Cost Effectiveness of PCSK9 Inhibitors for the Secondary Prevention of Cardiovascular Disease in Ireland: Considerations given to Population Heterogeneity

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

Helen Geraldine O'Donnell

BSc (Pharm), MPharm, MSc

National Centre for Pharmacoeconomics

&

Department of Pharmacology and Therapeutics,

Trinity College University of Dublin

2020

DECLARATION

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work. I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

_____

Helen O'Donnell

Date: 31/03/2020

SUMMARY

The aim of this thesis was to evaluate the cost-effectiveness of PCSK9 inhibitors (alirocumab and evolocumab for the secondary prevention of cardiovascular disease (CVD) in Ireland. Also, to identify the subgroups of this population in which PCSK9 inhibitors might be deemed cost-effective. Five studies were undertaken in pursuit of this aim.

A systematic review of the literature was conducted. The results were included in a meta-analysis to quantify the comparative effectiveness of PCSK9 inhibitors versus standard of care. It was found that PCSK9 inhibitors reduce the time to non-fatal myocardial infarction and non-fatal stroke. No treatment effect on CVD death was observed over the time period of the clinical trials. The quality of evidence was low. An adjusted indirect treatment comparison between alirocumab and evolocumab was also conducted. The similarity assumption was considered to hold sufficiently to justify the conclusion that there is no evidence of a difference in treatment effect between alirocumab and evolocumab.

The interaction that strategic behaviour such as price negotiations can introduce to economic evaluations was identified for the first time. If a pharmaceutical company offers a conditional discount, which is dependent on obtaining reimbursement in two subgroups an interaction is generated. This means that cost savings generated from the discount in one subgroup may be used to offset the incremental cost of extending reimbursement in the other. A framework was presented to guide the economic evaluation process in the presence of an interaction. It was shown that failure to account for the interaction can lead to incorrect conclusions regarding the cost-effectiveness of interventions. Adoption of the framework is expected to increase population health through the increased recognition of cost-effective interventions.

The links between The Irish Longitudinal Study on Aging (TILDA) and the EQ-5D-3L were identified. Mapping the TILDA dataset to the EQ-5D-3L was identified as a pragmatic method of generating utilities in the Irish setting in the absence of directly observed evidence (the data gap). A mapping model between them was derived in a population with CVD. To our knowledge, this is the first time that a mapping study of this kind has been performed to address such a data gap.

The mapping model was applied to the national TILDA population. Utility values for 23 subgroups of the secondary prevention population were derived. The use of patient-level data means that heterogeneity in the utility of the secondary prevention CVD population can be captured. Regression methods were used to estimate utility decrements for age, and to

estimate chronic utility decrements for myocardial infarction, stroke and multiple CV events. The importance of using adjusted or unadjusted utility values depending on the stage of the analysis was highlighted.

An incremental economic evaluation of PCSK9 inhibitors versus standard of care across 23 subgroups. A previously published economic model was used to extrapolate predicted costs and outcomes over a lifetime time horizon given the paucity of baseline risk data in the Irish setting.

The economic model allowed the capture of the effect of population heterogeneity on baseline risk, treatment effect, costs and utility values. Regardless of whether direct or indirect treatment effects are applied, the results show that PCSK9 inhibitors are not cost-effective in the secondary prevention CV population in Ireland. An incremental evaluation of alirocumab and evolocumab to standard of care (SOC) was also conducted. Neither alirocumab nor evolocumab are cost-effective relative to SOC. However, when compared to evolocumab, alirocumab is likely to represent a cost-effective alternative.

The outputs of this thesis also have a broader impact for economic evaluation in Ireland. Derivation of the mapping model to TILDA opens the door to the link between the EQ-5D-3L and the wealth of data in TILDA. The identification of the interaction between price negotiations and heterogeneity has many implications. Pharmaceutical companies now have additional considerations when devising reimbursement strategies. A link between game theory and economic evaluation has been identified. Several implications for policy and practice were identified. Evolocumab is currently reimbursed in Ireland in a very restricted population under the management of the Medicines Management Program. While the confidential price paid by the HSE is unknown, the results of the economic evaluation show that reimbursement of evolocumab is unlikely to represent a cost-effective use of scarce healthcare resources in the Irish setting. Consideration should be given to further restricting reimbursement criteria for future patients. At current list prices, alirocumab may represent a cost-effective alternative to evolocumab. If this conclusion holds at confidential net prices for both drugs, consideration should be given to reimbursing alirocumab. However, this is based on the provision that the reimbursement criteria for alirocumab are not broader than the comparable criteria for evolocumab. Building on the findings of this thesis, several future research opportunities were identified.

