D was reported for the whole disease, not stratified by scales or health states. Multiple studies reported on natural history of disease, such as survival [6, 7, 11, 39, 40]. A variety of disease characteristics were reported, such as stridor, falls, orthostatic hypotension, depression, apathy, delusions, and urinary incontinence [39–44]. Few studies reported time-to-event or probability-of-event analyses, which ought to have implications for early modelling, and they were falls per day, unintelligible speech, and time to aid-required walking, wheelchair confinement, and bedridden state [6, 11]. There seems to be little consensus on the inclusion of endpoints, which might be due to the multifaceted clinical representation of people living with MSA.

Establishing decision-analytic model structures for MSA could help guide the inclusion of health-economic endpoints, for example, in natural history of disease studies or clinical trials. This is important for cost and utility data, where recent data and sufficient stratification are missing for clinical events or milestones. Furthermore, reporting resource use rather than country-specific costs is recommendable for transferability. Lastly, research on the caregiver burden, such

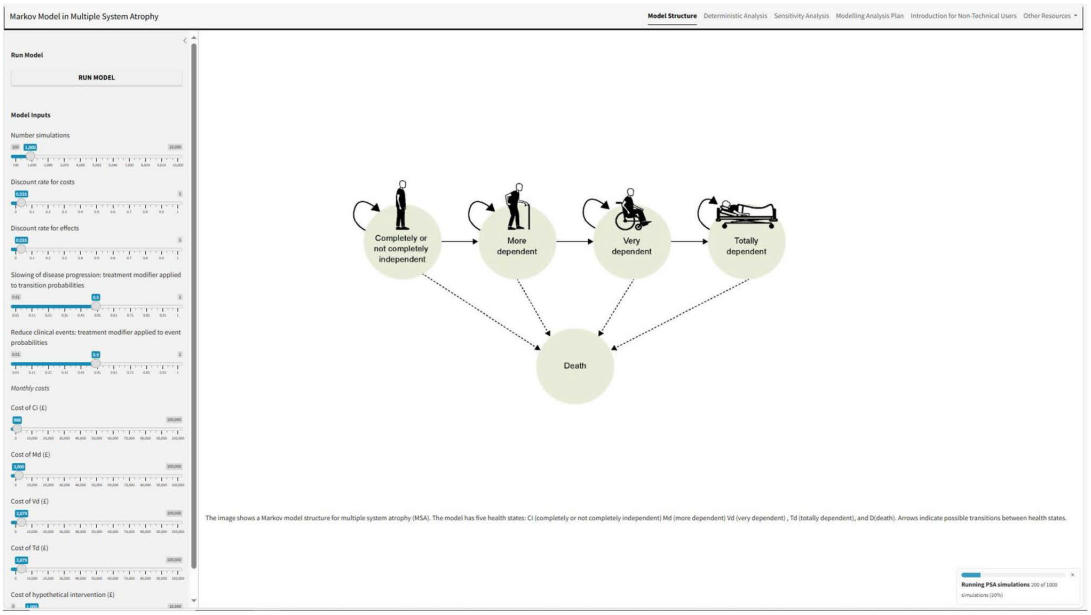


Fig. 4 Shiny app user interface

as associated indirect costs, was limited. Despite this, we found that the duration of care provided by caregivers was from 8 months to 8 years [53]. Caregivers provided 39 hours of practical care or 31 hours of emotional support per week if their patient had neurogenic orthostatic hypotension [54]. Moreover, 63% of caregivers to people living with MSA had depression [55]. A caregiver quote encapsulates the burden of informal care "I'm not going to ruin myself doing this, because if I ruin myself, I can't help him" [53].

4.2 Can Data Gaps Be Minimised, or Cost-Effectiveness Modelling Facilitated by Exploration of Rare Disease Analogues?

Cost-effectiveness analyses for disease analogues were plentiful [56–102]. The 5-year search limit was a minor limitation as a large proportion of cost-effectiveness analyses in advanced or late stage Parkinson's disease had an earlier publication peak [123]. This was a rather contradictory result in comparison with multiple sclerosis and it seemed to be caused by a different publication pattern, which indicated a continuum of publications rather than a peak as for Parkinson's disease [124, 125].

Health-economic parameters from cost-effectiveness analyses of disease analogues were often not transferable from papers that stratified by model health states. For example, for multiple sclerosis, costs and utilities were often stratified by an EDSS, from a variety of studies [56, 57, 59–73, 75–79, 81–83, 85, 86]. Similarly, utilities were stratified by Hoehn and Yahr and on- and off-states for Parkinson's disease [98, 99]. These insights could suggest that data were specifically generated for cost-effectiveness modelling purposes or that patient-reported outcomes were captured and, once collected, were used to inform health-economic models, thus further illustrating the benefit of early conceptual modelling to provide timely guidance for data generation. Useful analyses were those that referred to or provided thorough micro-costing approaches or reported relevant clinical events. For example, cost or disutility of urinary disorders, falling, depression, cardiovascular disorders, or walking aids, which are relevant events for MSA [5–7, 11, 39–41, 43, 44, 75, 87]. These findings further encourage the exploration of sourcing surrogate data from disease analogues, but there is a need for more research [16, 18].

However, certain modelling considerations could be carried over, such as assessing potential double counting if utility questionnaires were overlapping or if event-driven disutility increments were reported together with health state utilities [67, 77]. Moreover, there were examples where treatments were differentiated by their probability for various events, which supports the inclusion of a proposed event modifier (*rr_trt_mod_e*) in conceptual modelling [67, 68].

The recommendation to investigate disease analogues showed that the strategy was most useful for conceptualising model structures, for example, the overwhelming use of Markov models based on functional or disability rating scales such as Hoehn and Yahr or EDSS [18]. Similarly, for health-economic parameters, studies tended to be useful for inclusion or comparison if sufficient granularity was reported for costs and utilities, and when assessing applicability of frequent events such as urinary tract disorders or less frequent events such as house or car modifications [6, 7, 43, 44, 73, 87].

By contrast, more targeted searches for missing input parameters (for both rare and prevalent disease analogues) may be more time efficient in situations where model structure, including health-economic parameters, is already known [16, 18]. However, in all cases, seeking clinical advice is advisable to identify disease analogues and assess data eligibility for cost-effectiveness modelling.

4.3 Was Early HTA Feasible for MSA With Currently Available Evidence?

Multiple definitions exist for early HTA, but an amalgamation of them that aligns with this study is 'a systematic approach that syntheses data gaps to guide early evidence generation and to demonstrate an intervention's potential value for payers, patients, providers, and other healthcare stakeholders at an early stage of development' [29, 126].

4.3.1 Model Structure

Markov models constructed by functional or disability rating scales have seen widespread use in disease analogues, which can support the choice of UMSARS part IV to inform a model structure in MSA [56, 57, 59, 60, 62–85, 98–100, 102]. Hand searches for less prevalent disease analogues further corroborated these findings—for example, Huntington disease, where a total functional score was combined with active functional decline by Shoulson and Fahn categories, and spinocerebellar ataxia type 1, where early model health states were stratified by a scale for the assessment rating of ataxia [127, 128]. Moreover, it can conceivably be hypothesised that UMSARS part IV progression will correlate with increased costs and reduced quality of life, which is further indicated by published data and mean index scores derived from patient-level data in Table 4 [38]. What is more, McCrone and colleagues hold the view that their data support the hypothesis that disease-modifying interventions can decelerate or halt progression of disease, and hence reduce the economic burden [37].

However, there is abundant room for further progress in determining the most appropriate model structure for MSA. Which non-functional events such as depression, or

short-term events like hospitalisations due to falling should a model account for, and how, if they should be accounted for at all? Can or should stopping rules be applied, such as the ethical aspects of continued treatment in health states perceived worse than death [129]? What is the role of biomarkers in responder/non-responder analyses? Furthermore, aid-required walking, wheelchair confinement, and bedridden states have previously been reported as important motor function or clinical milestones [11]. Also, a noteworthy relationship exists between UMSARS part IV and these clinical milestones. For example, the fifth UMSARS part IV item states "totally dependent and helpless. Bedridden." By contrast, aid-required walking and wheelchair confinement are non-explicitly mentioned, but as previously highlighted, they have been illustrated alongside each other [44]. In theory, it should be possible to triangulate our results by partitioned survival modelling (time-to-event analyses), whereby patients move through health states based on parametric equations as opposed to transition probabilities for Markov models, for example, by digitising available Kaplan-Meier curves and reconstructing patient-level data on clinical progression milestones, but this would require *completely or not completely independent* to represent the state 'able to walk', despite being two separate items in UMSARS part IV [11, 130]. In this analysis, UMSARS IV items 1 (completely independent) and 2 (not completely independent) were grouped because cost and utility data indicated similar values, but this should be further investigated once more data becomes available as there might be an incremental difference in costs and effects that is not captured.

4.3.2 Early Cost-Effectiveness Modelling

The early cost-effectiveness modelling only had just enough data to populate all health-economic parameters. For example, the model was bound by available cost data which were aggregated for UMSARS part IV item 1 and 2, and 4 and 5. The Shiny app allows for variation in costs, but one could expand the model structure into all five UMSARS part IV items and generate state-specific costs. The natural history of disease data were crucial in estimating transition probabilities and utilities for health states. However, we encourage more research on estimation of transition probabilities in datasets with few subjects and short follow-up, as this is often the case for rare diseases. For example, our analyses showed sensitivity to the choice of methods, such as non-parametric count methods versus multistate modelling, which are available from the ESM [114].

Modelling results are highly uncertain at this early stage. However, model conceptualisation guidelines emphasise that models are a simplification of reality and that model specification is an iterative process, and future iterations may further add to a model as the understanding of the decision problem deepens [131, 132]. Model development should be independent from considerations on availability of data to inform health-economic parameters, which is why it is important not to solely appraise early economic evaluations of rare diseases based on data availability [132, 133]. We strongly oppose this practice because it is an inherent challenge and will hinder any effort to advance research on economic evaluations of rare diseases [18]. The use of a treatment modifier in early HTA could potentially be necessary to determine the efficacy required in clinical studies or for estimation of cost effectiveness with current evidence or efficacy projections, and subsequently inform go/no-go decisions in early pharmaceutical pipeline development, for example against country-specific HTA frameworks or cost-effectiveness thresholds [126]. It is useful when validating cost-effectiveness analyses during very early HTA to distinguish, according to guidelines on multi-application models, validation in the general sense (the MSA model itself) versus for a specific application (the effect of an intervention in MSA) [134]. The focus here was clearly on the model development, which from an early HTA perspective has value, whereas nothing was currently available to inform an intervention, which adds to the uncertainty of the modelling results. It has been stated in guidelines that sensitivity analyses are warranted for model parameters even if data do not exist (e.g. for an intervention) to assess a potential influence on results [132]. To summarise, our viewpoint on cost-effectiveness modelling in early HTA is that a treatment modifier has merit, and it is useful for determining efficacy levels required in clinical studies for go/no-go decisions, timely data generation, and for establishing model structures, despite uncertainty in results.

The PSA illustrated the uncertainty in data, while the scenario analyses illustrated the importance of, for example, generating state-specific cost data which had high influence on the results. However, cost effectiveness was achievable depending on user-specified input parameters. Moreover, by eliminating this uncertainty, one could potentially also start to see deterministic results that were cost effective. To conclude, we propose caution when interpreting results, but conceptual modelling has been useful in determining potential model structures and data gaps. It is advantageous to contextualise early conceptual modelling of MSA in an iterative modelling framework, because here we present only a humble beginning (first iteration) [131]. More research is needed on the most appropriate model structure and data gaps for MSA.

4.3.3 Frameworks for Early HTA

Practical frameworks vary for early HTA, but early-stage health-economic modelling is commonly included and has previously made use of hypothetical scenarios [29, 31].

Although the challenges for personalised medicines have been described, there is lack of focus on rare diseases [29, 126]. Interestingly, Wang and colleagues highlight the lack of a unified framework for early HTA as a limitation for wider use by manufacturers and innovators, which may have been partly induced by the inherent methodological affluency in an earlier definition as compared with the newer. For example, "*all methods used to…*" [29], is perhaps the reason for the suggested change to "*a systematic approach to*" [126], which ought to alleviate this issue by not encouraging the use of all methods.

4.4 Is It Worth Developing User-Friendly Interfaces for Early HTA in Rare Diseases?

Early HTA includes a broad range of stakeholders, which aligns well with our intentions for a Shiny app. Prof. Whitty wrote that "authors of modelling papers can do things to make their work substantially more useful for policy. The best is to provide an interactive interface, where if the policymaker does not agree with the starting assumptions of the model, they can change them" [135]. In our opinion, it is also one of several ways to promote transparency, validation, spark debate, and hopefully encourage others to progress a field of research such as a rare disease like MSA [134]. Perhaps Shiny apps, and R in particular, are becoming increasingly popular for these reasons, coupled with computational advantages and script-based coding resulting in less copy-paste errors as compared with spreadsheet-based modelling [21, 106, 136]. Nonetheless, tutorial papers on cost-effectiveness modelling in R have recently surged [21, 105, 136–141]. Interestingly, Hart and colleagues mention early cost-effectiveness modelling, PSA, changes to underlying data, and long run time if not conducted in R, as criteria that may help decide when R will lead to substantive improvements in efficiency and transparency [21]. We believe that this case study fulfilled most if not all criteria, and thus, early cost-effectiveness modelling in R outweighed the burden of programming and reviewing on a less familiar platform. To summarise, development of user-friendly interfaces is certainly possible with help from programming tutorials, even with modest R experience, and can be recommended for early conceptual cost-effectiveness modelling in rare diseases such as MSA [105, 107, 139, 142].

5 Limitations

A major limitation of this analysis was that cost data were only reported in two cost-of-illness studies by McCrone et al. and Winter et al., both dating back to 2011 [37, 38]. Winter and colleagues reported costs stratified by UMSARS part IV, but item 1 and 2, together with 4 and 5 were aggregated. The perspective was German, which required an assumption of generalisability to a UK setting [38]. In addition, a reduction in costs was not modelled despite a decrease in clinical events for the hypothetical intervention arm to avoid double-counting in relation to the use of aggregated cost data. The inclusion of clinical events was based on several factors, such as previous inclusion in cost-effectiveness analyses for disease analogues. However, there was no formal study, guidelines or established consensus that outlined the significance of disutility decrements and if events are sufficiently short to avoid double-counting, for example in relation to the use of 6-monthly EQ-5D questionnaires to inform health state utilities [57, 62, 86, 87]. The disutility values for urinary or bladder disorders and bowel disorders were estimated from EQ-5D catalogues [120, 121]. On the contrary, hypotension disorders (e.g., orthostatic hypotension) were estimated from an MSA population [50]. Lastly, the lack of disease-modifying interventions led to the use of treatment and event modifiers to assess the potential influence on cost-effectiveness results.

6 Conclusions

The literature searches for disease analogues were useful for providing information on model structures, which showed a widespread use of functional rating scales to inform Markov models. Although our analysis incorporated little surrogate data, it could potentially have been important had the natural history of disease data not been available to inform transition probabilities and utilities. It was feasible to conduct a first hypothetical cost-effectiveness analysis for MSA even though health-economic parameters were scarce. Early conceptual modelling is recommendable as it provides important insights on data gaps and evidence-generation needs. Moreover, it is recommendable to conduct iterative modelling, and our proposed model will almost certainly develop over time as new data emerge, such as expansion of health states, stopping rules, responder/non-responder analyses. R Shiny apps or user-friendly interfaces are preferable when conducting conceptual modelling for early HTA to facilitate inclusion of a broad range of stakeholders who might otherwise find it overly complicated to navigate the models. The skills required to code interfaces are attainable with published tutorials or R courses, but analysts should evaluate published criteria on potential pros and cons if they have less modelling experience in R.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s41669-025-00609-8.

Acknowledgements We would like to express our gratitude to all patients in the EMSA Natural History Study and their families, the Fifth Framework Programme of the European Community, and

Innovation Fund Denmark. In memory of Prof. Gregor Wenning [143]. We wish to pay our respects and acknowledge his invaluable contributions and expertise during the early study conceptualisation. In addition, we would like to thank Prof. Klaus Seppi for early advice and Geraldine Rylander for strategic and insightful comments on the draft manuscript.

Funding Tobias Sydendal Grand is an industrial PhD candidate at the University of Sheffield sponsored by Innovation Fund Denmark (case number: 2040-00017B) and Lundbeck A/S. The EMSA-SG Natural History Study was supported by funds from the fifth Framework Programme of the European Community (QLK6-CT-2000-0061).

Declarations

Conflict of interest The authors with non-Lundbeck affiliation have no competing interests to declare that are relevant to the content of this article. The following authors were affiliated with Lundbeck: Tobias Grand, Stefano Zanigni, Daniel Oudin Åström, Aroussi Bidani, and Stephane Regnier. Lundbeck A/S aspire to develop transformative treatments in neuroscience, such as for multiple system atrophy.

Availability of data and material Analysed data, except for patient-level data, from this study are included in this article or in supplementary information files.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Code availability The Shiny app is available from https://sheffield-university.shinyapps.io/Markov-Model-MSA/ and the R code from: https://github.com/TobiasGrand/CEA_MSA_ShinyApp. The data analysis report is available from https://tobiasgrand.github.io/CEA_MSA_DataAnalysis/.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Tobias Sydendal Grand. Praveen Thokala reviewed code for the Shiny app and double-coded the deterministic analysis in Excel. The first draft of the manuscript was written by Tobias Sydendal Grand, and all authors commented on previous versions of the modelling analysis plan and manuscript. All authors read and approved the final manuscript.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc/4.0/.

References

1. Richter T, Nestler-Parr S, Babela R, Khan ZM, Tesoro T, Molsen E, et al. Rare disease terminology and definitions—a systematic global review: report of the ISPOR rare disease special interest group. Value Health. 2015;18(6):906–14. https://doi.org/10.1016/j.jval.2015.05.008.
2. NHS Trust. Multiple system atrophy Leeds Centre for neurosciences; 2023. https://flipbooks.leedsth.nhs.uk/LN005500.pdf. Accessed 15 Jan 2025.
3. Kaplan S. Prevalence of multiple system atrophy: a literature review. Rev Neurol. 2024. https://doi.org/10.1016/j.neurol.2023.11.013.
4. Ben-Shlomo Y, Wenning GK, Tison F, Quinn NP. Survival of patients with pathologically proven multiple system atrophy. Neurology. 1997;48(2):384. https://doi.org/10.1212/WNL.48.2.384.
5. Wenning GK, Shlomo YB, Magalhães M, Danie SE, Quinn NP. Clinical features and natural history of multiple system atrophy: an analysis of 100 cases. Brain. 1994;117(4):835–45. https://doi.org/10.1093/brain/117.4.835.
6. Wenning GK, Geser F, Krismer F, Seppi K, Duerr S, Boesch S, et al. The natural history of multiple system atrophy: a prospective European cohort study. Lancet Neurol. 2013;12(3):264–74. https://doi.org/10.1016/s1474-4422(12)70327-7.
7. Low PA, Reich SG, Jankovic J, Shults CW, Stern MB, Novak P, et al. Natural history of multiple system atrophy in the USA: a prospective cohort study. Lancet Neurol. 2015;14(7):710–9. https://doi.org/10.1016/S1474-4422(15)00058-7.
8. Krismer F, Wenning GK. Multiple system atrophy: insights into a rare and debilitating movement disorder. Nat Rev Neurol. 2017;13(4):232–43. https://doi.org/10.1038/nrneurol.2017.26.
9. Mészáros L, Hoffmann A, Wihan J, Winkler J. Current symptomatic and disease-modifying treatments in multiple system atrophy. Int J Mol Sci. 2020. https://doi.org/10.3390/ijms21082775.
10. Gilman S, Wenning GK, Low PA, Brooks DJ, Mathias CJ, Trojanowski JQ, et al. Second consensus statement on the diagnosis of multiple system atrophy. Neurology. 2008;71(9):670. https://doi.org/10.1212/01.wnl.0000324625.00404.15.
11. Watanabe H, Saito Y, Terao S, Ando T, Kachi T, Mukai E, et al. Progression and prognosis in multiple system atrophy: an analysis of 230 Japanese patients. Brain. 2002;125(5):1070–83. https://doi.org/10.1093/brain/awf117.
12. Burns MR, McFarland NR. Current management and emerging therapies in multiple system atrophy. Neurotherapeutics. 2020;17(4):1582–602. https://doi.org/10.1007/s13311-020-00890-x.
13. Flabeau O, Meissner WG, Tison F. Multiple system atrophy: current and future approaches to management. Ther Adv Neurol Disord. 2010;3(4):249–63. https://doi.org/10.1177/1756285610375328.
14. Grossauer A, Sidoroff V, Heim B, Seppi K. Symptomatic care in multiple system atrophy: state of the art. Cerebellum. 2022. https://doi.org/10.1007/s12311-022-01411-6.
15. Ollendorf DA, Chapman RH, Pearson SD. Evaluating and valuing drugs for rare conditions: no easy answers. Value Health. 2018;21(5):547–52. https://doi.org/10.1016/j.jval.2018.01.008.
16. Pearson I, Rothwell B, Olaye A, Knight C. Economic modeling considerations for rare diseases. Value Health. 2018;21(5):515–24. https://doi.org/10.1016/j.jval.2018.02.008.
17. Nestler-Parr S, Korchagina D, Toumi M, Pashos CL, Blanchette C, Molsen E, et al. Challenges in research and health technology assessment of rare disease technologies: report of the ISPOR rare disease special interest group. Value Health. 2018;21(5):493–500. https://doi.org/10.1016/j.jval.2018.03.004.