ACKNOWLEDGEMENTS

First, I want to thank Professor Michael Barry for providing the opportunity to pursue this research, and for his support, encouragement and mentorship along the way. I also want to extend my heartfelt appreciation to Dr Laura McCullagh for her patience, knowledge and motivation.

Thank you to all my colleagues in the National Centre for Pharmacoeconomics. I am not only grateful for their knowledge and advice, but also for their endless support and encouragement. I particularly want to thank Dr Joy Leahy, Dr Lea Trela-Larsen, Dr David Mc Connell, and Professor Cathal Walsh for their statistical advice.

I also want to express my gratitude to Dr Andrew Maree, Professor Joe Harbison, and Dr Caroline Daly, and all in the Department of Cardiology in St. James's Hospital Dublin, for facilitating this research. Thanks also to all the patients who volunteered their time for data collection for this study.

My sincere thanks to Dr Miqdad Asaria, for permission to use and adapt the CALIBER economic model. I also want to thank the Irish Longitudinal study on Ageing (TILDA) and the Irish Social Science Data Archive for facilitating access to TILDA data.

To my friends and family, thank you for your kindness and patience. I particularly want to acknowledge my sister Sheila, and my deceased parents Michael and Kathleen, who supported and encouraged me in all my endeavours.

Finally, thank you to my fiancé Patrick for his love, patience, and endless support.

This thesis is dedicated to the memory of my parents,

Michael and Kathleen O'Donnell.

TABLE OF CONTENTS

Declaration ..... **i**

Summary ..... **ii**

Acknowledgements ..... **iv**

List of Tables ..... **xii**

List of Figures ..... **xiv**

List of Boxes ..... **xvi**

Abbreviations ..... **xvii**

List of Publications ..... **xix**

CHAPTER 1. INTRODUCTION 1

1.1 Introduction **2**

1.2 Aims and objectives **3**

1.3 Thesis outline **4**

CHAPTER 2. BACKGROUND 6

2.1 Health technology assessment **7**

2.1.1 Economic evaluation guidelines in Ireland 7

2.1.2 National Centre for Pharmacoeconomics 9

2.2 Heterogeneity **12**

2.3 Cardiovascular disease and the role of LDL-C **13**

2.3.1 Hypercholesterolemia 13

2.3.2 PCSK9 inhibitors 13

CHAPTER 3. EFFICACY OF PCSK9 INHIBITORS: SYSTEMATIC REVIEW AND EVIDENCE SYNTHESIS 20

3.1 Introduction **21**

3.1.1 Chapter outline 21

3.2 Literature review of previous meta-analysis and network meta-analysis of PCSK9 inhibitors 21

3.2.1 Pre-outcomes trials 21

3.2.2 Post-outcomes trials 23

3.2.3 Summary of literature review 26

3.3 Methods 26
3.3.1 Criteria for considering studies for this review 26
3.3.2 Search methods for identification of studies 28
3.3.3 Citation management 28
3.3.4 Assessment of risk of bias in included studies 29
3.3.5 Measure of treatment effect 29
3.3.6 Data synthesis 30
3.3.7 Subgroup analysis and assessment of heterogeneity 34
3.3.8 Assessment of quality and generation of summary of findings tables 34
3.4 Results 35
3.4.1 Excluded studies 37
3.4.2 Characteristics of included studies 37
3.4.3 Data availability 42
3.4.4 Risk of bias – primary efficacy outcomes 44
3.4.5 Other potential source of bias 44
3.4.6 Meta-analysis 44
3.4.7 Heterogeneity 49
3.4.8 Adjusted indirect comparison of alirocumab versus evolocumab 55
3.5 Discussion 59
3.5.1 Main findings 59
3.5.2 Strengths and limitations 61
3.5.3 Potential for further research 62
3.6 Conclusion 63
CHAPTER 4. THE INTERACTION BETWEEN PRICE NEGOTIATIONS AND HETEROGENEITY 64
4.1 Introduction 65
4.1.1 Aims 65
4.1.2 Chapter outline 66
4.2 Economic evaluation with population heterogeneity 66
4.2.1 Economic evaluation 66
4.2.2 Economic evaluation and heterogeneity 66
4.3 Strategic behaviour such as price negotiations can lead to interactions 67
4.3.1 Generation of the interaction 67
4.3.2 Introduction of the hybrid approach 68
4.4 Worked examples 69
4.4.1 Example 1 70