18. Grand TS, Ren S, Hall J, Åström DO, Regnier S, Thokala P. Issues, challenges and opportunities for economic evaluations of orphan drugs in rare diseases: an umbrella review. Pharmacoeconomics. 2024. https://doi.org/10.1007/s40273-024-01370-2.
19. Grand T, Thokala P, Ren K, Regnier S, Oudin Åström D, Awad S, et al. Use of disease analogues to inform health-economic parameters for economic evaluations of rare diseases: a case-study of multiple system atrophy Barcelona, Spain. Value Health. 2024. https://www.ispor.org/heor-resources/presentations-database/presentation/euro2024-4014/144187. Accessed 7 July 2024.
20. Ramsey S, Willke R, Briggs A, Brown R, Buxton M, Chawla A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA task force report. Value Health. 2005;8(5):521–33. https://doi.org/10.1111/j.1524-4733.2005.00045.x.
21. Hart R, Burns D, Ramaekers B, Ren S, Gladwell D, Sullivan W, et al. R and shiny for cost-effectiveness analyses: why and when? A hypothetical case study. Pharmacoeconomics. 2020;38(7):765–76. https://doi.org/10.1007/s40273-020-00903-9.
22. Chrysostome V, Tison F, Yekhlef F, Sourgen C, Baldi I, Dartigues JF. Epidemiology of multiple system atrophy: a prevalence and pilot risk factor study in Aquitaine, France. Neuroepidemiology. 2004;23(4):201–8. https://doi.org/10.1159/000078506.
23. Bower JH, Maraganore DM, McDonnell SK, Rocca WA. Incidence of progressive supranuclear palsy and multiple system atrophy in Olmsted County, Minnesota, 1976 to 1990. Neurology. 1997;49(5):1284–8. https://doi.org/10.1212/WNL.49.5.1284.
24. Osaki Y, Morita Y, Kuwahara T, Miyano I, Doi Y. Prevalence of Parkinson's disease and atypical parkinsonian syndromes in a rural Japanese district. Acta Neurol Scand. 2011;124(3):182–7. https://doi.org/10.1111/j.1600-0404.2010.01442.x.
25. Sakushima K, Nishimoto N, Nojima M, Matsushima M, Yabe I, Sato N, et al. Epidemiology of multiple system atrophy in Hokkaido, the northernmost island of Japan. Cerebellum. 2015;14(6):682–7. https://doi.org/10.1007/s12311-015-0668-6.
26. Bjornsdottir A, Gudmundsson G, Blondal H, Olafsson E. Incidence and prevalence of multiple system atrophy: a nationwide study in Iceland. J Neurol Neurosurg Psychiatry. 2013;84(2):136–40. https://doi.org/10.1136/jnnp-2012-302500.
27. Chiò A, Magnani C, Schiffer D. Prevalence of Parkinson's disease in Northwestern Italy: comparison of tracer methodology and clinical ascertainment of cases. Mov Disord. 1998;13(3):400–5. https://doi.org/10.1002/mds.870130305.
28. Mazzucato M, Visonà Dalla Pozza L, Manea S, Minichiello C, Facchin P. A population-based registry as a source of health indicators for rare diseases: the ten-year experience of the Veneto Region's rare diseases registry. Orphanet J Rare Dis. 2014;9:37. https://doi.org/10.1186/1750-1172-9-37.
29. Ijzerman MJ, Koffijberg H, Fenwick E, Krahn M. Emerging use of early health technology assessment in medical product development: a scoping review of the literature. Pharmacoeconomics. 2017;35(7):727–40. https://doi.org/10.1007/s40273-017-0509-1.
30. Shourick J, Wack M, Jannot A-S. Assessing rare diseases prevalence using literature quantification. Orphanet J Rare Dis. 2021;16(1):139. https://doi.org/10.1186/s13023-020-01639-7.
31. Wang Y, Rattanavipapong W, Teerawattananon Y. Using health technology assessment to set priority, inform target product profiles, and design clinical study for health innovation. Technol Forecast Soc Change. 2021;172:121000. https://doi.org/10.1016/j.techfore.2021.121000.
32. CADTH. Economic Evaluations & Models—MEDLINE. In: CADTH Search Filters Database. 2022. https://searchfilters.cadth.ca/link/16. Accessed 17 Nov 2022.
33. CADTH. Economic Evaluations & Models—Embase. In: CADTH Search Filters Database. 2022. https://searchfilters.cadth.ca/link/15. Accessed 23 Nov 2022.
34. CADTH. Economic—Health Utilities/Quality of Life—Standard—Embase 2023. https://searchfilters.cda-amc.ca/link/18. Accessed 1 May 2024.
35. CADTH. Economic—Health Utilities/Quality of Life—MEDLINE 2023. https://searchfilters.cda-amc.ca/link/19. Accessed 1 May 2024.
36. CADTH. Economic—Narrow—MEDLINE, Embase 2024. https://searchfilters.cda-amc.ca/link/20. Accessed 10 May 2024.
37. McCrone P, Payan CA, Knapp M, Ludolph A, Agid Y, Leigh PN, et al. The economic costs of progressive supranuclear palsy and multiple system atrophy in France, Germany and the United Kingdom. PLoS ONE. 2011;6(9):e24369. https://doi.org/10.1371/journal.pone.0024369.
38. Winter Y, Stamelou M, Cabanel N, Sixel-Döring F, Eggert K, Höglinger GU, et al. Cost-of-illness in multiple system atrophy and progressive supranuclear palsy. J Neurol. 2011;258(10):1827–34. https://doi.org/10.1007/s00415-011-6028-7.
39. Cao B, Zhang L, Zou Y, Wei Q, Ou R, Chen Y, et al. Survival analysis and prognostic nomogram model for multiple system atrophy. Parkinsonism Relat Disord. 2018;54:68–73. https://doi.org/10.1016/j.parkreldis.2018.04.016.
40. Foubert-Samier A, Pavy-Le Traon A, Guillet F, Le-Goff M, Helmer C, Tison F, et al. Disease progression and prognostic factors in multiple system atrophy: a prospective cohort study. Neurobiol Dis. 2020;139:104813. https://doi.org/10.1016/j.nbd.2020.104813.
41. Jecmenica-Lukic M, Petrovic IN, Pekmezovic T, Tomic A, Stankovic I, Svetel M, et al. The profile and evolution of neuropsychiatric symptoms in multiple system atrophy: self-and caregiver report. J Neuropsychiatry Clin Neurosci. 2021;33(2):124–31. https://doi.org/10.1176/appi.neuropsych.20030057.
42. Jecmenica-Lukic MV, Pekmezovic TD, Petrovic IN, Dragasevic NT, Kostic VS. Factors associated with deterioration of health-related quality of life in multiple system atrophy: 1-year follow-up study. Acta Neurol Belg. 2018;118(4):589–95. https://doi.org/10.1007/s13760-018-0962-4.
43. Kollensperger M, Stampfer-Kountchev M, Seppi K, Geser F, Frick C, Del Sorbo F, et al. Progression of dysautonomia in multiple system atrophy: a prospective study of self-perceived impairment. Eur J Neurol. 2007;14(1):66–72. https://doi.org/10.1111/j.1468-1331.2006.01554.x.
44. Saulnier T, Fabbri M, Le Goff M, Helmer C, Pavy-Le Traon A, Meissner WG, et al. Patient-perceived progression in multiple system atrophy: natural history of quality of life. J Neurol Neurosurg Psychiatry. 2024. https://doi.org/10.1136/jnnp-2023-332733.
45. Golden EP, McCreary M, Vernino S. Responsiveness of UMSARS and other clinical measures in a longitudinal structured care clinic for multiple system atrophy. Clin Auton Res. 2022;32(6):477–84. https://doi.org/10.1007/s10286-022-00898-7.
46. May S, Gilman S, Sowell BB, Thomas RG, Stern MB, Colcher A, et al. Potential outcome measures and trial design issues for multiple system atrophy. Mov Disord. 2007;22(16):2371–7. https://doi.org/10.1002/mds.21734.
47. Schrag A, Geser F, Stampfer-Kountchev M, Seppi K, Sawires M, Kollensperger M, et al. Health-related quality of life in multiple system atrophy. Mov Disord. 2006;21(6):809–15. https://doi.org/10.1002/mds.20808.
48. Schrag A, Selai C, Mathias C, Low P, Hobart J, Brady N, et al. Measuring health-related quality of life in MSA: the MSA-qol.

Mov Disord. 2007;22(16):2332–8. https://doi.org/10.1002/mds.21649.
49. Schrag A, Sheikh S, Quinn NP, Lees AJ, Selai C, Mathias C, et al. A comparison of depression, anxiety, and health status in patients with progressive supranuclear palsy and multiple system atrophy. Mov Disord. 2010;25(8):1077–81. https://doi.org/10.1002/mds.22794.
50. Xiao Y, Zhang L, Wei Q, Ou R, Hou Y, Liu K, et al. Health-related quality of life in patients with multiple system atrophy using the EQ-5D-5L. Brain Behav. 2022;12(10):e2774. https://doi.org/10.1002/brb3.2774.
51. Winter Y, Spottke AE, Stamelou M, Cabanel N, Eggert K, Hoglinger GU, et al. Health-related quality of life in multiple system atrophy and progressive supranuclear palsy. Neurodegener Dis. 2011;8(6):438–46. https://doi.org/10.1159/000325829.
52. Lin J-Y, Zhang L-Y, Cao B, Wei Q-Q, Ou R-W, Hou Y-B, et al. Sleep-related symptoms in multiple system atrophy: determinants and impact on disease severity. Chin Med J. 2020;134(6):690–8. https://doi.org/10.1097/CM9.0000000000001211.
53. Langford B, Zhou Y, Miyasaki JM. Multiple system atrophy caregivers' experience: a mixed methods study. Can J Neurol Sci. 2023;50(1):49–59. https://doi.org/10.1017/cjn.2021.252.
54. Gallop K, Pham N, Maclaine G, Saunders E, Black B, Hubig L, et al. Health-related quality-of-life and burden for caregivers of individuals with neurogenic orthostatic hypotension. Neurodegener Dis Manag. 2023;13(1):35–45. https://doi.org/10.2217/nmt-2022-0015.
55. Miyashita M, Narita Y, Sakamoto A, Kawada N, Akiyama M, Kayama M, et al. Care burden and depression in caregivers caring for patients with intractable neurological diseases at home in Japan. J Neurol Sci. 2009;276(1–2):148–52. https://doi.org/10.1016/j.jns.2008.09.022.
56. Ayati N, Fleifel L, Sharifi S, Sahraian MA, Nikfar S. Cladribine tablets are a cost-effective strategy in high-disease activity relapsing multiple sclerosis patients in Iran. Curr J Neurol. 2021;20(3):146–53. https://doi.org/10.18502/cjn.v20i3.7690.
57. Ayati N, Taheri S, Sahraian MA, Nikfar S. Cost-effectiveness of ocrelizumab for treatment of Iranian patients with relapsing multiple sclerosis. Curr J Neurol. 2021;20(3):154–61. https://doi.org/10.18502/cjn.v20i3.7691.
58. Bolanos-Diaz R, Sanabria-Montanez C, Custodio-Capunay N. Cost-effectiveness of interferon therapy for multiple sclerosis in Peru: impact of treatment adherence. J Pharm Health Serv Res. 2019;10(1):125–32. https://doi.org/10.1111/jphs.12279.
59. Furneri G, Santoni L, Ricella C, Prosperini L. Cost-effectiveness analysis of escalating to natalizumab or switching among immunomodulators in relapsing-remitting multiple sclerosis in Italy. BMC Health Serv Res. 2019;19(1):436. https://doi.org/10.1186/s12913-019-4264-1.
60. Kremer IEH, Hiligsmann M, Carlson J, Zimmermann M, Jongen PJ, Evers SMAA, et al. Exploring the cost effectiveness of shared decision making for choosing between disease-modifying drugs for relapsing-remitting multiple sclerosis in the Netherlands: a state transition model. Med Decis Mak. 2020;40(8):1003–19. https://doi.org/10.1177/0272989X20961091.
61. Lazzaro C, Bergamaschi R, Zaffaroni M, Totaro R, Paolicelli D. Cost-utility analysis of teriflunomide in naive vs. previously treated patients with relapsing-remitting multiple sclerosis in Italy. Neurol Sci. 2022;43(8):4933–44. https://doi.org/10.1007/s10072-022-06022-x.
62. Martins P, Vandewalle B, Felix J, Capela CM, Cerqueira JJ, Salgado AV, et al. Cost-effectiveness analysis of ocrelizumab for the treatment of relapsing and primary progressive multiple sclerosis in Portugal. PharmacoEcon Open. 2023;7(2):229–41. https://doi.org/10.1007/s41669-022-00381-z.
63. Michels RE, de Fransesco M, Mahajan K, Hengstman GJD, Schiffers KMH, Budhia S, et al. Cost effectiveness of Cladribine tablets for the treatment of relapsing-remitting multiple sclerosis in The Netherlands. Appl Health Econ Health Policy. 2019;17(6):857–73. https://doi.org/10.1007/s40258-019-00500-8.
64. Montgomery S, Woodhouse F, Vudumula U, Gudala K, Duddy M, Kroes M. Stick or twist? Cost-effectiveness of siponimod compared with continuing existing disease-modifying therapies in the treatment of active secondary progressive multiple sclerosis in the UK. J Med Econ. 2022;25(1):669–78. https://doi.org/10.1080/13696998.2022.2078103.
65. Nakhaipour HR, Vudumula U, Khurana V, Sebire G, Mah JK, Pohl D, et al. Cost-effectiveness of fingolimod versus interferon-beta1a for the treatment of pediatric-onset multiple sclerosis in Canada. J Med Econ. 2020;23(12):1525–33. https://doi.org/10.1080/13696998.2020.1840138.
66. Navarro CE, Betancur JE. Cost-utility analysis comparing ocrelizumab versus rituximab in the treatment of relapsing-remitting multiple sclerosis: the Colombian perspective. Value Health Reg Issues. 2023;36:83–91. https://doi.org/10.1016/j.vhri.2023.02.008.
67. Pinheiro B, Guerreiro R, Costa J, Miguel LS. Cost-effectiveness of cladribine tablets versus fingolimod in patients with highly active relapsing multiple sclerosis in Portugal. J Med Econ. 2020;23(5):484–91. https://doi.org/10.1080/13696998.2020.1717499.
68. Poveda JL, Trillo JL, Rubio-Terres C, Rubio-Rodriguez D, Polanco A, Torres C. Cost-effectiveness of cladribine tablets and fingolimod in the treatment of relapsing multiple sclerosis with high disease activity in Spain. Expert Rev Pharmacoecon Outcomes Res. 2020;20(3):295–303. https://doi.org/10.1080/14737167.2019.1635014.
69. Schur N, Gudala K, Vudumula U, Vadapalle S, Bhadhuri A, Casanova A, et al. Cost effectiveness and budget impact of siponimod compared to interferon beta-1a in the treatment of adult patients with secondary progressive multiple sclerosis with active disease in Switzerland. Pharmacoeconomics. 2021;39(5):563–77. https://doi.org/10.1007/s40273-021-01023-8.
70. Spelman T, Herring WL, Acosta C, Hyde R, Jokubaitis VG, Pucci E, et al. Comparative effectiveness and cost-effectiveness of natalizumab and fingolimod in rapidly evolving severe relapsing-remitting multiple sclerosis in the United Kingdom. J Med Econ. 2024;27(1):109–25. https://doi.org/10.1080/13696998.2023.2293379.
71. Spelman T, Herring WL, Zhang Y, Tempest M, Pearson I, Freudensprung U, et al. Comparative effectiveness and cost-effectiveness of natalizumab and fingolimod in patients with inadequate response to disease-modifying therapies in relapsing-remitting multiple sclerosis in the United Kingdom. Pharmacoeconomics. 2022;40(3):323–39. https://doi.org/10.1007/s40273-021-01106-6.
72. Stanisic S, Bertolotto A, Berto P, Di Procolo P, Morawski J. The cost-effectiveness of alemtuzumab in the management of relapse-remitting multiple sclerosis in Italy. Glob Reg Health Technol Assess. 2019. https://doi.org/10.1177/2284240319838524.
73. Taheri S, Sahraian MA, Yousefi N. Cost-effectiveness of alemtuzumab and natalizumab for relapsing-remitting multiple sclerosis treatment in Iran: decision analysis based on an indirect comparison. J Med Econ. 2019;22(1):71–84. https://doi.org/10.1080/13696998.2018.1543189.
74. Visser LA, Folcher M, Delgado Simao C, Gutierrez Arechederra B, Escudero E, Uyl-de Groot CA, et al. The potential cost-effectiveness of a cell-based bioelectronic implantable device delivering Interferon-beta1a therapy versus injectable Interferon-beta1a treatment in relapsing-remitting multiple sclerosis.

Pharmacoeconomics. 2022;40(1):91–108. https://doi.org/10.1007/s40273-021-01081-y.
75. Walter E, Berger T, Bajer-Kornek B, Deisenhammer F. Cost-utility analysis of alemtuzumab in comparison with interferon beta, fingolimod, and natalizumab treatment for relapsing-remitting multiple sclerosis in Austria. J Med Econ. 2019;22(3):226–37. https://doi.org/10.1080/13696998.2018.1556668.
76. Xu Y, Mao N, Chirikov V, Du F, Yeh Y-C, Liu L, et al. Cost-effectiveness of teriflunomide compared to interferon beta-1b for relapsing multiple sclerosis patients in China. Clin Drug Investig. 2019;39(3):331–40. https://doi.org/10.1007/s40261-019-00750-3.
77. Auguste P, Colquitt J, Connock M, Loveman E, Court R, Ciccarelli O, et al. Ocrelizumab for treating patients with primary progressive multiple sclerosis: an evidence review group perspective of a NICE single technology appraisal. Pharmacoeconomics. 2020;38(6):527–36. https://doi.org/10.1007/s40273-020-00889-4.
78. Baharnoori M, Bhan V, Clift F, Thomas K, Mouallif S, Adlard N, et al. Cost-effectiveness analysis of Ofatumumab for the treatment of relapsing-remitting multiple sclerosis in Canada. PharmacoEcon Open. 2022;6(6):859–70. https://doi.org/10.1007/s41669-022-00363-1.
79. Cortesi PA, Antonazzo IC, Gasperini C, Nica M, Ritrovato D, Mantovani LG. Cost-effectiveness and budget impact analysis of siponimod in the treatment of secondary progressive multiple sclerosis in Italy. PLoS ONE. 2022;17(3 March):e0264123. https://doi.org/10.1371/journal.pone.0264123.
80. Daroudi R, Mousavi M, Shirazikhah M, Alizadeh Zarei M, Hendi H, Joghataei F, et al. Cost-utility analysis of multiple sclerosis rehabilitation in Iran. Expert Rev Pharmacoecon Outcomes Res. 2023;23(10):1129–37. https://doi.org/10.1080/14737167.2023.2251689.
81. Ginestal R, Rubio-Terres C, Moran OD, Rubio-Rodriguez D, De Los Santos H, Ordonez C, et al. Cost-effectiveness of cladribine tablets and dimethyl fumarate in the treatment of relapsing remitting multiple sclerosis in Spain. J Comp Eff Res. 2023;12(2):e220193. https://doi.org/10.2217/cer-2022-0193.
82. Giovannoni G, Brex PA, Dhiraj D, Fullarton J, Freddi M, Rodgers-Gray B, et al. Glatiramer acetate as a clinically and cost-effective treatment of relapsing multiple sclerosis over 10 years of use within the National Health Service: final results from the UK risk sharing scheme. Mult Scler J Exp Transl Clin. 2019. https://doi.org/10.1177/2055217319893103.
83. Hua LH, Hersh CM, Morten P, Kusel J, Lin F, Cave J, et al. The impact of price reductions after loss of exclusivity in a cost-effectiveness analysis: fingolimod versus interferon beta-1a for the treatment of relapsing-remitting multiple sclerosis. J Manag Care Spec Pharm. 2019;25(4):490. https://doi.org/10.18553/jmcp.2019.25.4.490.
84. Lasalvia P, Hernandez F, Castaneda-Cardona C, Cuestas JA, Rosselli D. Cost-effectiveness of Natalizumab compared with Fingolimod for relapsing-remitting multiple sclerosis treatment in Colombia. Value Health Reg Issues. 2020;23:13–8. https://doi.org/10.1016/j.vhri.2019.08.481.
85. Rezaee M, Izadi S, Keshavarz K, Borhanihaghighi A, Ravangard R. Fingolimod versus natalizumab in patients with relapsing remitting multiple sclerosis: a cost-effectiveness and cost-utility study in Iran. J Med Econ. 2019;22(4):297–305. https://doi.org/10.1080/13696998.2018.1560750.
86. Acosta C, Gianinazzi M, Dort T, Armstrong N, Ryder S, Lundqvist T, et al. Modeling the cost-effectiveness of prolonged-release fampridine for the treatment of walking impairment in patients with multiple sclerosis in Sweden. J Med Econ. 2021;24(1):770–80. https://doi.org/10.1080/13696998.2021.1927746.
87. Hunter SF, Bindra J, Chopra I, Niewoehner J, Panaccio MP, Wan GJ. Cost-effectiveness of repository corticotropin injection for the treatment of acute exacerbations in multiple sclerosis. Clinicoecon Outcomes Res. 2021;13:883–92. https://doi.org/10.2147/CEOR.S330118.
88. Oppe M, Ortin-Sulbaran D, Vila Silvan C, Estevez-Carrillo A, Ramos-Goni JM. Cost-effectiveness of adding Sativex spray to spasticity care in Belgium: using bootstrapping instead of Monte Carlo simulation for probabilistic sensitivity analyses. Eur J Health Econ. 2021;22(5):711–21. https://doi.org/10.1007/s10198-021-01285-1.
89. Zhao X, Yang H, Wei T, Zhao J, Liu J, Huang Z, et al. Cost-effectiveness analysis of prolonged-release fampridine to treat walking disability of multiple sclerosis in China. J Comp Eff Res. 2022;11(14):1057–69. https://doi.org/10.2217/cer-2022-0109.
90. Palmer AJ, Zhao T, Taylor BV, van der Mei I, Campbell JA. Exploring the cost-effectiveness of EBV vaccination to prevent multiple sclerosis in an Australian setting. J Neurol Neurosurg Psychiatry. 2024;95(5):401–9. https://doi.org/10.1136/jnnp-2023-332161.
91. Dimitrova M, Seitaridou Y, Lazarova R, Petrova G, Mitov K, Milanov I, et al. Cost-effectiveness of disease-modifying treatments for multiple sclerosis in Bulgaria based on evidence from real world settings. Farmacia. 2023;71(6):1324–32. https://doi.org/10.31925/farmacia.2023.6.24.
92. Mantovani LG, Cozzolino P, Cortesi PA, Patti F. Cost-effectiveness analysis of cannabinoid oromucosal spray use for the management of spasticity in subjects with multiple sclerosis. Clin Drug Investig. 2020;40(4):319–26. https://doi.org/10.1007/s40261-020-00895-6.
93. Michel M, Le Page E, Laplaud DA, Wardi R, Lebrun C, Zagnoli F, et al. Cost-utility of oral methylprednisolone in the treatment of multiple sclerosis relapses: results from the COPOUSEP trial. Rev Neurol. 2022;178(3):241–8. https://doi.org/10.1016/j.neurol.2021.06.009.
94. Renfrew LM, Paul L, McFadyen A, Rafferty D, Moseley O, Lord AC, et al. The clinical- and cost-effectiveness of functional electrical stimulation and ankle-foot orthoses for foot drop in multiple sclerosis: a multicentre randomized trial. Clin Rehabil. 2019;33(7):1150–62. https://doi.org/10.1177/0269215519842254.
95. Rosato R, Pagano E, Giordano A, Farinotti M, Ponzio M, Veronese S, et al. Living with severe multiple sclerosis: cost-effectiveness of a palliative care intervention and cost of illness study. Mult Scler Relat Disord. 2021;49:102756. https://doi.org/10.1016/j.msard.2021.102756.
96. Huygens S, Versteegh M. Modeling the cost-utility of treatment sequences for multiple sclerosis. Value Health. 2021;24(11):1612–9. https://doi.org/10.1016/j.jval.2021.05.020.
97. Versteegh MM, Huygens SA, Wokke BWH, Smolders J. Effectiveness and cost-effectiveness of 360 disease-modifying treatment escalation sequences in multiple sclerosis. Value Health. 2022;25(6):984–91. https://doi.org/10.1016/j.jval.2021.11.1363.
98. Chaudhuri KR, Pickard AS, Alobaidi A, Jalundhwala YJ, Kandukuri PL, Bao Y, et al. The cost effectiveness of levodopa-carbidopa intestinal gel in the treatment of advanced Parkinson's disease in England. Pharmacoeconomics. 2022;40(5):559–74. https://doi.org/10.1007/s40273-022-01132-y.
99. Kalabina S, Belsey J, Pivonka D, Mohamed B, Thomas C, Paterson B. Cost-utility analysis of levodopa carbidopa intestinal gel (Duodopa) in the treatment of advanced Parkinson's disease in patients in Scotland and Wales. J Med Econ. 2019;22(3):215–25. https://doi.org/10.1080/13696998.2018.1553179.

100. Pahwa R, Merola A, Soileau M, Alobaidi A, Pickard AS, Kandukuri PL, et al. Cost-effectiveness of Carbidopa-Levodopa enteral suspension for advanced Parkinson's disease in the United States. Mov Disord. 2023;38(12):2308–12. https://doi.org/10.1002/mds.29624.
101. Gajdos O, Rozanek M, Donin G, Kamensky V. Cost-Utility Analysis of Home Mechanical Ventilation in Patients with Amyotrophic Lateral Sclerosis. Healthcare (Basel, Switzerland). 2021. https://doi.org/10.3390/healthcare9020142.
102. Thakore NJ, Pioro EP, Udeh BL, Lapin BR, Katzan IL. A cost-effectiveness framework for amyotrophic lateral sclerosis, applied to Riluzole. Value Health. 2020;23(12):1543–51. https://doi.org/10.1016/j.jval.2020.06.012.
103. Wenning GK, Tison F, Seppi K, Sampaio C, Diem A, Yekhlef F, et al. Development and validation of the unified multiple system atrophy rating scale (UMSARS). Mov Disord. 2004;19(12):1391–402. https://doi.org/10.1002/mds.20255.
104. Chang W, Cheng J, Allaire J, Sievert C, Schloerke B, Xie Y, et al. shiny: Web Application Framework for R 2023. https://shiny.rstudio.com/. Accessed 1 Mar 2023.
105. Smith R, Schneider P. Making health economic models shiny: a tutorial. Wellcome Open Res. 2020;5:69. https://doi.org/10.12688/wellcomeopenres.15807.2.
106. Alarid-Escudero F, Krijkamp EM, Pechlivanoglou P, Jalal H, Kao SZ, Yang A, et al. A need for change! A coding framework for improving transparency in decision modeling. Pharmacoeconomics. 2019;37(11):1329–39. https://doi.org/10.1007/s40273-019-00837-x.
107. Alarid-Escudero F, Krijkamp E, Enns EA, Yang A, Hunink MGM, Pechlivanoglou P, et al. A tutorial on time-dependent cohort state-transition models in R using a cost-effectiveness analysis example. Med Decis Mak. 2023;43(1):21–41. https://doi.org/10.1177/0272989x221121747.
108. R Core Team. R: a language and environment for statistical computing: R Foundation; 2021. https://www.r-project.org/. Accessed 1 Mar 2023.
109. Bouvy J. Should NICE's cost-effectiveness thresholds change? 2024. https://www.nice.org.uk/news/blogs/should-nice-s-cost-effectiveness-thresholds-change-. Accessed 9 Jan 2024.
110. York Health Economics Consortium. Highly Specialised Technologies (UK NICE) [online] 2016. https://yhec.co.uk/glossary/highly-specialised-technologies-uk-nice/. Accessed 6 Mar 2024.
111. National Institute for Health and Care Excellence. NICE process and methods. Appendix 1: highly specialised technologies 2024. https://www.nice.org.uk/process/pmg46/resources/highly-specialised-technologies-nice-prioritisation-board-routing-criteria-pdf-19830673472965. Accessed 28 Aug 2025.
112. Fowler C. MSA Trust Research Strategy 2011. https://www.msatrust.org.uk/wp-content/uploads/2015/11/MSA-Trust-Research-Strategy.pdf. Accessed 9 Jan 2024.
113. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual 2022. https://www.nice.org.uk/process/pmg36/chapter/economic-evaluation-2. Accessed 23 Jan 2025.
114. Srivastava T, Latimer NR, Tappenden P. Estimation of transition probabilities for state-transition models: a review of NICE appraisals. Pharmacoeconomics. 2021;39(8):869–78. https://doi.org/10.1007/s40273-021-01034-5.
115. Jackson C. Multi-state models for panel data: the msm package for R. J Stat Softw. 2011;38(8):1–28. https://doi.org/10.18637/jss.v038.i08.
116. Chancellor JV, Hill AM, Sabin CA, Simpson KN, Youle M. Modelling the cost effectiveness of Lamivudine/Zidovudine combination therapy in HIV infection. Pharmacoeconomics. 1997;12(1):54–66. https://doi.org/10.2165/00019053-199712010-00006.
117. Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.
118. York Health Economics Consortium. Preferences-based measures [online] 2016. https://yhec.co.uk/glossary/preference-based-measures/. Accessed 2 May 2025.
119. Morton F, Singh J. eq5d: Methods for Analysing 'EQ-5D' Data and Calculating 'EQ-5D' Index Scores 2024. R package version 0.15.4. https://cran.r-project.org/web/packages/eq5d/index.html. Accessed 31 Oct 2024.
120. Hvidberg MF, Alava MH. Catalogues of EQ-5D-3L health-related quality of life scores for 199 chronic conditions and health risks for use in the UK and the USA. Pharmacoeconomics. 2023;41(10):1287–388. https://doi.org/10.1007/s40273-023-01285-4.
121. Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V. Catalogue of EQ-5D scores for the United Kingdom. Med Decis Mak. 2011;31(6):800–4. https://doi.org/10.1177/0272989x11401031.
122. Rowen D, Brazier J, Roberts J. Mapping SF-36 onto the EQ-5D index: how reliable is the relationship? Health Qual Life Outcomes. 2009;7(1):27. https://doi.org/10.1186/1477-7525-7-27.
123. Dams J, Zapp J-J, König H-H. Modelling the cost effectiveness of treatments for Parkinson's disease: an updated methodological review. Pharmacoeconomics. 2023;41(10):1205–28. https://doi.org/10.1007/s40273-023-01289-0.
124. Navarro CE, Ordóñez-Callamand E, Alzate JP. Disease modifying therapies in multiple sclerosis: cost-effectiveness systematic review. Farm Hosp. 2020;44(2):68–76. https://doi.org/10.7399/fh.11385.
125. Gallehzan NA, Khosravi M, Jamebozorgi K, Mir N, Jalilian H, Soleimanpour S, et al. Cost-utility and cost-effectiveness analysis of disease-modifying drugs of relapsing–remitting multiple sclerosis: a systematic review. Health Econ Rev. 2024;14(1):12. https://doi.org/10.1186/s13561-024-00478-7.
126. Llorian E, Waliji LA, Dragojlovic N, Michaux KD, Nagase F, Lynd LD. Frameworks for health technology assessment at an early stage of product development: a review and roadmap to guide applications. Value Health. 2023;26(8):1258–69. https://doi.org/10.1016/j.jval.2023.03.009.
127. Guzauskas GF, Tabrizi SJ, Long JD, Arnesen A, Hamilton JL, Claassen DO, et al. Long-term health outcomes of Huntington disease and the impact of future disease-modifying treatments: a decision-modeling analysis. Neurol Clin Pract. 2024;14(5):e200340. https://doi.org/10.1212/cpj.0000000000200340.
128. van Prooije T, Ruigrok S, van den Berkmortel N, Maas RPPWM, Wijn S, van Roon-Mom WMC, et al. The potential value of disease-modifying therapy in patients with spinocerebellar ataxia type 1: an early health economic modeling study. J Neurol. 2023;270(8):3788–98. https://doi.org/10.1007/s00415-023-11704-3.
129. Clarke S, Ellis M, Brownrigg J. The impact of rarity in NICE's health technology appraisals. Orphanet J Rare Dis. 2021;16(1):218. https://doi.org/10.1186/s13023-021-01845-x.
130. York Health Economics Consortium. Partitioned Survival Model [online] 2016. https://yhec.co.uk/glossary/partitioned-survival-model/. Accessed 11 Oct 2024.
131. Kunst N, Burger EA, Coupé VMH, Kuntz KM, Aas E. A guide to an iterative approach to model-based decision making in health and medicine: an iterative decision-making framework. Pharmacoeconomics. 2023. https://doi.org/10.1007/s40273-023-01341-z.
132. 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 Mak. 2012;32(5):678–89. https://doi.org/10.1177/0272989x12454941.

133. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices–overview: a report of the ISPOR-SMDM modeling good research practices task force-1. Med Decis Mak. 2012;32(5):667–77. https://doi.org/10.1177/0272989x12454577.
134. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM modeling good research practices task force-7. Med Decis Mak. 2012;32(5):733–43. https://doi.org/10.1177/0272989x12454579.
135. Whitty CJM. What makes an academic paper useful for health policy? BMC Med. 2015;13(1):301. https://doi.org/10.1186/s12916-015-0544-8.
136. Xin Y, Gray E, Robles-Zurita JA, Haghpanahan H, Heggie R, Kohli-Lynch C, et al. From spreadsheets to script: experiences from converting a Scottish cardiovascular disease policy model into R. Appl Health Econ Health Policy. 2022;20(2):149–58. https://doi.org/10.1007/s40258-021-00684-y.
137. Naylor NR, Williams J, Green N, Lamrock F, Briggs A. Extensions of health economic evaluations in R for Microsoft Excel users: a tutorial for incorporating heterogeneity and conducting value of information analyses. Pharmacoeconomics. 2023;41(1):21–32. https://doi.org/10.1007/s40273-022-01203-0.
138. Krijkamp EM, Alarid-Escudero F, Enns EA, Jalal HJ, Hunink MGM, Pechlivanoglou P. Microsimulation modeling for health decision sciences using R: a tutorial. Med Decis Mak. 2018;38(3):400–22. https://doi.org/10.1177/0272989x18754513.
139. Green N, Lamrock F, Naylor N, Williams J, Briggs A. Health economic evaluation using Markov models in R for Microsoft Excel users: a tutorial. Pharmacoeconomics. 2023;41(1):5–19. https://doi.org/10.1007/s40273-022-01199-7.
140. Ben ÂJ, van Dongen JM, El Alili M, Esser JL, Broulíková HM, Bosmans JE. Conducting trial-based economic evaluations using R: a tutorial. Pharmacoeconomics. 2023;41(11):1403–13. https://doi.org/10.1007/s40273-023-01301-7.
141. Smith RA, Mohammed W, Schneider PP. Packaging cost-effectiveness models in R: A tutorial. [version 1; peer review: 2 approved with reservations]. Wellcome Open Res. 2023(8):419. https://doi.org/10.12688/wellcomeopenres.19656.1.
142. Krijkamp EM, Alarid-Escudero F, Enns EA, Pechlivanoglou P, Hunink MGM, Yang A, et al. A multidimensional array representation of state-transition model dynamics. Med Decis Mak. 2020;40(2):242–8. https://doi.org/10.1177/0272989x19893973.
143. Poewe W, Kaufmann H, Quinn N. In Memoriam Gregor K. Wenning (1964–2024). Mov Disord. 2024;39(7):1085–6. https://doi.org/10.1002/mds.29896.

Authors and Affiliations

Tobias Sydendal Grand[1,2] · Shijie Ren[1] · Praveen Thokala[1,7] · Stefano Zanigni[2] · Daniel Oudin Åström[2] · Aroussi Bidani[2] · Günter Höglinger[3,4,5] · Werner Poewe[6] · Florian Krismer[6] · Stephane Regnier[2] · the EMSA-SG Natural History Study Investigators

✉ Tobias Sydendal Grand
tsgrand1@sheffield.ac.uk

Shijie Ren
s.ren@sheffield.ac.uk

Praveen Thokala
praveen.thokala@gmail.com

Stefano Zanigni
STZN@lundbeck.com

Daniel Oudin Åström
AAST@lundbeck.com

Aroussi Bidani
ARBD@lundbeck.com

Günter Höglinger
Guenter.Hoeglinger@med.uni-muenchen.de

Werner Poewe
werner.poewe@i-med.ac.at

Florian Krismer
florian.krismer@i-med.ac.at

Stephane Regnier
STRR@lundbeck.com

1 Sheffield Centre for Health and Related Research (SCHARR), University of Sheffield, 30 Regent St, Sheffield City Centre, Sheffield S1 4DA, UK

2 Lundbeck A/S, Ottiliavej 9, 2500 Copenhagen, Denmark

3 Ludwig Maximilian University of Munich, Geschwister-Scholl-Platz 1, 80539 Munich, Germany

4 German Center for Neurodegenerative Diseases (DZNE), Von-Siebold-Straße 3A, 37075 Göttingen, Germany

5 Munich Cluster for Systems Neurology (SyNergy), Feodor-Lynen-Straße 17, 81377 München, Germany

6 Department of Neurology, Medical University Innsbruck, Christoph-Probst-Platz, Innrain 52, 6020 Innsbruck, Germany

7 PT Health Economics Limited, 41a, Ashland Road, Sheffield S7 1RH, UK

IV.III Conceptual modelling of functional milestones in multiple system atrophy: comparing Markov and partitioned survival models using digitised time-to-event data

Most authors have approved this version of the manuscript. I have redacted the author names to maintain confidentiality until the submission-ready version is finalised.

Conceptual modelling of functional milestones in multiple system atrophy: comparing Markov and partitioned survival models using digitised time-to-event data

Abstract

Background and objectives

Multiple system atrophy is a devastating, fatal, and rare disease with a prevalence of approximately 5 cases per 100,000 people in the United Kingdom. This study explores functional milestones using Markov and partitioned survival modelling for multiple system atrophy to triangulate previous modelling results informed by disability scales. We firstly assess health-economic parameters for modelling, e.g., if individual patient-level data are available or need estimation. We then evaluate functional milestones and their advantages and disadvantages for model strategies (Markov and partitioned survival models). Lastly, we explore if it is worthwhile exploring expected value of perfect information in rare disease using multiple system atrophy as a proxy.

Methods

The analysis applied a United Kingdom national health service perspective with a cost-effectiveness threshold of £30,000 per quality-adjusted life year. Kaplan-Meier curves for functional milestones were digitised and individual patient-level data reconstructed. Markov and partitioned survival models were constructed in R for cost-effectiveness analysis using similar health-economic parameters, apart from the model-specific parameters, e.g., transition probabilities and parametric equations, respectively.

Results

The deterministic incremental cost-effectiveness ratios for Markov and partitioned survival models were £179,362 and £131,380, respectively. The curves that represented expected value of perfect information had similar shapes, but the curve was slightly higher for the partitioned survival model compared to the Markov model from approx. £30,000 – 170,000 willingness to pay thresholds, whereafter the curves converged.

Discussion

The pros and cons of using functional milestones to inform cost-effectiveness modelling in comparison to disability scales (questionnaires) were investigated. Partitioned survival modelling was feasible despite data scarcities, and it could be considered for further examination. The relevance of value of information analysis in rare diseases warrants further research.

Key points for decision makers

Partitioned survival models in multiple system atrophy informed by functional milestones were explored as a potential alternative to Markov models informed by disability scales. Functional milestones for modelling seem to confirm robustness of previous Markov models informed by disability scales as results were similar despite differences in model type and in certain health-economic parameters. It may be worth exploring partitioned survival modelling using disability scales.

Value of information analysis in rare diseases requires special considerations and further research to determine applicability and relevance.

Digitised and reconstructed data are useful for triangulation or conceptual modelling if individual patient-level data are not available despite potential limitations.

1 Background

Multiple system atrophy (MSA) is a debilitating and rare disease with a prevalence of approximately 5 cases per 100,000 people in the United Kingdom (UK) [1, 2]. The mean age at onset is 50 to 59 years with an expected survival from onset of symptoms of 6.2 to 9.8 years [3-7]. Age is the only accepted risk factor. Diagnosis is difficult and often mistaken for other diseases, which often results in inadequate treatment [7, 8]. Guidelines define clinical symptoms by two phenotypes, MSA-P where symptoms of parkinsonism are dominating. For example, bradykinesia, rigidity, tremor, or freezing gait [7, 8]. MSA-C shows primarily cerebellar symptoms such as gait, limb ataxia, explosive speech, and cerebellar oculomotor disturbances [7-9].