4.4.2 Example 2 76
4.4.3 Further examples 76
4.5 Strategic behaviour can increase the reimbursement price of new drugs 78
4.6 Discussion 78
4.6.1 Main findings 78
4.6.2 Limitations 81
4.6.3 Future Research 81
4.7 Conclusion 82
CHAPTER 5. BRIDGING THE DATA GAP: MAPPING TILDA TO THE EQ-5D-3L 83
5.1 Introduction 84
5.1.1 Aim 85
5.1.2 Chapter outline 85
5.2 Mapping in the literature 85
5.3 Methods 92
5.3.1 Overview of study design 92
5.3.2 Independent Sample 93
5.3.3 Outcome measures 94
5.3.4 Statistics and econometrics 98
5.3.5 Validation 105
5.3.6 Selection of the chosen model 106
5.4 Results 106
5.4.1 Independent Sample characteristics 106
5.4.2 Correlation 110
5.4.3 Estimation 111
5.4.4 Prediction 111
5.4.5 Validation 113
5.4.6 Selection of the best fitting model 113
5.5 Discussion 114
5.5.1 Main findings 114
5.5.2 Limitations 114
5.5.3 Potential for future research 115
5.5.4 Lessons for a future mapping study 116
5.6 Conclusion 117

CHAPTER 6. UTILITY VALUES FOR AN ECONOMIC MODEL IN CVD 118

6.1 Introduction 119

6.1.1 Aim 119

6.1.2 Chapter outline 119

6.2 Considerations when deriving utility values for cardiovascular disease 119

6.2.1 Irish population health and preferences 120

6.2.2 Baseline health 121

6.2.3 Joint health conditions 121

6.2.4 Causal health decrements 125

6.2.5 Account for acute and chronic effects of events 126

6.2.6 Adverse event associated with treatment 127

6.2.7 Sensitivity and scenario analyses 127

6.2.8 Summary 127

6.3 Cardiovascular utilities derived from the EQ-5D in the literature 127

6.3.1 Utility catalogues estimated by Sullivan et al. 127

6.3.2 Health Survey for England data 130

6.3.3 Other studies 131

6.4 Approaches taken in other models 133

6.4.1 CALIBER model approach. 133

6.4.2 Ara et al assessment of cardiovascular models. 133

6.4.3 Utilities adopted in NICE technology appraisal of PCSK9 inhibitors. 134

6.5 Generation of cardiovascular utilities for the economic model 134

6.5.1 Methods 135

6.5.2 Results 140

6.5.3 Summary of utilities used in the economic model 145

6.6 Discussion 147

6.6.1 Main findings 147

6.6.2 Limitations and potential for future research 149

6.7 Conclusion 151

CHAPTER 7. ECONOMIC EVALUATION OF PCSK9 INHIBITORS FOR THE SECONDARY PREVENTION OF CVD 152

7.1 Introduction 153

7.1.1 Aim 153

7.1.2 Chapter outline 153

7.2 Previous economic evaluations of PCSK9 inhibitors 154

7.3 The CALIBER CV model 155

7.3.1 Population 155

7.3.2 Model structure 155

7.3.3 Baseline risk 157

7.3.4 Costs 162

7.3.5 Utilities 163

7.3.6 Treatment effect scenarios 163

7.3.7 Results 164

7.3.8 Transferability of the economic model 164

7.4 Development of the CALIBER economic model for the economic evaluation of PCSK9 inhibitors in Ireland. 165

7.4.1 Irish reference case 165

7.4.2 Population 165

7.4.3 Intervention 167

7.4.4 Comparators 168

7.4.5 Model structure 168

7.4.6 Baseline risk 168

7.4.7 Adverse events 170

7.4.8 Resource use and costs 170

7.4.9 Utilities 172

7.4.10 Treatment effectiveness 172

7.4.11 Discontinuation rates 175

7.4.12 Probabilistic analysis 175

7.5 Results 176

7.5.1 Population characteristics 176

7.5.2 PCSK9 inhibitors versus standard of care 178

7.5.3 Evolocumab and alirocumab versus standard of care 189

7.5.4 Alirocumab versus evolocumab 194

7.6 Discussion 195

7.6.1 Deterministic versus probabilistic results 195

7.6.2 Alirocumab versus evolocumab 196

7.6.3 Benefits of real-world data 196

7.6.4 Estimation of baseline risk 196

7.6.5 Treatment effect 196

7.6.6 Utilities 197

7.6.7 Heterogeneity 197

7.6.8 Modelling in R 198
7.6.9 Limitations 199
7.7 Conclusion 199

CHAPTER 8. CONCLUSION 201
8.1 Introduction 202
8.2 Main findings 202
8.3 Implications for policy and practice 203
8.4 Future research 204
8.5 Conclusion 205
References 206
Appendices 229

LIST OF TABLES

Table 2.1 Summary of the Irish reference case 7
Table 2.2 ESC/EAS risk categories and LDL-C targets (32) 17
Table 2.3 PCSK9 inhibitor reimbursement criteria for NICE and the SMC (33–36) 19
Table 3.1 Quality of evidence grades under the GRADE approach (77) 35