Treatment aims at managing symptoms, because no disease modifying treatments have been licensed, which include multiple specialities like neurology, cardiology, gastroenterology, and urology. Social services play a large part for end-of-life planning and for informal caregivers [10-12]. The functional milestones are aid-required walking, wheelchair confinement, bedridden and death [13]. In addition, key milestones beyond ambulation include intelligible speech, swallowing issues, tube feeding, indwelling of catheters or severe orthostatic hypotension with frequent dizziness [2].

Grand et al. (2024) found that the main challenges for economic evaluation of rare diseases were lack of evidence and a scarcity of economic evaluations, but suggested timely and iterative cost-effectiveness modelling, for example, with expected value of perfect information (EVPI) analyses to mitigate challenges [14]. EVPI indicates if additional research is potentially worthwhile [15, 16]. It is essentially a comparison between a scenario with current evidence as opposed to a scenario with perfect evidence, e.g., assuming an infinite sample size [17]. Grand et al. (2025) conducted an early health technology assessment (HTA) with a user-friendly model of MSA wherein a unified MSA rating scale (UMSARS) informed the model structure [18].

This study was motivated by a recommendation for exploration of partitioned survival modelling using functional milestones (time-to-event analyses) for model triangulation purposes [18]. However, from a clinician or health systems point of view, there may be an added benefit of functional milestones if an intervention is approved in the future for an eligible population, which is that it is much easier to intuitively confirm if someone is able to walk as opposed filling out an UMSARS questionnaire, which requires more time, and might be difficult for the patient or even impossible in late disease stages. In addition, responses to questionnaires may fluctuate over shorter periods, whereas functional milestones ought to follow a fixed sequential pattern as there are no disease modifying treatments to reverse disease progression.

We firstly investigate if individual patient-level data (IPD) can be digitised and reconstructed from available publications in MSA. We then examine if alternative model structures can be constructed using functional milestones to evaluate advantages and disadvantages of model strategies (Markov and partitioned survival models). Lastly, we explore if it is worthwhile exploring EVPI for MSA.

2 Methods

This section outlines the methods used for model development and estimation of health-economic parameters.

2.1 Study population and perspective

The perspective of the analysis was a UK national health service (NHS) perspective with a cost-effectiveness threshold of £30,000 per quality-adjusted life year (QALY) [19]. Discount rates were 3.5% for costs and effects [20]. A health economic analysis plan was not developed for this study because the analysis was primarily pragmatic or conceptual in nature. The diagnosis criteria for the modelled population were the Gilman criteria [9]. Discrete-time points from Kaplan-

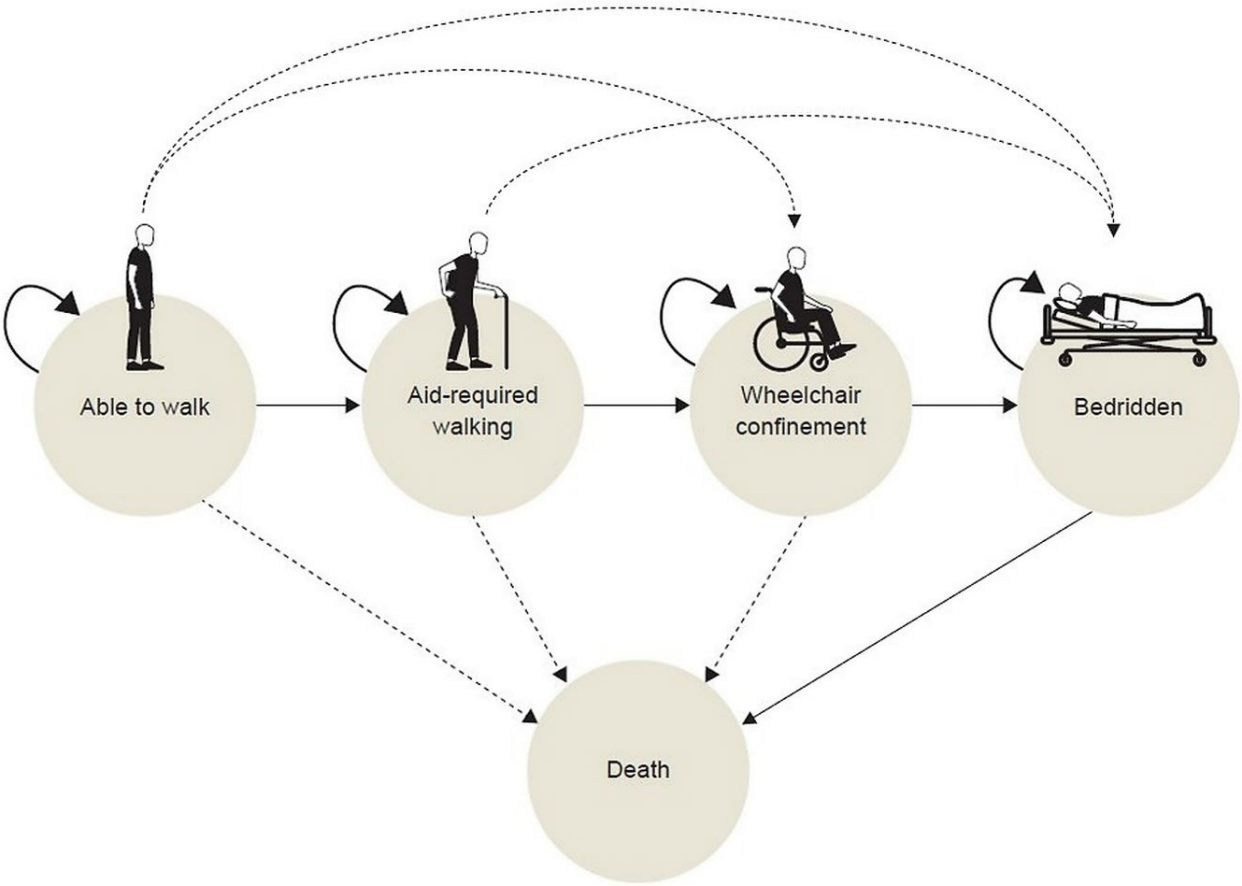
Meier curves were reported annually, which guided the decision to use an annual cycle length in the models [13]. The time horizon was 32 years, because few subjects were still at risk of death (meaning they remained alive) as illustrated by published Kaplan-Meier curves at 12 and 25 years, depending on whether their duration of survival was estimated from the time point of probable diagnosis or symptom onset, respectively [5, 13]. In addition, a modelled treatment modifier in the study resulted in slowing of disease progression, which meant that a longer time horizon was needed to capture downstream consequences, e.g., effects and costs from the intervention arm. For example, the minimum time horizon to achieve this was 32 years.

2.2 Model structure and development

An R project includes code for data analysis and results (See code availability). The modelling report is a Markdown report (html file) with a dropdown button named "show", which opens relevant code chunks when clicked [21, 22]. The Markdown report is available from: https://tobiasgrand.github.io/Report_PSM_Markov_MSA/.

The two models were based on either a partitioned survival or Markov modelling approach [22]. They shared the model structure from Fig. 1, which had five health states informed by functional milestones: able to walk, aid-required walking, wheelchair confinement, bedridden and death. The progression was sequential through model health states and mortality was only allowed from the bedridden state, but we expand the explanation of the individual model structures in Sections 2.3.2 and 2.3.3. The comparators were best supportive care (BSC) versus a hypothetical drug given alongside BSC because disease modifying treatments are currently missing.

Fig. 1: Model structure with functional milestones in MSA



* Note that in Fig 1. The sequential transitions are illustrated by the unbroken lines. The dotted lines represent non-sequential transitions that are allowed in the Markov model because they represent a series of sequential transitions occurring within one model cycle, which results in an apparent transition between non-sequential health states within one cycle. See sections 2.3.2 and 2.3.3 for additional details.

2.3 Estimated health-economic parameters

Health-economic parameters were categorised as either estimated or other health economic parameters. The estimated health-economic parameters were those that required significant resources and efforts to derive. For example, it was necessary to digitise Kaplan-Meier curves and reconstruct IPD data to estimate transition probabilities for the Markov model and parametric survival models for the partitioned survival model. By contrast, other health-economic parameters were those that could be sourced from published papers, e.g., costs, utilities, and disutilities.

2.3.1 Digitised functional milestones and reconstructed IPD data

Kaplan-Meier curves for functional milestones were digitised from a natural history of disease study in the absence of access to original IPD [13]. This study was to our knowledge the only published study that reported time to all five functional milestones as Kaplan-Meier curves in MSA [13]. The study used Gilman 1999 diagnosis criteria, and had up to 17 years of follow-up data [23].

The R package IPDfromKM, which is informed by an iterative Kaplan-Meier estimation algorithm developed by Guyot and colleagues, was used to reconstruct pseudo IPD. As only the total number at risk was available in the source paper, a strong assumption of no censoring was required when reconstructing the pseudo IPD due to the lack of information [24, 25]. It means that uncertainty in reconstructed hazard rates was underestimated [24].

The event curves were reported in four different figures in the source paper. Data for time to aid-required walking/wheelchair confinement/bedridden/death were reconstructed independently. The reconstructed data were then visually compared for accuracy and revised to improve accuracy if required, e.g., to check for manual clicking errors, against the original figure. See the Markdown report (section 1.1.3) for illustrations of the reconstructed Kaplan-Meier curves: https://tobiasgrand.github.io/Report_PSM_Markov_MSA/.

In addition, to constructing a dataset for a hypothetical drug & BSC an efficacy modifier was implemented. It meant that the time to an event (milestones) was divided by an arbitrary hazard rate of 0.67 to reflect an intervention that delays disease progression by approximately 50%. Efficacy modifications have been used and discussed elsewhere for early modelling in the absence of a licensed disease-modifying treatment [18].

The process for reconstructing IPD had multiple steps. First, subjects (n = 230) were generated using their sequence number (1 to 230) as subject id's, for each of the comparator arms, which resulted in 460 (1-230; 231-460) reconstructed patients assigned to the events. The subjects were assigned their functional milestones (events) using each of the four event datasets for the two interventions. The process for this required looping through annual cycles whilst assigning events based on the cumulative probabilities from the curves and incremental number of subjects with events. The subjects were sequentially assigned these events, e.g., looping from 1 to 230 while assigning events, if any subjects remained unassigned at later cycles, they were assigned the maximum observed time for their event. It meant that subjects with lower subject id's were consistently assigned the earlier events, which resulted in those subjects having faster progression of events.

Our approach meant that subjects were forced sequentially through the model health states, and this made sure that the distributional properties of the reconstructed Kaplan-Meier curves were preserved, and that the reconstructed IPD could be used for fitting distributions prior to cost-effectiveness modelling. This approach allowed for consistency between the

Markov and partitioned survival models, as both models used the same sequential progression assumption in their underlying data.

2.3.2 Markov model

The Markov model was hardcoded but required that transition probabilities were estimated from the reconstructed datasets, which were organised for use in the msm R package that requires that time is a continuous process underlying the data [26]. The dataset fulfilled the minimum criteria that were: subject id, time, and state, but were moreover ordered by intervention. This package facilitates multistate modelling that estimates transition probabilities by maximum likelihood estimation [26]. Multistate modelling has not seen widespread use in health economics yet, but a review found that it was the third most frequent method for estimating transition probabilities in National Institute for Health and Care Excellence (NICE) appraisals after survival analysis and count methods [27].

In the msm R package, the Q matrix (transition intensity matrix) defines which transitions between states are allowed to reflect the model structure, for example, the assumption of sequential transitions. The transition intensity values were manually calibrated in R until a reasonable relationship between observed and predicted prevalence by time was visually confirmed. A higher intensity value indicates that a transition is more likely to occur whereas zeros indicate disallowed transitions. It meant that the strict assumption of sequential progression was applied in the coding by specifying the Q matrix accordingly. However, it is important to highlight that unobserved transitions may occur at any time, because of the underlying assumption of continuous time. This assumption was partly violated as reconstructed data were panel data, e.g., observed at discrete time intervals. For example, for longer cycle lengths, transitions across multiple health states may be observed, e.g., from able to walk directly to wheelchair confinement despite not being permitted by the Q matrix. They do, however, represent a sequence of unobserved transitions occurring between discrete time points. For example, an intermediate sequence from able to walk to aid-required walking to wheelchair confinement within an annual cycle. This problem could likely be solved with a shorter cycle length, but an annual cycle length was enforced for consistency with the partitioned survival model [26]. It was decided to reflect the observed transitions as dotted lines in the model structure (Fig. 1) despite representing an unobserved sequence of events. It represents a limitation of the Markov modelling approach, because cost and QALYs were not accounted for in the intermediate (unobserved) states. A Dirichlet distribution, which is a multivariate generalisation of the beta distribution, was assigned to transition probabilities for the probabilistic sensitivity analysis (PSA), which is a method to assess parameter uncertainty through random sampling from probability distributions [28].

2.3.3 Partitioned survival model

The partitioned survival model naturally required distinct methodological techniques that differed from those employed in the Markov model. For example, the digitised time-to-event datasets were fitted individually for each event curve using exponential, Weibull, logistic, log-normal, and exponential distributions and their fit evaluated by Akaike's information criterion (AIC) and Bayesian information criterion (BIC). Weibull distributions provided the best fit based on these criteria for all event curves which meant that Weibull distributions were fitted to each of the events when the event datasets were combined by subject id into parametric models for the two comparators.

The partitioned survival modelling was conducted with the hesim R package for which the authors had written a tutorial for a 4-state partitioned survival model, which supported the decision to use this R package [29, 30]. Importantly, the subjects with faster progressions, e.g., those that caused unobserved transitions in the multistate modelling (direct transitions through intermediate states), were handled with greater flexibility in the partitioned survival modelling. The fractional time spent in a health state within a model cycle was obtained from the area differences in cumulative event probability curves

[31]. This approach ensured that subjects who progressed through multiple states during one model cycle still accumulated costs and effects from their time in these intermediate states, which contrasts with the Markov model approach.

The deterministic analysis made use of a single simulation with no sampling from distributions for the parametric time-to-event models and health-economic parameters. The probabilistic sensitivity analysis ran 500 simulations to limit the computational burden, and a normal distribution was applied for time-to-event models.

2.4 Other health-economic parameters

2.4.1 Costs

A previous literature review found two cost-of-illness studies from 2011, which meant that health-economic parameters for MSA were scarce [18, 32, 33]. However, a recent study focussed on resource utilisation (frequency and odds ratios), which was not sufficient to inform the modelling [34]. Importantly, stratification of costs was only reported in one study by UMSARS Part IV, which meant that it was necessary to assume that costs stratified by UMSARS Part IV were approximate of functional milestones, which was an apparent limitation [35]. The German cost data were adjusted for inflation, euros converted to British pounds, converted from six-monthly to annual costs, and they were assumed generalisable to a UK setting, which was also necessary in a previous cost-effectiveness analysis [18]. For transparency, an R script was coded for the sourcing and preparation of cost data. The drug costs were £12,000 per year, which was an arbitrary cost estimate. The drug costs were not applied to the bedridden state, which was regarded as a stopping rule by clinical experts. Table 1 includes the cost data across health states.

Table 1: Cost data and parameters for functional health states

State	Mean costs	Distribution	Shape*	Scale*
Able to walk	£11,853	Gamma	46	257.67
Aid-required walking	£24,002			521.79
Wheelchair confinement	£44,152			959.83
Bedridden	£44,152			959.83

*Note that standard deviations to derive shape and scale were estimated based on total costs and number of subjects. Values are rounded.

2.4.2 Utilities

Utility data were not available for functional milestones, which meant that a similar assumption of UMSARS Part IV approximating functional milestones was necessary. It allowed for the data to be sourced from a previous paper, which are reported in Table 2 [18].

Table 2: Mean index scores for functional milestones

Functional milestones	Mean Index Score*
Able to walk	0.58
Aid-required waling	0.43
Wheelchair confinement	0.14
Bedridden	-0.13

*Point estimate for mean EQ-5D-3L index score (UMSARS Part IV). Note that numeric values are rounded to two decimals for illustration, but all decimals were used for modelling.

The disutilities from Table 3 were sourced from literature following a previous approach, which selected the most conservative estimates for inclusion in a cost-effectiveness analysis [18].

Table 3: Disutilities and probabilities for clinical events

Clinical events	Probabilities*	Reference	Disutilities*	References
Hypotension disorders	56.7 %	Wenning and colleagues [5]	-0.07	Xiao and colleagues [36]
Urinary or bladder disorders	51.1 %		-0.01	Sullivan and colleagues [37]
Bowel disorders	58.2 %		-0.19	Hvidbjerg and Alava [38]

* Note all numeric values are rounded to two decimals.

The disutilities were multiplied by their annual probabilities, which produced disutility decrements that were subtracted from health-state utilities, which is an approach previously used elsewhere [18]. However, for the hypothetical drug & BSC, an efficacy modifier was implemented to account for a reduction in clinical events resulting from an expected treatment effect. The event modifier had a value of 0.5, which was multiplied with event probabilities, and the utility decrement subtracted from health-state utilities. In PSA, a beta distribution was only applied in cases where alpha and beta for the beta distribution were positive, which were otherwise fixed. The alpha and beta values were estimated from standard deviations [18]. The derived health states utilities and parameters are shown in Table 4.

Table 4: Mean health-state utilities after adjusting for disutility increments, and distributions

Comparators	Health states	Mean utility	Distribution	Alpha	Beta
BSC	Able to walk	0.43	Beta	1.80	2.43
	Aid-required walking	0.27		0.48	1.28
	Wheelchair confinement	-0.01	fixed		
	Bedridden	-0.28			
Hypothetical drug & BSC	Able to walk	0.50	Beta	2.18	2.17
	Aid-required walking	0.35		0.76	1.41
	Wheelchair confinement	0.07	fixed		
	Bedridden	-0.28			

*Note that numeric values are rounded to two decimals. BSC (Best Supportive Care).

2.5 Reporting of analyses

Multiple analyses were reported for both models and their treatment strategies. First, incremental cost-effectiveness ratio (ICER) tables were produced for comparison of deterministic and probabilistic results. Second, line plots informed by Markov traces illustrated the state occupancy across annual model cycles. Third, PSA results were visually represented by a cost-effectiveness plane, cost-acceptability curves, and used to calculate EVPI.

Lastly, results were explored using a QALY shortfall calculator while following relevant guidance from NICE's decision support unit on severity modification by QALY weights [39-41]. Age at disease onset was 55.4 years and female study participation approx. 43% to align with the digitised data [13]. The QALY weights should be applied to the incremental QALYs, which results in the following formula for calculating weighted ICERs:

$$ICER_{weighted} = \frac{Costs_{Hypothetical\ Drug\ \&\ BSC} - Costs_{BSC}}{\left(QALYs_{Hypothetical\ Drug\ \&\ BSC} - QALYs_{BSC}\right) x\ QALY\ Weight}$$

3. Results

3.1 Cost-effectiveness results

The cost-effectiveness results in Table 5 shows differences between the deterministic and probabilistic analyses of £21,000 and £5,000 for the partitioned survival and Markov model, respectively. There was a larger difference when comparing the different model types. For example, across the model types when comparing deterministic analyses, they were £131,380 and £179,362 per QALY, and for probabilistic £151,944 and £183,808 per QALY. The partitioned survival model provided the most cost-effective estimates that were in favour of the hypothetical drug.

Table 5: ICER table with cost-effectiveness results across models and analyses

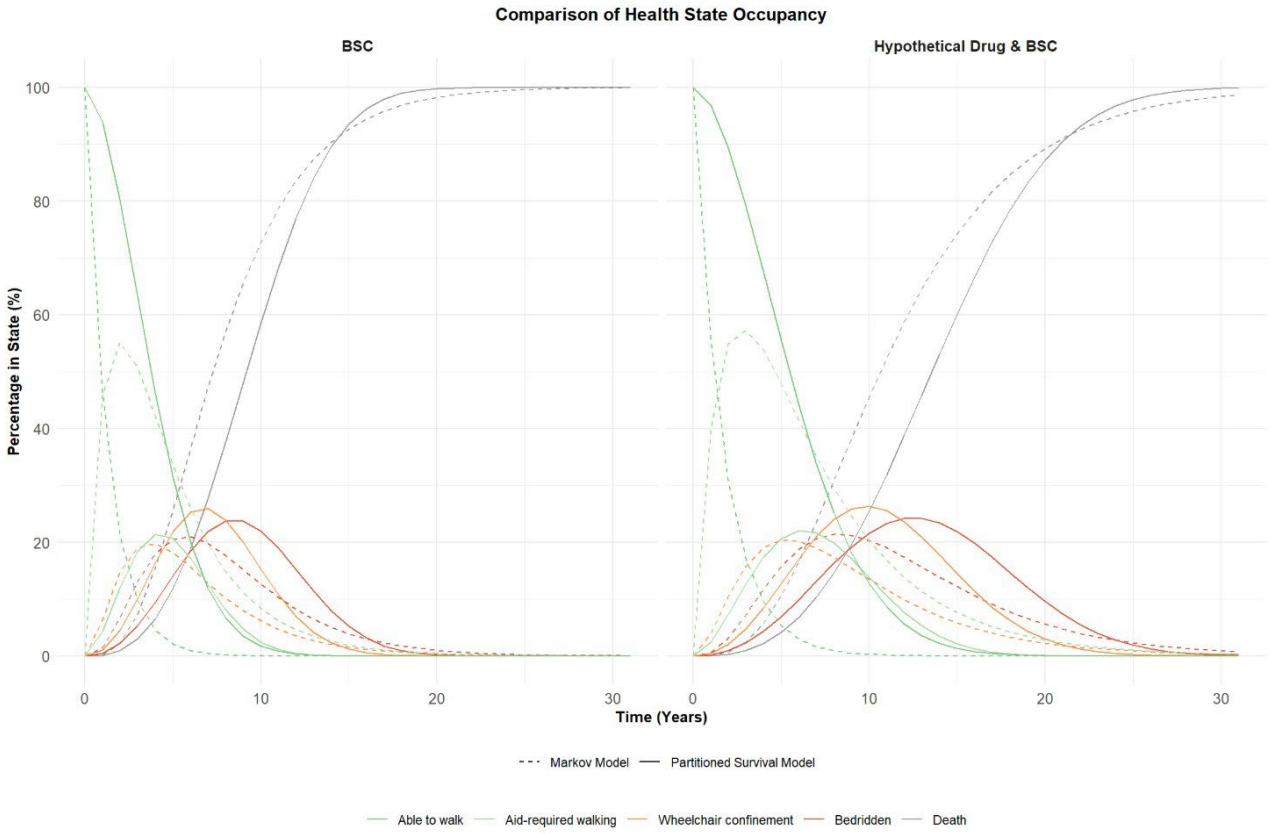
Model	Analysis	Intervention	Costs	QALYs	ICER
Partition Survival model	Deterministic	Hypothetical Drug & BSC	£357,937	2.42	
		BSC	£189,607	1.14	131,380
	Probabilistic	Hypothetical Drug & BSC	£374,840	2.66	
		BSC	£201,966	1.53	151,944
Markov model	Deterministic	Hypothetical Drug & BSC	£375,882	2.08	
		BSC	£208,437	1.14	179,362
	Probabilistic	Hypothetical Drug & BSC	£377,759	2.10	
		BSC	£209,328	1.18	183,808

* Note that the table contains rounded values, but all decimals were used in the R code (calculations). BSC (Best Supportive Care).

3.2 State occupancy

Fig. 2 contains an illustration of state occupancy by percentage in state and time in years on the y- and x-axis, respectively from both Markov and partitioned survival models. The graphics show that a hypothetical drug & BSC prolongs the time spent in earlier states as compared to BSC alone. In addition, it is also clear that the two model types use different methods to estimate state occupancy resulting in differently shaped curves. For example, the curve for the able to walk state is less steep in the partitioned survival model, which equates to increased time in this state.

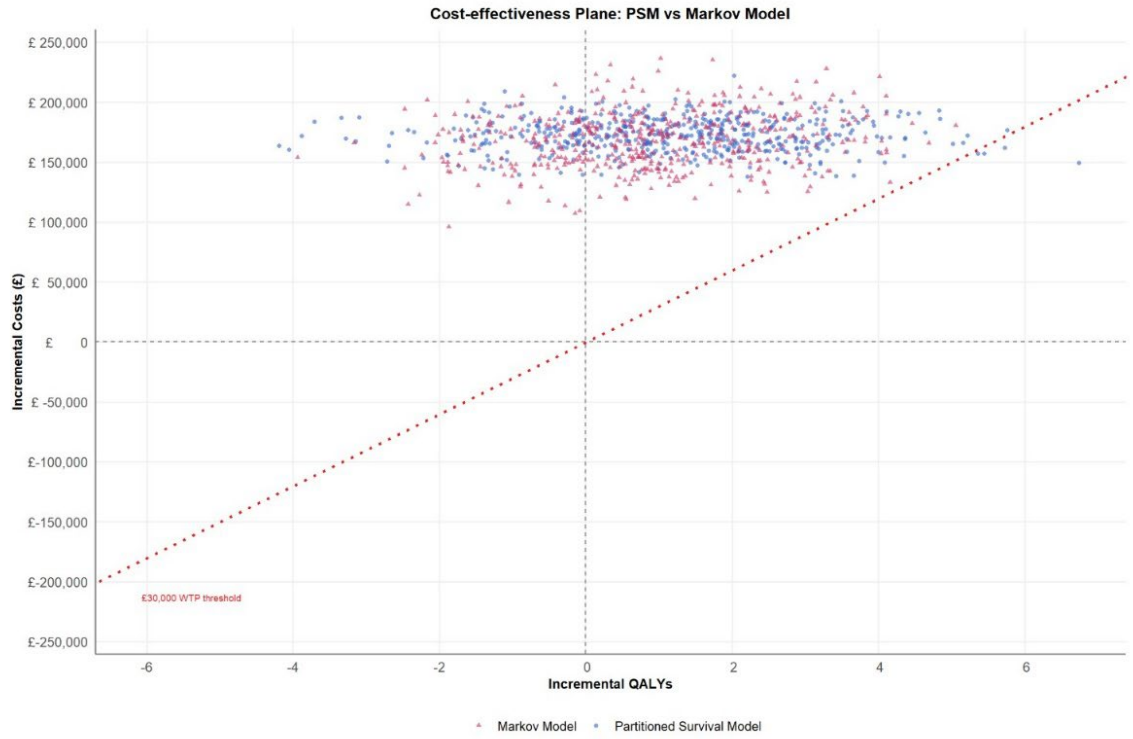
Fig. 2: Comparison of health state occupancy



3.3 Cost-effectiveness plane and acceptability curves

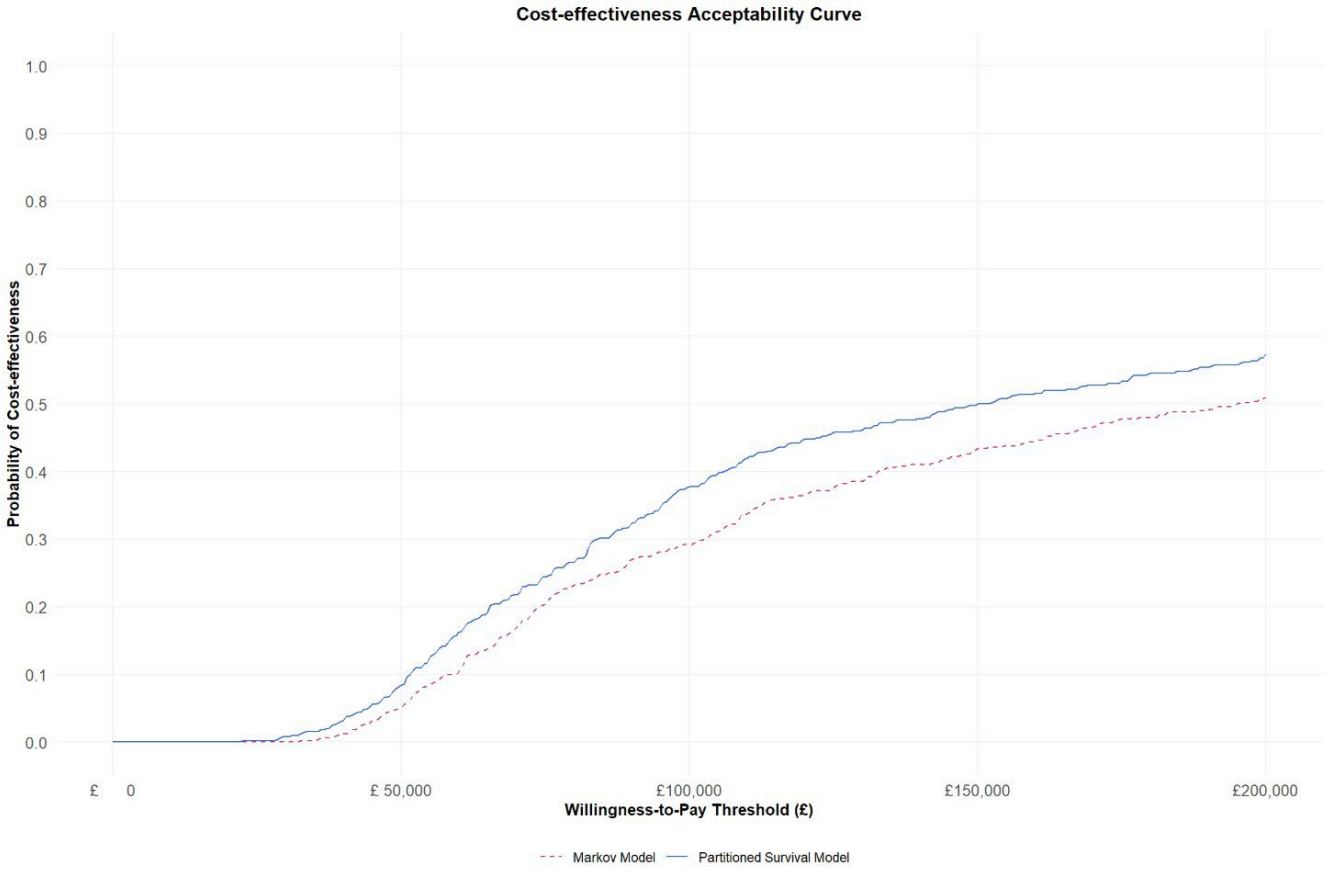
The cost-effectiveness plane is illustrated in Fig. 3, which largely showed that the simulations from probabilistic sensitivity analyses were overlapping, but different patterns are noticeable. For example, the partitioned survival model seems to generate simulations with a wider range for incremental QALYs, and less incremental cost differences as compared to the Markov model.

Fig. 3: Cost-effectiveness plane



The cost-effectiveness acceptability curves in Fig. 4 showed that a 0.5 probability of cost-effectiveness was reached at around a £150,000 willingness-to-pay threshold for the partitioned survival model, but closer to a threshold of £195,000 for the Markov model.

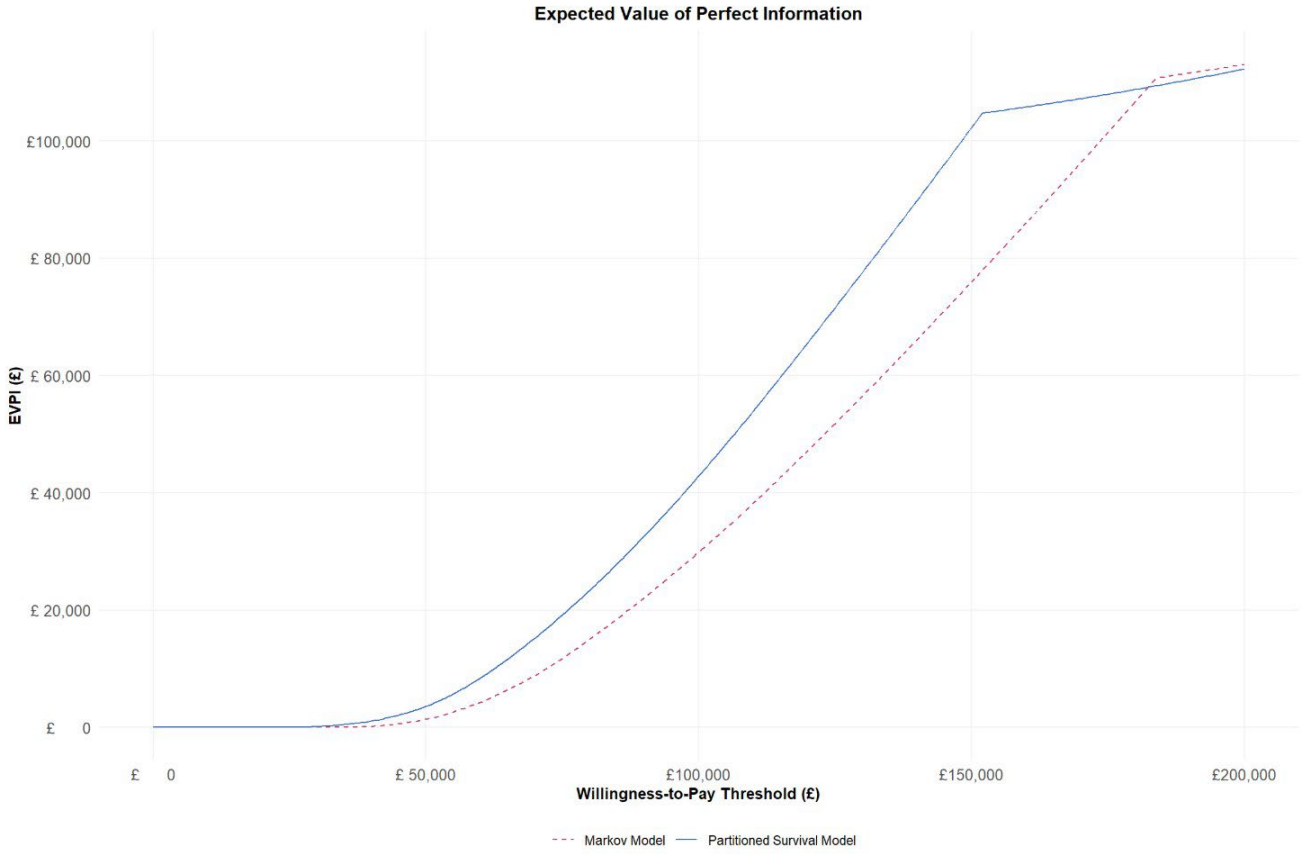
Fig. 4: The cost-effectiveness acceptability curves



3.4 EVPI

Fig. 5 shows the EVPI curves, which have similar shape, but EVPI is slightly higher for the partitioned survival model compared to the Markov model from approx. £30,000 – 170,000 willingness to pay thresholds, whereafter the curves converge.

Fig. 5: EVPI at the individual level



3.4 QALY weights for disease severity

NICE guidance states that QALY weights can be applied to reflect disease severity, with the weighting dependent on the lifetime QALY shortfall under BSC [40]. The QALY shortfall calculator uses outputs from cost-effectiveness analyses and population characteristics to provide a QALY weight, which was 1.2 for all four analyses examined [39-41]. This was then applied to estimate the weighted ICERs. The ICERs with weights applied for disease severity, show in Table 6, are naturally lower than results without the weighting, shown in Table 5. However, they are still above the cost-effectiveness threshold of £30,000 per QALY.

Table 5: Results with QALY weights from the shortfall calculator

Model	Analysis	Intervention	Costs	QALYs	Weight	$ICER_{weighted}$
Partition Survival model	Deterministic	Hypothetical Drug & BSC	£357,937	2.42		
		BSC	£189,607	1,14	x 1.2	131,508
	Probabilistic	Hypothetical Drug & BSC	£374,840	2.66		
		BSC	£201,966	1.53	x 1.2	127,488

Markov model	Deterministic	Hypothetical Drug & BSC	£375,882	2.08		
		BSC	£208,437	1,14	x 1.2	148,444
	Probabilistic	Hypothetical Drug & BSC	£377,759	2.10		
		BSC	£209,328	1.18	x 1.2	152,564

* Note the QALY shortfall calculator settings: discount rate 3.5%, mean age 54 years, female 43%, scenario "reference case", and remaining discounted QALYs from BSC (called current standard of care in Shiny app). In addition, the rounded data from Table 5 were used for these calculations with the shortfall calculator. BSC (Best Supportive Care)

4. Discussion

With this present study, we have demonstrated that it is possible to reconstruct and digitise IPD data for cost-effectiveness modelling. Moreover, by using functional milestones for modelling using, e.g., partitioned survival models, we increase confidence in previous modelling results, despite signification limitations in terms of data availability. To our knowledge, this is the first study to apply and discuss the potential of EVPI to inform data collection priorities in rare disease using MSA as a proxy.

4.1 What are the implications of using alternative model structures and strategies?

It is fundamental for the scientific merit of this study to highlight the differences and similarities between UMSARS Part IV and functional milestones, and why it makes sense to explore the functional milestones through conceptual modelling. The very first study of cost effectiveness in MSA argued for the use of disability scales, e.g., UMSARS Part IV to inform model structures, which was based on a thorough investigation of model structures for disease analogues, which often use disability or functional scales [18]. The disease analogues were progressive diseases of the central nervous system featuring symptoms overlapping with MSA. It was further hypothesised that increased costs and reduced quality-of-life data were correlated with disease progression on the disease-specific disability scale UMSARS Part IV [18, 33]. The relationship between UMSARS Part IV and functional milestones was first illustrated by Saulnier and colleagues, and Table 7 contains a simplified illustration for comparison [42]. It is important to note that the first two items from UMSARS Part IV will need to represent the milestone, able to walk, despite being separate items, for a modelling comparison to be feasible. Noteworthy is that these two items were collapsed in a cost-of-illness study, which could potentially indicate comparable costs, because previously, utilities were found to be of comparable values in a cost-effectiveness analysis [18, 32]. However, collapsing health states may result in small incremental differences being lost. Bedridden appears both in UMSARS Part IV item 5, but is also a functional milestone, which is interesting since the functional milestones aid-required walking and wheelchair are not explicitly mentioned.

Table 7: Comparison of UMSARS Part IV and functional milestones

UMSARS Part IV	Completely independent. Able to do all chores with minimal difficulty or impairment.	Not completely independent. Needs help with some chores.	More dependent. Help with half of chores. Spends a large part of	Very dependent. Now and then does a few chores alone or begins	Totally dependent and helpless. Bedridden.

	Essentially normal. Unaware of any difficulty.	the day with chores.	alone. Much help needed.	
Functional milestones	Able to walk	Aid-required walking	Wheelchair	Bedridden

Data have shown that UMSARS Part IV responses may fluctuate over shorter periods, e.g., six-monthly intervals, likely due to day-to-day symptoms, but this is to a limited degree [18]. However, a Markov model allowing backwards transitions would be needed if researchers were to account for this, but it may depend on the dataset at hand. Moreover, this issue may be more pronounced when follow-up intervals are short, because people living with MSA are likely to progress in terms of severity measured by UMSARS Part IV over time. By contrast, functional milestones are likely to have a stronger adherence to a sequential progression assumption, e.g., progression from able to walk to aid-required walking and so forth. Although, e.g., wheelchair use will gradually increase. It is almost certain once this milestone is reached that a wheelchair will never not be needed, only if bedridden, because disease progression will not revert. This finding has important implications for model development, which is that partitioned survival modelling becomes a viable option, if sequential progression can be assumed, which would be interesting to explore for UMSARS Part IV. Naturally, Markov modelling will remain an option with or without this progression assumption.

4.2 How do model structures and data sources influence cost-effectiveness?

The first cost-effectiveness analysis in MSA, which was a Markov model informed by UMSARS Part IV health states, reported an ICER for the deterministic analysis of £142,036, which is lower than the £179,362 for the Markov model informed by functional milestones but higher than the £131,380 from the partitioned survival model in this analysis (without QALY weights applied) [18]. The present study findings largely confirm previous study results and contributes additional evidence that suggests that models are robust, despite differences in study designs and data inputs.

Nonetheless, the result differences that were observed between cost-effectiveness analyses and model types were likely attributed to structural and parameter differences, e.g., Markov versus partitioned survival model. For example, the ability of partitioned survival models to handle fractional state occupancy within one model cycle through area differences in cumulative event probability curves. In addition, the previous cost-effectiveness analysis was informed by IPD from European MSA Study Group (EMSA-SG) with a 2-year follow-up for UMSARS Part IV whereas this analysis digitised functional milestones where subjects were followed for 17 years [18]. The digitisation and reconstructed IPD adds to the uncertainty of cost-effectiveness results in this present analysis, e.g., assumptions on censoring. Moreover, this analysis required a sequential progression assumption, which was not the case in the previous cost-effectiveness analysis since they had access to IPD which was major advantage, e.g., allowing deaths to occur from all health states. In addition, the previous model used a shorter cycle length and thereby avoided intermediate state transitions from multistate modelling [18]. The efficacy modification was also different in that their method for efficacy incorporation modified transition probabilities directly in a transition matrix. This contrasts with the method used in this study where time to events were increased in datasets using an arbitrary hazard rate, which meant that the same dataset could later be used in both model types [18, 28, 43]. The hazard rate was applied to all health states, e.g., also for bedridden and death, which was done to avoid an overlap or unrealistic sequence of events, e.g., bedridden occurring before wheelchair confinement. It means that there was a costing discrepancy in that drug costs were halted, but the treatment modification technically persists at the

bedridden state. Whilst this was a limitation, this assumption was kept to align with a previous cost-effectiveness analysis where treatment costs were assumed to be halted for ethical reasons [18].