Table 3.2 Baseline characteristics of trial populations 40
Table 3.3 Data availability for outcomes of interest across included studies 43
Table 3.4 Summary of findings PCSK9 inhibitors compared to placebo for the secondary prevention of cardiovascular disease across primary outcomes 47
Table 3.5 Effect of PCSK9 inhibitors versus placebo on time to CV death by baseline LDL-C (136) 51
Table 3.6 Review outcomes by baseline LDL-C and history of CVD 53
Table 3.7 Summary of findings table alirocumab compared to evolocumab for prevention of cardiovascular disease 58
Table 3.8 Adjusted indirect comparison alirocumab v evolocumab for the primary review outcomes 58
Table 4.1 Total and incremental costs and QALYS under the stratified and weighted average approaches in Examples 1 and 2 (equal patient numbers are assumed in both subpopulations) 71
Table 4.2 Total and incremental costs and QALYS for Examples 1 and 2 under the hybrid approach 75
Table 5.1 Extract of mapping studies to the EQ-5D-3L, in CVD, from the Health Economics Research Centre mapping database (199,204,205) 91
Table 5.2 'Independent Sample' characteristics 108
Table 5.3 Spearman correlation coefficients between EQ-5D-3L domains and comparable TILDA variables included in the mapping model 110
Table 5.4 Predicted model characteristics and performance 113
Table 6.1 Age standardised country-specific reported problems in EQ-5D-5L in populations with CHD 121
Table 6.2 Summary of utility catalogue characteristics estimated by Sullivan et al (233–235) 128
Table 6.3 Utility multipliers calculated from Health Survey for England data reported by Ara et al (253). 131
Table 6.4 Mean utility values in the ODYSSEY program 132

Table 6.5 Utilities used in the CALIBER economic model by Asaria et al (based on Sullivan et al 2011 catalogue) (15,234) 133
Table 6.6 Coefficients and standard errors for age and sex for predicted utility regressed on sociodemographic and cv variables in an ordinary least squares (OLS) model. 141
Table 6.7 Acute event utility multipliers estimated from Ara et al (253) 142
Table 6.8 Coefficients and standard errors for age and sex for predicted utility regressed on sociodemographic and cardiovascular variables in an OLS model 143
Table 6.9 Coefficients and standard errors for MI and stroke coefficients for predicted utility regressed on sociodemographic and cardiovascular variables in an OLS model 144
Table 6.10 Calculation of utility multipliers from the national TILDA population 145
Table 6.11 Utilities used in base case and scenario analysis for the economic model evaluating PCSK9 Inhibitors in Chapter 7 146
Table 7.1 Adjusted hazard ratios for 5-year risk of CV events for multiple versus single vascular beds adjusted for age and sex from the Oxford Vascular Study (301) 169
Table 7.2 Evolocumab and alirocumab drug costs 172
Table 7.3 Hazard ratios applied in base-case and scenario analysis for estimating the effect of PCSK9 inhibitors versus placebo 175
Table 7.4 Patient characteristics for TILDA population 1 (defined as having experienced MI, angina, TIA or stroke) 177
Table 7.5 Probabilistic results of the basecase analysis for the comparison of PCSK9 inhibitors versus standard of care 180
Table 7.6 Total and incremental costs and outcomes for PCSK9 inhibitors versus standard of care for 23 subgroups TILDA with a modelled of 4mmol/L , under the direct treatment effect approach (deterministic analysis) 183
Table 7.7 Total and incremental costs and outcomes for PCSK9 inhibitors versus standard of care for TILDA population 1 by LDL-C threshold. 184
Table 7.8 Deterministic costs and outcomes for TILDA population at an LDL-C treatment threshold of 4mmol/L under alternative treatment effect assumptions. 185
Table 7.9 Scenario analysis for basecase analysis PCSK9 inhibitors under different utility source under direct and indirect treatment effect assumptions. 187
Table 7.10 Scenario Analysis examining cost-effectiveness under different price assumptions 188
Table 7.11 Mean expected costed and outcomes for the comparison of alirocumab, evolocumab and SoC 191

LIST OF FIGURES

Figure 2.1 Timeline of PCSK9 inhibitor therapy from discovery to the present time. 15
Figure 3.1 PRISMA diagram 36
Figure 3.2 Efficacy of PCSK9 inhibitors versus placebo for reducing the risk of the primary outcomes 46
Figure 3.3 Meta-analysis of secondary outcomes for PCSK9 inhibitors versus placebo 48
Figure 3.4 Network diagram of RCT evidence 56
Figure 4.1 Incremental costs and QALYs of drug 'D' v. standard of care 'S' under the stratified approach by subpopulation and price for Examples 1 and 2 70
Figure 4.2 Incremental costs and QALYs of drug 'D' v. standard of care under the weighted average approach for the total licensed population at the negotiated price for Examples 1 and 2 72
Figure 4.3 Incremental analysis of alternatives under the hybrid approach for Examples 1 and 2. 73
Figure 4.4 Further examples (a) and (b) of potential reimbursement outcomes under the hybrid approach. 77
Figure 5.1 Overview of mapping study design 93
Figure 5.2 'Independent Sample' participant flow diagram 107
Figure 5.3 Distribution of the EQ-5D-3L in the Independent Sample 108
Figure 5.4 Independent Sample responses to the Question Subset variables included in the mapping model. 109
Figure 5.5 Mean observed and predicted utility values (± standard error) for the six mapping models estimated across independent variables. 112
Figure 7.1 Economic model and risk equations adapted from Asaria et al. (15). 157
Figure 7.2 Proportion of patients in each of the six primary health states over time as predicted by the multi-state model in the base-case analysis of PCSK9 inhibitors versus standard of care. 179
Figure 7.3 Scatter plot of incremental comparison of expected costs and QALYs for PCSK9 inhibitors versus standard of care at an LDL-C threshold of 4mmol/L. (Direct Treatment Effect) 180
Figure 7.4 Cost effectiveness acceptability curve for PCSK9 inhibitors versus standard of care under the basecase analysis 181
Figure 7.5 Tornado diagram of one way sensitivity analysis for PCSK9 inhibitors versus standard of care under the base-case analysis. 189

Figure 7.6 Proportion of patients in each of the six primary health states over time as predicted by the multi-state model in the base-case analysis of alirocumab, evolocumab and standard of care..... 190
Figure 7.7 Scatter plot of incremental costs and QALYS for alirocumab and evolocumab versus standard of care..... 192
Figure 7.8 Cost-Effectiveness acceptability curve for the comparison of evolocumab , alirocumab and standard of care. 193
Figure 7.9 Cost Effectiveness acceptability curve for alirocumab versus evolocumab in TILDA population 1 at an LDL-C threshold of 4mmol/L assuming direct treatment effect..... 194

LIST OF BOXES

Box 2.1 Possible rapid review outcomes 10
Box 2.2 Full HTA recommendations 10
Box 5.1 Summary of pre-modelling recommendations by Wailoo et al (200) 87
Box 5.2 Summary of modelling and data analysis recommendations by Wailoo et al (200) 87
Box 5.3 Summary of reporting of mapping studies recommendations by Wailoo et al (200) ... 88
Box 5.4 Mapping study inclusion and exclusion criteria 94