Although, the results are not identical, they could be seen as supportive, because different approaches (models) and inputs (reconstructed data for functional milestones) produce comparable results as compared to a previously published analysis [18]. However, it may be worth exploring partitioned survival approaches in more detail if IPD were available to relax the sequential progression assumption, e.g., to account for censoring or apply this approach using UMSARS Part IV, because this approach seems to provide cost-effectiveness estimates favouring an intervention. At least, as a sensitivity analysis.

4.3 The importance of QALY shortfall multipliers

The QALY shortfall multiplier was included in this analysis for two reasons: first, MSA is unlikely to fulfil one of the routing criteria for highly specialised technology appraisal which is a point prevalence of 1:50,000 cases or less in England, when the prevalence for MSA is estimated at approximately 5 cases per 100,000 people in the UK [1, 44]. Other decision modifiers can potentially be applied to those rare diseases that are selected for highly specialised technology appraisal, but it depends on the size of benefit. Second, the severity calculated by both proportional and absolute QALY shortfalls indicated sufficient severity to allow for a QALY weight, (multiplier) of 1.2. It is potentially quite important for manufacturers to understand these relatively new methods, e.g., absolute shortfall (tend to favour younger patients) and proportional shortfall (tend to favour older patients). Also, it would be pertinent to know which HTA bodies apply these measures, e.g., Norway, Netherlands, and UK as this might influence cost-effectiveness [45]. For example, the deterministic analysis from our Markov model was reduced from £179,362 to £148,444 per QALY, for the ICER and weighted ICER, respectively.

4.4 Is it worth conducting EVPI analyses for rare diseases?

Pearson et al. (2018) and Grand et al. (2024) did not identify examples of the application of value of information analysis for economic modelling in rare disease nor did an umbrella review of economic evaluations in rare diseases, but both papers recommend that it is used to inform data collection priorities [14, 46]. The EVPI will indicate if it is worthwhile to fund and conduct further research, e.g., if the value of eliminating uncertainty is higher than the costs of collecting new evidence [15-17]. It was calculated at the individual level (per person) in this analysis, but the population EVPI is more useful, because it provides the upper limit for the value of a study aimed at reducing uncertainty across model parameters. However, it requires that the eligible treatment population is known including the period for which the decision problem is relevant [47].

Although these analyses may be more useful at the population level, they require additional considerations for rare diseases, especially during early cost-effectiveness modelling. First, a rare disease is often defined partly by its low prevalence, and multiplying individual-level EVPI with a small population will inherently contribute to lower EVPI than for more prevalent diseases [48]. Thus, minimising the likelihood that additional research will be worthwhile. On the other hand, uncertainty is most likely higher in cost-effectiveness models for rare diseases [14, 46, 49]. So, an important question remains unanswered, that is “which of these factors contribute most to such a reciprocal relationship?” Second, rare diseases tend to suffer from multiple phenotypes, delayed diagnosis or misdiagnosis, and fewer clinical experts to consult for advice, which can make it difficult to define the eligible population. Furthermore, defining a time horizon for the decision problem may be difficult due to missing natural history of disease data or unknown disease trajectories [14]. For these reasons, it may also be worth for researchers to consider if managed-entry agreements could play a role in reducing uncertainty for a specific rare disease prior to investing in additional studies solely based on EVPI analyses.

For increased robustness, it may be worthwhile also exploring expected value of partial perfect information (EVPPI) for a parameter or sets of parameters, which can help determine which parameters contribute most to uncertainty [28, 47]. For example, to further evaluate additional research versus managed-entry agreements, as the evidence generation plan may depend on which parameters are causing the uncertainty. Moreover, one can take additional steps and evaluate sample sizes or study designs if the costs of conducting research for, e.g., a set of parameters are lower than EVPPI. These methods are more complex because their methods depend on study design, model type and expertise, and they are called expected value of sample information (EVSI) and expected net benefit of sampling (ENBS). For example, these methods help determine study types and their optimal design to reduce uncertainty in decision-analytic models [47]. Generally, research on the implications and relevance of these advanced methods in rare diseases are warranted.

4.5 Limitations

The reconstruction of IPD was challenging. Obviously, working with IPD of sufficient quality rather than having to reconstruct data from published Kaplan-Meier curves with limited information, is highly preferable. Only the total number at risks were reported, and the reconstruction of IPD relies on the strong assumption that there is no censoring. The subject numbers at risk or number of events would vastly have improved the quality of digitisation, but they were not reported in original source paper for every event, except, indirectly for aid-required walking if assumed that the study population was event-free at study entry [13]. The only comment on censoring was that subjects (n = 40) undergoing tracheostomy (surgery) were treated as censored data [13].

The lack of data on censoring led to a likely overestimation of survival. For example, a sequential progression assumption was needed, e.g., death could not occur from all health states, which would be expected from MSA. For example, severe urinary tract infections in earlier disease states, which is one of the leading causes of death [50]. In terms of cost-effectiveness modelling this approach may also overestimate costs, because everyone needs to progress through all states, and accumulate more time in states that represent higher costs or lower quality of life, for example, bedridden. However, it was a necessary simplification given the reconstruction of IPD from available Kaplan-Meier curves, which nonetheless reflect the progressive and fatal nature of MSA. So even if functional milestones to inform partitioned survival models are preferable, much data would need to be generated to increase robustness of results. Future studies can increase their applicability or relevance for modelling by including utilities and cost data for UMSAR Part IV and functional milestones to further compare outcomes. However, there remain overwhelming precedence for using functional or disability rating scales to inform decision-analytic modelling in diseases similar to MSA, which means that UMSARS Part IV, as it stands, is currently the most suitable option to define model health states [18]

A digitisation approach that explicitly captures censoring (as a vector of censoring times) has been proposed for survival modelling, but this method was not applicable as it requires that censoring times are clearly marked and distinguishable on the curves, and similarly to the IPDfromKM package, it requires as a minimum, the total number at risk at baseline (time = 0) for all event curves [51]. However, this cost-effectiveness analysis nonetheless required merging of multiple Kaplan-Meier curves, as opposed to one event only, e.g., death, to reconstruct a study population. In addition, events were assigned to subjects by looping through subject id's for each of the event datasets, which meant that subjects with lower id numbers were consistently assigned the earlier events, and it resulted in them having faster disease progression. Altogether, these factors added to the complexity of reconstructing IPD for cost-effectiveness modelling. It would be interesting for researchers to further investigate digitisation of multiple event curves and its integration into a unified reconstructed dataset.

Much research is needed to further examine the relationship between functional milestones as opposed to disability or functional scales, e.g., UMSARS Part IV to inform modelling structures and strategies for MSA. Irrespective of this relationship, a major limitation of this study was that cost and utility data were missing for functional milestones, which meant that these health-economic parameters were sourced as surrogate data from other studies that stratified by UMSARS Part IV [5, 18, 32].

This analysis took a UK NHS perspective, but the time-to-event data were sourced from a Japanese population, which most likely have a different prevalence and distribution of phenotypes [13, 52]. In addition, it seemed that Winter and colleagues did not include social services costs in their cost categories (inpatient care, drugs, special equipment, home care, and other direct costs) stratified by UMSARS Part IV, which play a large part for end-of-life planning and caregivers. A UK NHS and personal social services would have been preferable [10-12, 32].

5. Conclusions

This analysis explored cost-effectiveness modelling for MSA using reconstructed data for time to functional milestones from a publication. Functional milestones may be modelled using partitioned survival modelling if progression of events is assumed sequential, but more research is needed to assess this model type for UMSARS Part IV. The Markov and partitioned survival models showed that results were comparable to a previously published Markov model informed by UMSARS Part IV despite differences in model structures. The partitioned survival model produced cost-effectiveness estimates that were more favourable to the hypothetical drug as opposed to the Markov model, but generally, there exists a paucity of health-economic data for functional milestones that make results highly uncertain. Lastly, more research is needed to establish the relevance of EVPI analysis in rare diseases, and this analysis made use of MSA as proxy to illustrate this point.

6. Statements and Declarations

6.1 Funding

Tobias Sydendal Grand is an industrial PhD candidate at the University of Sheffield sponsored by Innovation Fund Denmark (case number: 2040-00017B) and Lundbeck A/S.

6.2 Conflicts of interest

The authors with non-Lundbeck affiliation have no competing interests to declare that are relevant to the content of this article. The following authors were affiliated with Lundbeck A/S: *Redacted for now*. Lundbeck A/S aspire to develop transformative treatments in neuroscience, e.g., multiple system atrophy.

6.3 Availability of data and materials

The data sources are outlined in the manuscript.

6.4 Ethic approval

Not applicable

6.5 Consent to participate

Not applicable

6.6 Consent for publication

Not applicable

6.7 Code availability

The R project is available here: https://github.com/TobiasGrand/PSM-and-Markov-for-milestones-in-MSA. Note that captured Figures (png files) are not included in the R project due to copyright issues. The Markdown report is here: https://tobiasgrand.github.io/Report_PSM_Markov_MSA/.

6.8 Authors contribution

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Tobias Sydendal Grand. Daniel Oudin Åström reviewed the code. The first draft of the manuscript was written by Tobias Sydendal Grand, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

7. References

1. NHS Trust. Multiple System Atrophy Leeds Centre for Neurosciences; 2023. https://flipbooks.leedsth.nhs.uk/LN005500.pdf. Accessed 15 January 2025.

2. Krismer F, Fanciulli A, Meissner WG, Coon EA, Wenning GK. Multiple system atrophy: advances in pathophysiology, diagnosis, and treatment. Lancet Neurol. 2024;23(12):1252–66. https://doi.org/10.1016/s1474-4422(24)00396-x

3. Ben-Shlomo Y, Wenning GK, Tison F, Quinn NP. Survival of patients with pathologically proven multiple system atrophy. Neurology. 1997;48(2):384. https://doi.org/10.1212/WNL.48.2.384

4. Wenning GK, Shlomo YB, Magalhães M, Danie SE, Quinn NP. Clinical features and natural history of multiple system atrophy: An analysis of 100 cases. Brain. 1994;117(4):835–45. https://doi.org/10.1093/brain/117.4.835

5. Wenning GK, Geser F, Krismer F, Seppi K, Duerr S, Boesch S, et al. The natural history of multiple system atrophy: a prospective European cohort study. Lancet Neurol. 2013;12(3):264–74. https://doi.org/10.1016/s1474-4422(12)70327-7

6. Low PA, Reich SG, Jankovic J, Shults CW, Stern MB, Novak P, et al. Natural history of multiple system atrophy in the USA: a prospective cohort study. The Lancet Neurology. 2015;14(7):710–9. https://doi.org/10.1016/S1474-4422(15)00058-7

7. Mészáros L, Hoffmann A, Wihan J, Winkler J. Current Symptomatic and Disease-Modifying Treatments in Multiple System Atrophy. Int J Mol Sci. 2020;21(8). https://doi.org/10.3390/ijms21082775

8. Krismer F, Wenning GK. Multiple system atrophy: insights into a rare and debilitating movement disorder. Nat Rev Neurol. 2017;13(4):232–43. https://doi.org/10.1038/nrneurol.2017.26

9. Gilman S, Wenning GK, Low PA, Brooks DJ, Mathias CJ, Trojanowski JQ, et al. Second consensus statement on the diagnosis of multiple system atrophy. Neurology. 2008;71(9):670. https://doi.org/10.1212/01.wnl.0000324625.00404.15

10. Burns MR, McFarland NR. Current Management and Emerging Therapies in Multiple System Atrophy. Neurotherapeutics. 2020;17(4):1582–602. https://doi.org/10.1007/s13311-020-00890-x

11. Flabeau O, Meissner WG, Tison F. Multiple system atrophy: current and future approaches to management. Ther Adv Neurol Disord. 2010;3(4):249–63. https://doi.org/10.1177/1756285610375328

12. Grossauer A, Sidoroff V, Heim B, Seppi K. Symptomatic Care in Multiple System Atrophy: State of the Art. The Cerebellum. 2022. https://doi.org/10.1007/s12311-022-01411-6

13. Watanabe H, Saito Y, Terao S, Ando T, Kachi T, Mukai E, et al. Progression and prognosis in multiple system atrophy: An analysis of 230 Japanese patients. Brain. 2002;125(5):1070–83. https://doi.org/10.1093/brain/awf117

14. Grand TS, Ren S, Hall J, Åström DO, Regnier S, Thokala P. Issues, Challenges and Opportunities for Economic Evaluations of Orphan Drugs in Rare Diseases: An Umbrella Review. Pharmacoeconomics. 2024. https://doi.org/10.1007/s40273-024-01370-2

15. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. Value in Health. 2012;15(6):835–42. https://doi.org/10.1016/j.jval.2012.04.014

16. Barton GR, Briggs AH, Fenwick EA. Optimal cost-effectiveness decisions: the role of the cost-effectiveness acceptability curve (CEAC), the cost-effectiveness acceptability frontier (CEAF), and the expected value of perfection information (EVPI). Value Health. 2008;11(5):886–97. https://doi.org/10.1111/j.1524-4733.2008.00358.x

17. Kunst N, Burger EA, Coupé VMH, Kuntz KM, Aas E. A Guide to an Iterative Approach to Model-Based Decision Making in Health and Medicine: An Iterative Decision-Making Framework. Pharmacoeconomics. 2023. https://doi.org/10.1007/s40273-023-01341-z

18. Grand TS, Ren S, Thokala P, Zanigni S, Åström DO, Bidani A, et al. Cost-effectiveness modelling of multiple system atrophy for early health technology assessment. PharmacoEconomics Open. 2025. https://doi.org/10.1007/s41669-025-00609-8

19. Bouvy J. Should NICE's cost-effectiveness thresholds change? 2024. https://www.nice.org.uk/news/blogs/should-nice-s-cost-effectiveness-thresholds-change-. Accessed 9 January 2024.

20. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual 2022. https://www.nice.org.uk/process/pmg36/chapter/economic-evaluation-2. Accessed 23 January 2025.

21. Allaire J, Xie Y, Dervieux C, McPherson J, Luraschi J, Ushey K, et al. Rmarkdown: Dynamic documents for r 2023. https://github.com/rstudio/rmarkdown. Accessed 3 July 2025.

22. R Core Team. R: A language and environment for statistical computing: R Foundation; 2021. https://www.r-project.org/. Accessed 1 March 2023.

23. Gilman S, Low PA, Quinn N, Albanese A, Ben-Shlomo Y, Fowler CJ, et al. Consensus statement on the diagnosis of multiple system atrophy. J Neurol Sci. 1999;163(1):94–8. https://doi.org/10.1016/s0022-510x(98)00304-9

24. Guyot P, Ades AE, Ouwens MJNM, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Medical Research Methodology. 2012;12(1):9. https://doi.org/10.1186/1471-2288-12-9

25. Lui N, Zhou Y, Lee J. IPDfromKM: reconstruct individual patient data from published Kaplan-Meier survival curves. 2021;2021. https://doi.org/10.1186/s12874-021-01308-8

26. Jackson C. Multi-State Models for Panel Data: The msm Package for R. Journal of Statistical Software. 2011;38(8):1 – 28. https://doi.org/10.18637/jss.v038.i08

27. Srivastava T, Latimer NR, Tappenden P. Estimation of Transition Probabilities for State-Transition Models: A Review of NICE Appraisals. Pharmacoeconomics. 2021;39(8):869–78. https://doi.org/10.1007/s40273-021-01034-5

28. Briggs A, Sculpher M, Claxton K. Decision Modelling for Health Economic Evaluation: Oxford University Press; 2006.

29. Incerti D, Jansen J. hesim: Health Economic Simulation Modeling and Decision Analysis. 2021. https://arxiv.org/abs/2102.09437. Accessed 3 July 2025.

30. Incerti D, Jansen JP, Clements M. Partitioned survival models 2025. https://hesim-dev.github.io/hesim/articles/psm.html. Accessed 15 Jan 2025.

31. Williams C, Lewsey JD, Mackay DF, Briggs AH. Estimation of Survival Probabilities for Use in Cost-effectiveness Analyses: A Comparison of a Multi-state Modeling Survival Analysis Approach with Partitioned Survival and Markov Decision-Analytic Modeling. Med Decis Making. 2017;37(4):427–39. https://doi.org/10.1177/0272989x16670617

32. Winter Y, Stamelou M, Cabanel N, Sixel-Döring F, Eggert K, Höglinger GU, et al. Cost-of-illness in multiple system atrophy and progressive supranuclear palsy. J Neurol. 2011;258(10):1827–34. https://doi.org/10.1007/s00415-011-6028-7

33. McCrone P, Payan CA, Knapp M, Ludolph A, Agid Y, Leigh PN, et al. The economic costs of progressive supranuclear palsy and multiple system atrophy in France, Germany and the United Kingdom. PLoS One. 2011;6(9):e24369. https://doi.org/10.1371/journal.pone.0024369

34. Kaplan S, Xie H, Kumar A. Diagnostic symptoms, treatment patterns, and healthcare resource utilization in multiple system atrophy in the United States. Neurodegenerative Disease Management. 2025:1–10. https://doi.org/10.1080/17582024.2025.2554490

35. Wenning GK, Tison F, Seppi K, Sampaio C, Diem A, Yekhlef F, et al. Development and validation of the Unified Multiple System Atrophy Rating Scale (UMSARS). Mov Disord. 2004;19(12):1391–402. https://doi.org/10.1002/mds.20255

36. Xiao Y, Zhang L, Wei Q, Ou R, Hou Y, Liu K, et al. Health-related quality of life in patients with multiple system atrophy using the EQ-5D-5L. Brain and Behavior. 2022;12(10):e2774. https://doi.org/10.1002/brb3.2774

37. Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V. Catalogue of EQ-5D scores for the United Kingdom. Med Decis Making. 2011;31(6):800–4. https://doi.org/10.1177/0272989x11401031

38. Hvidberg MF, Alava MH. Catalogues of EQ-5D-3L Health-Related Quality of Life Scores for 199 Chronic Conditions and Health Risks for Use in the UK and the USA. Pharmacoeconomics. 2023;41(10):1287–388. https://doi.org/10.1007/s40273-023-01285-4

39. McNamara S, Schneider PP, Love-Koh J, Doran T, Gutacker N. Quality-Adjusted Life Expectancy Norms for the English Population. Value in Health. 2023;26(2):163–9. https://doi.org/10.1016/j.jval.2022.07.005

40. Wailoo A. A guide to calculating burden of illness for NICE Evaluations: Technical Support Document 23 2024. https://sheffield.ac.uk/nice-dsu/tsds/full-list. Accessed August 25 2025.

41. Schneider P, McNamara S, Love-Koh J, Doran T, Gutacker N. QALY Shortfall Calculator 2021. https://shiny.york.ac.uk/shortfall. Accessed 1 September 2025.

42. Saulnier T, Fabbri M, Le Goff M, Helmer C, Pavy-Le Traon A, Meissner WG, et al. Patient-perceived progression in multiple system atrophy: natural history of quality of life. Journal of neurology, neurosurgery, and psychiatry. 2024. https://doi.org/10.1136/jnnp-2023-332733

43. Chancellor JV, Hill AM, Sabin CA, Simpson KN, Youle M. Modelling the Cost Effectiveness of Lamivudine/Zidovudine Combination Therapy in HIV Infection. Pharmacoeconomics. 1997;12(1):54–66. https://doi.org/10.2165/00019053-199712010-00006

44. National Institute for Health and Care Excellence. NICE process and methods. Appendix 1: highly specialised technologies 2024. https://www.nice.org.uk/process/pmg46/resources/highly-specialised-technologies-nice-prioritisation-board-routing-criteria-pdf-19830673472965. Accessed 28 August 2025.

45. Ali S, Haines A, Lee K. Severity Weighting of Quality-Adjusted Life-Years for Economic Evaluation: What, How, and Where Next? Canadian Journal of Health Technologies. https://doi.org/10.51731/cjht.25.1063

46. Pearson I, Rothwell B, Olaye A, Knight C. Economic Modeling Considerations for Rare Diseases. Value Health. 2018;21(5):515–24. https://doi.org/10.1016/j.jval.2018.02.008

47. Heath A, Kunst N, Jackson C. Value of Information for Healthcare Decision-Making: CRC Press; 2024.

48. Richter T, Nestler-Parr S, Babela R, Khan ZM, Tesoro T, Molsen E, et al. Rare Disease Terminology and Definitions - A Systematic Global Review: Report of the ISPOR Rare Disease Special Interest Group. Value Health. 2015;18(6):906–14. https://doi.org/10.1016/j.jval.2015.05.008

49. Nestler-Parr S, Korchagina D, Toumi M, Pashos CL, Blanchette C, Molsen E, et al. Challenges in Research and Health Technology Assessment of Rare Disease Technologies: Report of the ISPOR Rare Disease Special Interest Group. Value Health. 2018;21(5):493–500. https://doi.org/10.1016/j.jval.2018.03.004

50. Papapetropoulos S, Tuchman A, Laufer D, Papatsoris AG, Papapetropoulos N, Mash DC. Causes of death in multiple system atrophy. Journal of Neurology, Neurosurgery & Psychiatry. 2007;78(3):327. https://doi.org/10.1136/jnnp.2006.103929

51. Rogula B, Lozano-Ortega G, Johnston KM. A Method for Reconstructing Individual Patient Data From Kaplan-Meier Survival Curves That Incorporate Marked Censoring Times. MDM Policy & Practice. 2022;7(1):23814683221077643. https://doi.org/10.1177/23814683221077643

52. Kaplan S. Prevalence of multiple system atrophy: A literature review. Revue Neurologique. 2024. https://doi.org/10.1016/j.neurol.2023.11.013

V Appendix (Use of GenAI declaration)

Acknowledge, Describe, Evidence template

Please include a completed version of this template as an appendix to any submitted work that has involved the use of generative AI (GenAI) technology (further information on GenAI and assessment).