ABBREVIATIONS

ACS	Acute coronary syndrome
AIC	Akaike information criteria
ALDV	Adjusted limited dependent Variable
BIC	Bayesian information criteria
CABG	Coronary artery bypass graft
CAD	Coronary artery disease
CCC	Clinical classification categories
CES-D	Centre for Epidemiological Studies Depression Scale
CHD	Coronary heart disease
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CLAD	Censored least absolute deviation
CPU	Corporate pharmaceutical unit
CTTC	Cholesterol treatment trialists collaboration
CV	Cardiovascular
CVD	Cardiovascular disease
EMA	European medicines agency
FH	Familial hypercholesterolemia
HADS-A	Hospital anxiety depression scale – anxiety subscale
HeFH	Heterozygous familial hypercholesterolemia
HIQA	Health Information and Quality Authority
HoFH	Homozygous familial hypercholesterolemia
HR	Hazard ratio
HRQoL	Health related quality of life
HSE	Health service executive
ICD-9	9th Version International Classification of Diseases
IPHA	Irish pharmaceutical healthcare association
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
LDL-C	Low density lipoprotein cholesterol
MACE	Major adverse cardiovascular events
MAE	Mean absolute error

MAPS	Mapping onto Preference-based measures reporting standards
MI	Myocardial infarction
MINAP	Myocardial Ischaemia National Audit Project Registry
mmol/L	Millimoles per Litre
NCPE	National Centre for Pharmacoeconomics
NICE	National Institute for Health and Care Excellence
NSTEMI	Non-ST elevation myocardial infarction
OLS	Ordinary least squares
OR	Odds ratio
PAD	Peripheral arterial disease
PCSK9	Proprotein convertase subtilisin/kexin type 9
PSA	Probabilistic sensitivity analysis
QALY	Quality adjusted life years
QPC	Quality priority condition
RCT	Randomised controlled trial
ROB 2	Cochrane Risk of Bias 2
RMSE	Root mean square error
RR	Risk ratio
SE	Standard error
SMC	Scottish Medicines Consortium
SOC	Standard of care
STEMI	ST elevation myocardial infarction
TIA	Transient ischemic attack
TILDA	The Irish Longitudinal Study on Aging

LIST OF PUBLICATIONS

First Author Publications

O'Donnell H, McCullagh L, Barry M, Walsh C. The Interaction between Price Negotiations and Heterogeneity: Implications for Economic Evaluations. Medical Decision Making. 2020;40(2)144-155

Peer Reviewed Oral Presentations

O'Donnell H, McCullagh L, Barry M, Walsh C. A Framework for Economic Evaluations with Interactions between Groups. HTAi. Cologne. Jun 17 2019.

O'Donnell H, McCullagh L, Barry M. Mapping A National Longitudinal Survey to the EQ-5D-3L for HTA. HTAi Annual Conference. Vancouver. June 2018.

Peer Reviewed Poster Presentations

O'Donnell H, McCullagh L, Barry M. Mapping The Irish Longitudinal Study on Aging (TILDA) to the EQ-5D-3L for Economic Evaluation. SPhERE Network Annual Conference. Dublin. February 26 2019.

O'Donnell H, McCullagh L, Barry M, Walsh C. Negotiating the Relationship between Price and Volume: Implications for ICER Calculations in Heterogenous Treatment Populations. SMDM 17th Annual Biennial Conference. Leiden. June 10-12 2018
Winner 2nd Place Lee Lusted Student prize for outstanding presentation of research.

O'Donnell H, McCullagh L, Barry M. Secondary Prevention Cardiovascular Patients in Ireland – Population Characteristics for Cost-Effectiveness Analysis. ISPOR Glasgow. Nov 04-08 2017.

O'Donnell H, McCullagh L, Barry M. The Potential Budget Impact of PCSK9 inhibitors in Ireland. ISPOR Vienna. Oct 29- Nov 02 2016.

Additional publications unrelated to the work in this thesis.

To HTA or Not to HTA: Identifying the Factors Influencing the Rapid Review Outcome in Ireland. O'Donnell H, Lamrock L, Tilson L, Barry M. Value in Health 2020;23(2)274-275

Connolly E, O'Donnell H, Lamrock L, Tilson L, Barry M. Health Technology Assessment of Drugs in Ireland: An analysis of Timelines. Pharmacoeconomics (Open) doi.org/10.1007/s41669-019-00177-8

Gorry C, McCullagh L, O'Donnell H, Barrett S, Schmitz S, Barry M, Curtin K, Beausang E, Barry R, Coyne I. Neoadjuvant treatment for malignant and metastatic cutaneous melanoma (Protocol). Cochrane Database of Systematic Reviews. 2018, Issue 3. Art No.: CD012974

Chapter 1. Introduction

1.1 Introduction

Elevated levels of low density lipoprotein cholesterol (LDL-C) is one of many known risk factors for cardiovascular disease (CVD) (1). While statins are the backbone of treatment for hypercholesterolemia, a new class of therapy, proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitors, has emerged. Evolocumab and alirocumab are the first monoclonal antibodies licensed for the treatment of hypercholesterolemia. They have been shown to reduce LDL-C levels by up to 70% (2,3). In 2015, both were licensed by the European Medicines Agency (EMA) for the treatment of patients who fail to reach LDL-C goals on maximum tolerated dose of statins alone and in those who are statin intolerant (4,5). The ability to reduce the risk of cardiovascular (CV) events was subsequently confirmed in populations with a previous history of CV events (6,7). This broad license of PCSK9 inhibitors means that they are licensed for both primary prevention and secondary prevention of CVD. Primary prevention refers to prevention of a first CV event such as myocardial infarction (MI) or stroke in those who have no history of CVD. Secondary prevention refers to the prevention of further CV events in those who already have a history of a CVD.

The large number of patients potentially eligible for treatment coupled with an indefinite treatment duration means that indiscriminate use of PCSK9 inhibitors could place a substantial cost on the Irish healthcare system. Economic evaluation is a comparison of the costs and consequences of a health technology, with one or more alternatives, to determine if the incremental benefits of intervention are worth the increased costs (8). In any healthcare system with a limited budget, choices must be made in order to optimise the use of limited healthcare resources.

Population heterogeneity means that there are differences in demographic or clinical characteristics across the patient population. This may introduce clinical heterogeneity where the relative and absolute clinical benefit of PCSK9 inhibitors varies across the licensed population. Population heterogeneity may also translate to economic heterogeneity where there are differences in cost-effectiveness of PCSK9 inhibitors across subgroups (where subgroups are defined by named patient characteristics). If PCSK9 inhibitors are not cost-effective in all subgroups in which they are reimbursed, their high cost could impose a significant opportunity cost into the healthcare system in the form of health forgone. Therefore, it is important to identify the subgroups of the licensed population in which they are most cost-effective.

In line with drug reimbursement policy in Ireland, the National Centre for Pharmacoeconomics (NCPE) in Ireland has critically assessed the clinical and economic evaluations submitted by the marketing authorisation holders of evolocumab and alirocumab (9–11). The NCPE had multiple concerns regarding the modelling approaches taken by the applicants. Under the NCPE's adjusted base case, evolocumab and alirocumab were not cost-effective versus standard of care in any of the subgroups examined. Reimbursement was not recommended unless cost-effectiveness could be improved relative to existing treatments (9–11).

Following confidential price negotiations, evolocumab was approved for reimbursement. But only in subgroups of the licensed population coupled with strict clinical criteria (12). Alirocumab has not been approved for reimbursement to date. An independent multi-technology appraisal, where the cost-effectiveness of evolocumab versus alirocumab is determined, would inform future reimbursement decisions.

The treatment landscape of hypercholesterolemia continues to change. Bempedoic acid is an oral non-statin LDL-C lowering drug which received Committee for Medicinal Products for Human Use (CHMP) positive opinion in January 2020 from the EMA (13). An application for marketing authorisation for inclisiran, a PCSK9 synthesis inhibitor, was filed with the EMA in February 2020 (14). Robust estimates of the cost-effectiveness of current therapies are required to order to inform future evaluations of these drugs.

Given the limitations of previous evaluations, independent estimates of the cost-effectiveness of PCKS9 inhibitors are required in order to identify subgroups of the population in which they are cost-effective.

1.2 Aims and objectives

The aim of this thesis is to quantify the cost-effectiveness of PCSK9 inhibitors (evolocumab and alirocumab) for the secondary prevention of CVD in Ireland. Also, to identify the subgroups of the population in which PCSK9 inhibitors could be deemed to be cost-effective.

The following objectives in pursuit of these aims were identified:

- Develop an understanding of the relationship between price negotiations and population heterogeneity and the implications for economic evaluation.
- Review and synthesise the clinical evidence for PCSK9 inhibitors (for reducing the risk of cardiovascular events) and examine if this varies by subgroup.
- Derive utility values in CVD reflective of the population in Ireland.

- Develop a decision analytic model economic model, reflective of the population in Ireland with CVD, and which can provide results stratified by subgroups within the licensed population.

Evaluation of the clinical evidence and the cost-effectiveness of PCSK9 inhibitors for the primary prevention of CVD in Ireland is outside the scope of this research.