Acknowledge

I acknowledge the use of generative AI and for the following purposes:

to generate materials for background research and self-study in the drafting of this assessment.

to generate materials that were included within my final assessment in modified form.

Describe

Please provide a short summary of how you used generative AI in your assignment. You may wish to include the following information:

What prompts did you use?

What outputs did you generate?

How did you use/adapt/develop the outputs?

Summary: First, I want to state that generative AI has improved significantly over the period where I have conducted my PhD. Initially, it was almost useless, but now it is truly helpful. However, it is not flawless. The art is knowing how and what to use it for, and what NOT to use it for. Thus, I appreciate the approach taken by the University of Sheffield, which is a positive approach, e.g., providing all students with an opportunity to develop their AI literacy during their time at the University (https://sheffield.ac.uk/saas/elevate/ai-principles). For these reasons, I have only used it for purposes that could not affect the results of my research. However, I have used GenAI to improve the presentation of my research, e.g., illustrations, to increase confidence in my research, e.g., debugging, and lastly, when I could not sufficiently do something, e.g., steps for inserting a pdf into a Word document.

GenAI has not been used for writing – only in very few instances to paraphrase very important sentences, where I wanted to check if the wording and meaning made sense to others, and if synonyms had the meaning, I thought they did. I am not a native speaker, but I know what I would like to express in writing and GenAI can check this for me.

To conclude, ChatGPT, Claude and Copilot were used to:

Improve the presentation of research

Examples:

Changing colours in ggplot to reflect disease severity, e.g., green, light green, orange, red.

Changing from a $ sign to £ sign in ggplot.

Troubleshooting

Examples

Code is not running, and you already spent hours looking for the error. Then a prompt could be, why is the code not running, and usually it has to do with a typo. Very simple, but GenAI is much faster at identifying typos if code is lengthy, e.g., over thousand lines of code.

GenAI has been used to refine existing code for better performance and readability, e.g., in the Shiny app, I asked why the model was running so slow, which occurred after I made two wrapping functions. It helped a bit to move the deterministic analysis outside the Shiny Server loop (action button), which was a GenAI suggestion.

GenAI has been useful for debugging, e.g., optimising debugging statements by turning them into loops, which is more elegant and takes up less coding space. For example, checking that transition probabilities sum to one.

Asking what should be included in the thesis summary? It is guidance that is helpful for providing a first draft for later review by supervisors. The alternative is to constantly disturbing them with questions.

Advise on processes

Examples

My first step is always looking online, e.g., watch YouTube videos on how to upload files to Github (things are always difficult the first time), but next step could be to ask GenAI.

How to insert pdfs into word without losing the intended formatting.

Evidence

Please provide evidence of the outputs that you generated by copying and pasting below or by providing a screenshot.

Improve the presentation of research

Generative AI system: ChatGPT 5

 # 1 Prompt: I have a thesis title "issues, challenges, and opportunities for economic evaluation in rare diseases" and three independent projects that serve to answer or inform the title: Issues, challenges and opportunities for economic evaluations of orphan drugs in rare diseases: an umbrella review, Cost-effectiveness modelling of multiple system atrophy for early HTA, Conceptual modelling of functional milestones in multiple system atrophy: comparing Markov and partitioned survival models using digitised time-to-event data. I would like to create an image that shows the relationship between them. What information do you need from me?

 # 1 Output: Perfect — that's a great foundation for a clear visual! To design an image that accurately shows the relationships between your thesis and its three projects, I'll need a few more details so the structure, flow, and emphasis are right. Here's what would help:

1. Relationship Details

How do the three projects connect to the overall thesis?

For example, do they each address different aspects of "issues, challenges, and opportunities," or do they build sequentially (e.g., literature review → modelling study → conceptual exploration)?

Is there a specific order or hierarchy (e.g., one informs the others, or they are parallel and equally important)?

2. Conceptual Emphasis

Would you like the image to highlight methodological relationships (e.g., how each method contributes to understanding economic evaluation) or thematic relationships (issues → challenges → opportunities)?

Should it focus on research flow (inputs → analyses → outputs) or on conceptual overlap (shared themes across projects)?

3. Visual Style

What kind of figure are you envisioning?

Examples: flowchart, concept map, Venn diagram, layered pyramid, radial (hub-and-spoke) layout, etc.

Should it look academic and schematic (e.g., figure for a thesis document) or more illustrative and communicative (e.g., presentation or poster)?

4. Any Key Elements or Outputs

For each project, what is its main purpose, outcome, or contribution?

e.g., “Umbrella review → identifies methodological gaps,” “Cost-effectiveness model → tests early HTA approach,” etc.

Are there any shared datasets, frameworks, or methods across projects that you’d like visualised?

Once I have that, I can design a clear conceptual diagram — e.g., showing your thesis as the central theme with the three projects positioned around it, connected by arrows or labelled pathways.

Would you like the final image to be simple and publication-ready (e.g., one clean diagram for your thesis document), or a more detailed version (e.g., for a presentation explaining how the pieces fit)?

 # 2 Prompt: Before we proceed. Please provide examples of other layouts than the Flow diagram, but style is academic and minimalist, fine with bullets as long as they can be changed

 # 2 Output: Excellent — since you want to explore alternative layouts (but keep the academic and minimalist look, with editable bullet points), here are three strong options besides the Layered Flow Diagram. Each of these could be provided as an editable PowerPoint layout with simple boxes, lines, and text placeholders.

---

1. Radial “Hub-and-Spoke” Layout

Purpose: Emphasizes that the three projects independently contribute to the central thesis, rather than forming a sequential process.

Structure:

[Project 1]

|

[Project 2] ---- [Thesis Title] ---- [Project 3]

|

[Integrated Insights]

Design notes:

Thesis title in the center (main box).

Three surrounding boxes connected with simple straight lines (like spokes).

Optional lower box for “Integrated insights / synthesis.”

Clean academic look (e.g., thin black lines, light grey/blue fills).

Works very well if all three studies are distinct but conceptually linked.

---

2. Venn / Thematic Overlap Diagram

Purpose: Shows conceptual overlaps — e.g., all projects address “data limitations” or “feasibility,” but in different ways.

Structure:

(Umbrella Review)

∩

(Early HTA)

∩

(Conceptual Modelling)

Overlapping regions can have placeholders like:

“Data uncertainty”

“Rare disease modelling challenges”

“Methodological innovation”

Central overlap labeled with your thesis title or core question.

Design notes:

Minimalist circles with semi-transparent fills.

Useful for conceptual or discussion sections of the thesis.

---

3. Concept Map / Hierarchical Tree

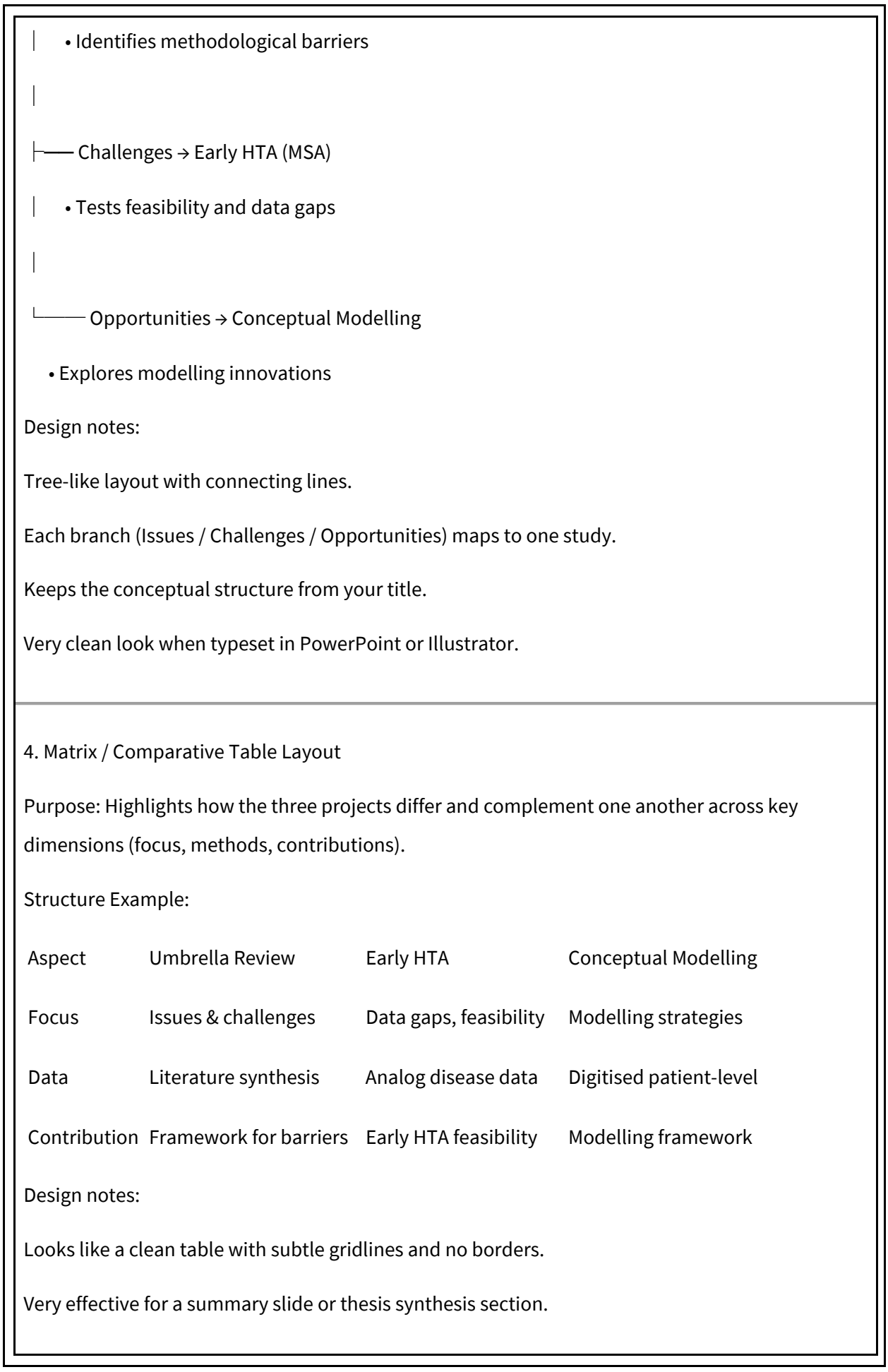
Purpose: Ideal if you want to show how the thesis breaks into key questions and how each project answers specific sub-questions.

Structure:

Thesis: Issues, Challenges, and Opportunities

|

├── Issues → Umbrella Review



Design notes:

Tree-like layout with connecting lines.

Each branch (Issues / Challenges / Opportunities) maps to one study.

Keeps the conceptual structure from your title.

Very clean look when typeset in PowerPoint or Illustrator.

---

4. Matrix / Comparative Table Layout

Purpose: Highlights how the three projects differ and complement one another across key dimensions (focus, methods, contributions).

Structure Example:

Aspect	Umbrella Review	Early HTA	Conceptual Modelling
Focus	Issues & challenges	Data gaps, feasibility	Modelling strategies
Data	Literature synthesis	Analog disease data	Digitised patient-level
Contribution	Framework for barriers	Early HTA feasibility	Modelling framework

Design notes:

Looks like a clean table with subtle gridlines and no borders.

Very effective for a summary slide or thesis synthesis section.

Recommendation Summary

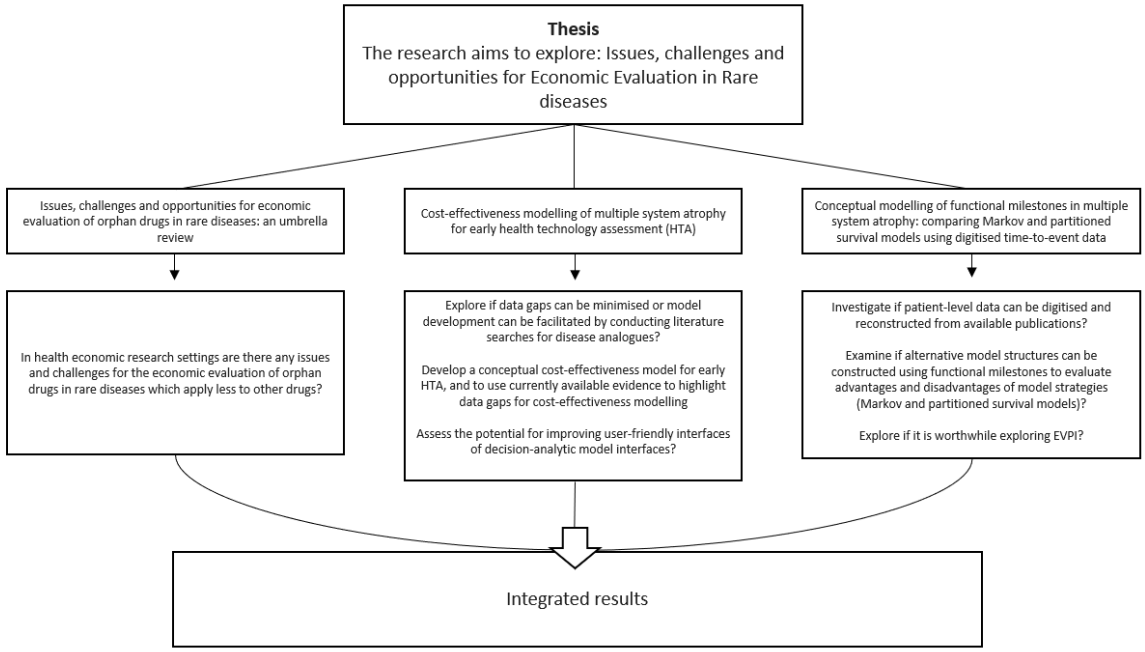
Layout	Best for	Visual tone	Strength
Layered Flow	Showing research progression	Hierarchical	Sequential clarity
Radial Hub	Equal contributions	Balanced	Emphasizes unity
Venn / Overlap	Conceptual relationships	Interpretive	Shows shared ideas
Tree / Concept Map	Linking to thesis title categories	Hierarchical	Thematic clarity
Matrix	Comparing projects	Structured	Summary view

Would you like me to prepare 2–3 of these layouts (editable PowerPoint format) so you can test which fits best visually?
If yes, please tell me which two you'd like (e.g., "Hub-and-Spoke" + "Concept Map"), and I'll create them in an academic minimalist style with editable bullets.

 # Interim Action: I now had the idea to create a Venn diagram and layered flow diagram, which I did in PowerPoint, because this is something that ChatGPT cannot do. It will gladly try, but the output looks horrible. However, after I made my image in PowerPoint, I wanted additional help to improve the visual expression.

 # 3 Prompt:

I now did this in PowerPoint any design suggestions for improvement?



 # 3 Output: Fantastic work — this version already looks clear, logical, and academically professional. The layout is well-balanced and your text hierarchy is strong.
Here are a few targeted design refinements that will make it look even cleaner and more publication-ready for a thesis or presentation:

🎨 1. Visual Hierarchy & Alignment

✅ Keep

The overall structure and flow (top → middle → bottom) is perfect.

Box spacing and arrow connections are consistent.

🔧 Improve

Vertically align the top three middle boxes more precisely (

 # 4 Updated Prompt (interrupted its thinking)

Other things?

 # 4 Output

Excellent — since your layout is already solid and communicates the logic clearly, you're now at the stage of polishing visual clarity, academic tone, and readability. Here's a detailed set of design improvements to take it from "good" to "publication-quality":

🎯 1. Visual Hierarchy (Make Key Ideas Pop)

Goal: Help readers instantly grasp structure and relationships.

Bold the box titles (e.g., "Issues, challenges and opportunities…", "Cost-effectiveness modelling…") but keep the question text in regular weight.

Slightly enlarge the thesis box font (e.g., 13–14 pt vs. 12 pt for the others).

Consider adding a subtle grey fill (5–10 % tint) to the thesis box and the integrated results box to visually "bookend" the figure (top and bottom).

📏 2. Symmetry and Spacing

Goal: Make the layout feel perfectly balanced.

Horizontally align all boxes using PowerPoint's "Align Middle" tool (especially the three middle-level boxes).

Ensure equal vertical spacing between each layer (thesis → middle → question → integrated results).

Increase the arrow length slightly from the middle to lower boxes — this creates better air and readability.

🧭 3. Arrow Consistency

Goal: Maintain visual rhythm and readability.

Use uniform arrow thickness (1–1.25 pt) across all connectors.

Consider curved connectors for the bottom ones (as in your earlier image), which feel more natural and help guide the eye inward.

Use straight, non-overlapping arrows elsewhere to emphasize logical flow.

🔠 4. Typography & Text Layout

Goal: Improve readability and reduce visual clutter.

Keep Calibri 12 pt but increase line spacing slightly (1.1–1.15) for readability.

Use sentence case (not title case) inside boxes for consistency with academic writing.

Add subtle indentation for bullet points to create a clearer text hierarchy.

If text in a box feels dense, slightly widen the box instead of reducing font size — clarity over compactness.

🎨 5. Subtle Visual Styling

Goal: Professional polish without distracting color.

Keep the overall monochrome scheme, but you can introduce light differentiation:

Thesis and results boxes: pale grey background (#F2F2F2).

Project boxes: white background, black border.

Add a thin shadow or 1 pt darker border to the thesis and integrated boxes to frame them.

📚 6. Optional Academic Touches

If this figure goes into your thesis manuscript:

Add a figure title beneath:
“Figure X. Relationship between thesis components and sub-projects addressing issues, challenges, and opportunities for economic evaluation in rare diseases.”

Keep it on one page, no page breaks or text wrapping.

Ensure alt-text or a short caption description if submitting digitally (for accessibility).

✅ Example of the Final Visual Logic Flow

Thesis (grey highlight)

|

├── Issues → Umbrella Review → Research Question

├── Challenges → Early HTA (MSA) → Research Questions

└──── Opportunities → Conceptual Modelling (MSA) → Research Questions

|

Integrated Results (grey highlight)

Would you like me to suggest an exact colour palette and font sizing scheme (with RGB codes and sizes) so it stays consistent with your thesis formatting guidelines?

 # 5 Prompt

Yes - please suggest RGB codes and for which boxes

 # 5 Output

Excellent — let's give your figure a clean, professional academic palette that enhances readability while staying subtle and printable in grayscale.

Here's a tested combination used in many health economics theses and HTA reports:

🎨 Recommended Colour Scheme (with RGB values)

Element	Fill Colour	Border Colour	Font Colour	Notes
Thesis Box	RGB(230, 230, 230) → Light grey	RGB(0, 0, 0)	RGB(0, 0, 0)	Gives subtle emphasis without dominating; top-level focus.
Middle Project Boxes	RGB(255, 255, 255) → White	RGB(0, 0, 0)	RGB(0, 0, 0)	Keeps neutrality; separates nice against the light-grey backgrou

Lower Question Boxes	RGB(247, 247, 247) → Very pale grey	RGB(0, 0, 0)	RGB(0, 0, 0)	Slightly differentiated from project boxes; helps visual grouping.
Integrated Results Box	RGB(230, 230, 230) → Same as Thesis Box	RGB(0, 0, 0)	RGB(0, 0, 0)	Creates visual symmetry (top and bottom "anchor" boxes).
Background of Slide	RGB(242, 242, 242) → Light neutral grey	—	—	Gives a subtle contrast to white boxes; easy on the eyes.
Arrows / Lines	RGB(0, 0, 0)	—	—	Consistent black for clarity and print reliability.

🧩 Font & Text Styling

Element	Font	Size	Style	Alignment
Thesis title	Calibri	13–14 pt	Bold	Center
Project titles	Calibri	12 pt	Bold	Center
Question text	Calibri	11–12 pt	Regular	Left-aligned within box
Integrated results	Calibri	12 pt	Regular (first line bold optional)	Center

Use 1.15 line spacing and avoid italics unless emphasizing technical terms (e.g., *EVPI*, *HTA*).

🖍️ Extra Refinement Tips

Add 3–4 pt padding inside each text box (Format Shape → Text Options → Internal Margins) for breathing space.

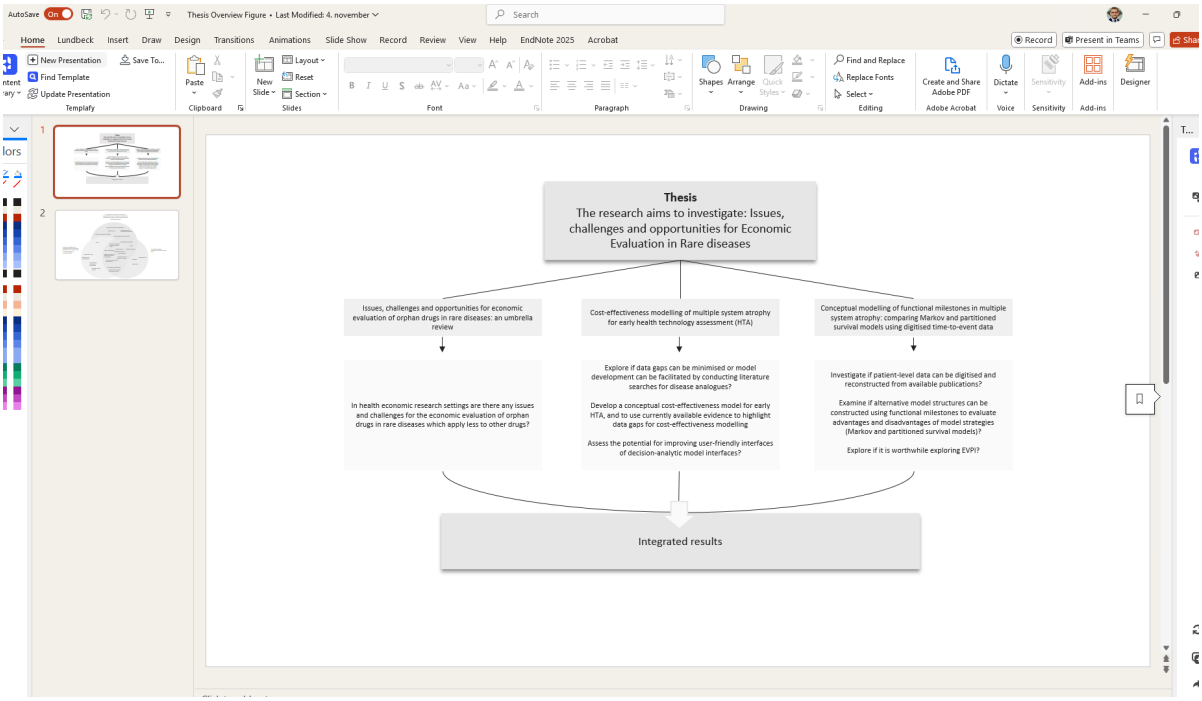
Keep arrow heads small (Size ≈ 4 pt width) and consistent.

To print cleanly in black and white, ensure grey tones differ by at least 15 RGB points between levels.

Would you like me to suggest a matching academic colour variant (for example, subtle blue-grey or teal-grey tones) in case you later want a coloured version for a presentation?

 # Results and interpretation: Based on this guidance from ChatGPT I ended up with this Figure that I created in PowerPoint and a visual design that I could implement across Figures in the Thesis. I did the work myself but was guided by GenAI, especially in terms of the figures and colouring. Also, I aligned the middle box as suggested.

The interpretation is that the prompting reflects the guidelines in section 3. Dialogic prompting (https://sheffield.ac.uk/study-skills/digital/generative-ai/critical), because here we dig deeper and produce a nuanced output that ensure a better understanding of the concept (PhD) by organizing the various chapters in the PhD as an illustration. In this way – readers can obtain a quick overview, but the drawing was made manually but guided by GenAI.



Improve the presentation of research (second example)

Generative AI system: ChatGPT – 4 (older version)

 # 1 Prompt: I have created a plot with ggplot using outputs from the hesim package, but I cannot find information online on how to change from $ sign to £ sign in the plot. Can you help me change the currency from dollars to pounds in my cost-effectiveness plot?

 # 1 Output:

Create a custom formatting function:

```
format_dollar <- function(x) {
  paste0("£ ", formatC(x, format = "d", big.mark = ","))
}
```

And then using it in the scale functions:

```
scale_y_continuous(labels = format_dollar)
```

This function:

Adds the £ symbol prefix

Formats as integer (format = "d")

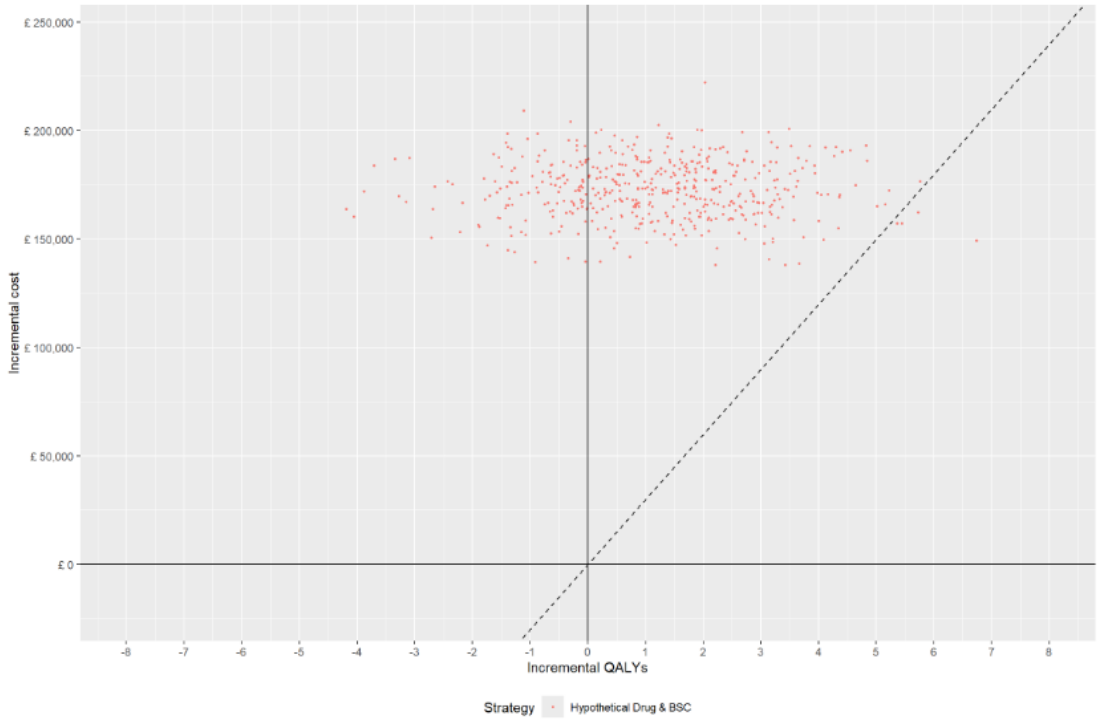
Adds comma thousand separators (big.mark = ",")

 # Results and interpretation: Based on this tiny update to my ggplot I managed to overwrite the default $ sign from the hesim package. I remember spending countless hours on websites to learn this but found nothing. This was in my early days of coding.

The interpretation here is that there is no conflict with guidelines as we are dealing with a minor change. The prompt bank highlights accuracy as an issue to be aware about (https://sheffield.ac.uk/study-skills/digital/generative-ai/prompt-bank). This was very easy to assess here, as the only expected change in the expected output was a change from $ to £. Indeed, this was the only change observed.

```r
  format_dollar <- function(x) {
    paste0("£ ", formatC(x, format = "d", big.mark = ","))
  }

ylim <- max(cea_pw_out$delta[, ic]) * 1.1
xlim <- ceiling(max(cea_pw_out$delta[, ie]) * 1.1)
ggplot(cea_pw_out$delta,
       aes(x = ie, y = ic, col = strategy_factor(strategy_id))) +
  geom_jitter(size = .5)  +
  xlab("Incremental QALYs") +
  ylab("Incremental cost") +
  scale_y_continuous(limits = c(-22000, ylim), #here you can -ylim to 0
                     labels = format_dollar) +
  scale_x_continuous(limits = c(-xlim, xlim), breaks = seq(-10, 10, 1)) +
  theme(legend.position = "bottom") +
  scale_colour_discrete(name = "Strategy") +
  geom_abline(slope = 30000, linetype = "dashed") +
  geom_hline(yintercept = 0) +
  geom_vline(xintercept = 0)
```



Troubleshooting

Generative AI system: Claude (older version)

 # 1 Prompt: My Shiny app has become slower after I constructed two wrapper functions. Can we optimise the performance so it loads faster?

 # 1 Output: I can help optimize your Shiny app's performance! Let me analyze your code structure.

I see the issue - you're running both wrapper functions reactively, which is causing unnecessary recalculations. Here's what's happening:

Current Problem:

f_wrapper_det() runs deterministic analysis with fixed parameters

f_wrapper_psa() runs probabilistic sensitivity analysis with user-defined parameters

Both are being called inside your reactive context, meaning the deterministic analysis recalculates unnecessarily

Solution:

Since your deterministic results never change (they use base-case parameters from the publication), you should run f_wrapper_det() once when the app loads, not reactively.

Here's what to do:

Move the deterministic analysis outside the server function (before the UI definition):

############################ Load the data ###################################deterministic_results <- f_wrapper_det() # Generates deterministic results once at app startup########## ################ Define UI ######################################

Keep only PSA inside the observeEvent in your server function:

observeEvent(input$run_model, { # Only run PSA here - deterministic_results already exists results <- f_wrapper_psa(n_sim = input$SI_n_sim, ...)})

Access deterministic results directly in your render functions:

output$SO_icer_Table_des <- renderUI({ results_format <- deterministic_results$results_det # Uses pre-computed results # ... rest of your code})

Why this works:

Deterministic analysis runs once at startup (~2-3 seconds)

PSA only runs when user clicks "Run Model" button

No redundant calculations

~50% faster app loading and interaction

This is a common Shiny optimization pattern: separate static computations (run once) from dynamic computations (run on user input).

Would you like me to help implement these changes?

 # 1 Results and interpretation: The deterministic analysis was moved outside the "action button" or observeEvent function. GenAI failed to consider that the dynamic computations always run upon loading the Shiny App, but it was useful advice. It improved the speed of the model slightly – it was still slower than before we created two wrappers instead of having only one, but two wrappers were suggested by my supervisors to make sure the deterministic and probabilistic analysis were passed through the same model code f_MM_MSA.R. This was regarded best practice but was massive undertaking because no guidance or tutorials on this existed, but I adapted and made it work.

Our interpretation here is that this is an example of a PRO. GenAI can help you identify areas for further investigation, which usually happen to be areas of improvement as it was in this case (https://sheffield.ac.uk/study-skills/digital/generative-ai/critical). Although, the code example was provided, it was unnecessary. Just the guidance to move the code would have been enough for me to do it.

Troubleshooting (second example)

Generative AI system: ChatGPT 5

 # 1 Prompt: This is the structure of the discussion. How much should be included for the summary of thesis? Discussion • Summary of the thesis • Novel contributions of your PhD • Umbrella review on rare diseases • Early HTA Modelling for MSA • Flexible RShiny Model/living HTA • Considering data/structural uncertainty (Markov vs PSA) • Implications of your work • Recommendations for future analysts/modellers • Nomenclature • Early HTA • Structured elicitation? • Limitations of your work • List all the drawbacks of your work • What is missing from your PhD? • What aspects are uncertain? • Future work • What will you do if you had 3 more years? • Conclusions

 # 1 Output:

That's a very solid and comprehensive Discussion chapter structure — you're clearly aiming for a full integrative synthesis, which is exactly what examiners want in a PhD built around multiple studies.

Your question — *how much should be included for the "Summary of the thesis"* — is an important one. The key is to set the tone and orient the reader before going into the deep analytical and reflective parts.

Here's a detailed guideline for what to include and how much space to give it 👇

🎯 Purpose of the "Summary of the Thesis" Section

The *Summary of the Thesis* is not a rehash of the abstract or results, but a bird's-eye overview that ties together the logic, flow, and outcomes of your PhD.
Think of it as the narrative bridge between your empirical chapters (6–8) and your interpretive discussion.

It should:

Briefly restate the overarching aim and research questions

Summarize how each project contributes to those aims

Highlight the logical progression between studies (umbrella → model → comparative)

Mention key findings at a high level, without diving into detailed results

Transition into the "Novel contributions" section smoothly

🧩 Recommended Length and Depth

Thesis length	Suggested length for "Summary of the thesis"	Approx. word cou
60,000–80,000 words (typical PhD)	2–4 pages (5–7% of the Discussion chapter)	1,000–1,500 word

That's usually 1.5–2 single-spaced pages, or roughly 5–7 paragraphs.

✍️ Suggested Structure for "Summary of the Thesis"

Here's a clean and defensible outline you can follow:

1. Restate the Aim and Research Questions (≈1 paragraph)

Reiterate what your PhD set out to investigate and why it matters (e.g., challenges in health-economic evaluation of orphan drugs and early HTA in rare diseases).

2. Describe the Overall Design and Flow (≈1 paragraph)

Explain how the three projects build upon one another — e.g.,

Project 1 (Umbrella review) maps the field and identifies gaps;

Project 2 (Early HTA model) operationalizes a case study (MSA);

Project 3 (Comparative modelling) explores structural and data uncertainty.

3. Summarize Each Project's Core Findings (≈3 short paragraphs)

One paragraph per study, focusing on the main insight and how it advances your overarching aim. Keep it synthetic — e.g.:

What did the umbrella review reveal about methodological gaps?

What did the early HTA model demonstrate in terms of feasibility?

What did the comparative modelling contribute about uncertainty or structure?

4. Integrative Summary (≈1 paragraph)

Bring the three strands together:

"Taken together, these studies demonstrate that early HTA in rare diseases can be strengthened by transparent, flexible models and by systematic approaches to handling uncertainty and data limitations."

5. Transition Sentence

End with a short bridge leading to the next section:

"The following sections discuss the novel contributions of this work, its implications for health-economic analysis in rare diseases, and directions for future research."

💡 Practical Tips

Avoid redundancy: Don't restate all results — just the "so what."

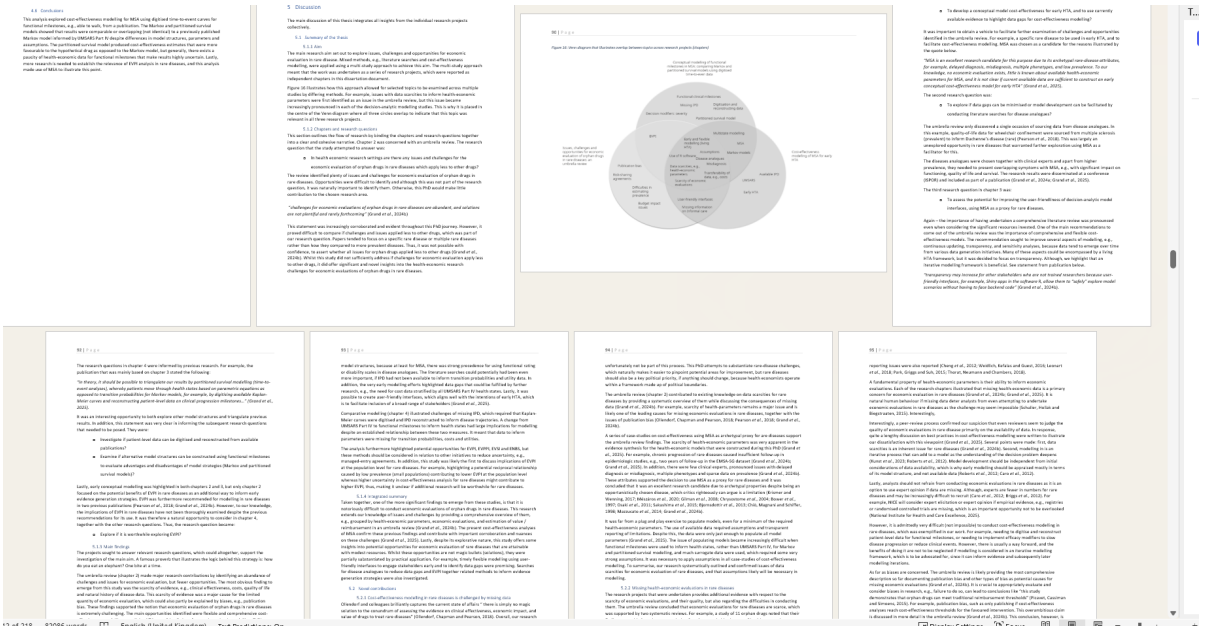
Be synthetic: Focus on how the three studies *connect* rather than describing them separately.

Think examiners: This is often the first part they reread before your viva; make it crisp, readable, and logically structured.

Would you like me to draft a model version (≈500 words) of what your *"Summary of the Thesis"* could look like, using your actual studies (umbrella review, early HTA modelling for MSA, Markov vs PSA comparison)? It could serve as a template for your own writing.

 # 1 Results and interpretation: It may be difficult to see in the picture below, but I tried to follow this guidance to write the summary of thesis. It is four pages as suggested incl. an image that takes up one page. I use two paragraphs to reinstate the aims, one more than suggested, but do also include an image to show the PhD topics. The flow diagram was moved to the introduction. I use almost two pages to describe each of the research projects findings, which is much more than the suggested three paragraphs. I managed to keep the integrative summary to one paragraph as suggested. See images below, however, the thesis was later reviewed by supervisors, so the snapshot below is from the unedited version based of GenAI guidance for the structure. Version Thesis_251111.

Again – here we are dealing with a PRO in that GenAi helped me structure and breakdown concepts and ideas to make them better accessible. I could have written the section without, but it made it much easier to get started. I could also have asked my supervisors, but why do that, when you can get help elsewhere. Section 1.4 and 1.5 are now very different. In the end, not a lot remains from the initial GenAI proposal after a review by my supervisors.



Advise on processes

 # 1 Prompt: How do I best copy pdf into word without missing any pages

 # 1 Output: Here's how to copy a PDF into Word without losing pages or formatting, step by step:

Option 1: Use Word's Built-In PDF Import

If you have Microsoft Word 2013 or later, this is usually the easiest.

Open Word.

Go to File → Open → Browse and select your PDF.

Word will show a message:

"Word will now convert your PDF into an editable Word document."

Click OK.

It will import *all* pages into Word, preserving text, images, and layout.

☑ Best for: Text-based PDFs (reports, articles, manuals).

⚠ Caution: Complex layouts (tables, multi-columns) may shift slightly.

Option 2: Use Adobe Acrobat (if you have it)

If you have Adobe Acrobat Pro, you'll get the most accurate results.

Open the PDF in Acrobat Pro.

Go to File → Export To → Microsoft Word → Word Document.

Choose settings (e.g., retain layout, include images).

Save the file and open it in Word.

☑ Best for: Documents with graphics, columns, or forms.

💡 Acrobat's export keeps page integrity and rarely misses any pages.

🔄 Option 3: Use an Online Converter

If you don't have Acrobat, use trusted tools like:

Smallpdf

Adobe's free online converter

ILovePDF

Upload your PDF.

Convert and download the .docx.

Open it in Word and check page count.

☑ Best for: Occasional use.

⚠ Caution: Avoid uploading sensitive documents unless the service is secure.

📄 Option 4: For Image-Only PDFs (Scanned Documents)

If your PDF is made up of scanned images (not selectable text):

Use OCR (Optical Character Recognition) in:

Adobe Acrobat → Tools → Scan & OCR

Google Drive → Open with Google Docs

Online tools like OnlineOCR.net

☑ This converts images into editable text, preserving page order.

🧠 Tips to Ensure No Pages Are Missing

Always check the page count before and after conversion.

If pages are missing, split the PDF (e.g., with SmallPDF) and convert in smaller chunks.

Use “Keep source formatting” when pasting into existing Word docs.

In Word, enable “Show all pages” in View → Print Layout to confirm all imported.

Would you like me to show you how to do this using Word or Adobe Acrobat specifically, step-by-step with screenshots?

 # Results and interpretation: It took some trial and error, but I managed to copy my full pdf publication into the Appendix C (Publications) but the formatting is still not optimal because there are certain format requirements for the thesis that conflicts with the formatting of the published papers. I added statements to explain this in the appendix and hope my assessors go online and read the papers.

Well – here I did not get the result that I wanted. It could have been better, but it is likely better than if I tried copy pasting by myself, which I initially did. It resulted in only the first pdf page showing. Again – this was a simple case because I could compare the pdf before and after pasting into word for accuracy.

I hope these examples provide a good overview of how GenAI was used in accordance with university guidelines.

Declaration

I confirm that no content created by generative AI technologies has been presented as my own work.

[ii]