1.3 Thesis outline

The structure of the thesis is outlined below:

Chapter 2 provides background information. It describes the clinical information of relevance for this thesis, including the relationship between hypercholesterolemia and CVD, and the role of PCSK9 inhibitors and other lipid modifying therapies. Methods of economic evaluation are introduced and their use in Ireland is described. The concept of population heterogeneity is introduced and the importance of accounting for heterogeneity in economic evaluation is outlined.

Chapter 3 presents a systematic review of the evidence of PCSK9 inhibitors for the secondary prevention of CVD. The evidence identified is synthesised in a meta-analysis quantifying the efficacy of PCSK9 inhibitors versus placebo and in an adjusted indirect comparison of alirocumab versus evolocumab.

Chapter 4 recognises that individual countries have engaged in price negotiations with pharmaceutical companies and have restricted the reimbursement of PCKS9 inhibitors to only those subgroups of the population in whom they are cost-effective. This chapter describes a heterogeneity-price interaction that can be introduced into economic evaluations. A framework is presented for how drugs such as PCSK9 inhibitors should be assessed in these scenarios.

Chapter 5 presents a pragmatic solution to the paucity of utility values in the Irish setting. The development (and the outputs) of a mapping model, between The Irish Longitudinal Study on Aging (TILDA) and the EQ-5D-3L are described.

Chapter 6 describes considerations to be undertaken when populating an economic model with utility values. The mapping model (developed in Chapter 5) is applied to the national TILDA population and health state utility values for a decision analytic model in CVD are derived.

Chapter 7 describes an economic evaluation of PCSK9 inhibitors, for the secondary prevention of CVD, in the Irish healthcare setting. A model derived by Asaria et al (that predicts costs and outcomes in coronary heart disease (CHD) in the UK) is adapted for the Irish healthcare setting.

The model is expanded to account for the efficacy of PCSK9 inhibitors in a population with CVD in Ireland (15). The effectiveness and cost effectiveness of PCSK9 inhibitors in a range of subgroups of a heterogeneous population (where subgroups are identified by patient characteristics) is investigated.

Model inputs include efficacy outputs from Chapter 3 and the health state utility values from work presented in Chapter 5 and 6.

Chapter 8 summarises the main findings in this thesis. The contribution of this research to the literature is described. Potential implications for policy and practice are identified.

Chapter 2. Background

2.1 Health technology assessment

Health Technology Assessment (HTA) has been defined as "*a multidisciplinary process that summarises information about the medical, social, economic, and ethical issues related to the use of health technology in a systematic, transparent, unbiased robust manne*r" (16). Its aim is defined as informing the "*formulation of safe, effective, health policies that are patient focused and seek to achieve best value*" (16).

In Ireland, both the NCPE and the Health Information and Quality Authority (HIQA) have defined roles regarding the health technology assessment of Ireland. The NCPE assess pharmaceutical technologies while HIQA assess other technologies.

Within the framework of HTA, the NCPE critically assesses economic evaluations submitted by pharmaceutical companies in support of their drug. Based on this evidence, and independent systematic review, the NCPE issues recommendations for reimbursement.

2.1.1 Economic evaluation guidelines in Ireland

As described in Chapter 1, economic evaluation is a comparison of the costs and consequences of a health technology, with one or more alternatives, to determine if the incremental benefits of intervention are worth the increased costs (8). Guidelines for economic evaluation are published by HIQA (8). These outline the principles or reference case that should be followed when conducting an evaluation in Ireland. The reference case is summarised in Table 2.1.

Table 2.1 Summary of the Irish reference case

Element	Reference Case
Evaluation type	Cost-Utility Analysis
Perspective on costs	The publicly funded health and social care system in Ireland (HSE)
Perspective on outcomes	All health benefits accruing to individuals
Choice of comparator	Routine care in Ireland
Synthesis of effectiveness	Based on systematic review
Outcome measurement	QALYs
Discount rate	Apply an annual rate of 4% on costs and outcomes after the first year
Sensitivity analyses	Probabilistic and deterministic sensitivity analysis
Equity rating	Equal weighting should be applied to the outcome measure

HSE, Health Service Executive; QALYs, Quality Adjusted Life Years.

The preferred evaluation type in the Irish reference case is a cost utility analysis. Cost utility analyses use quality adjusted life years (QALYs) as the outcome measure. QALYs simultaneously account for changes in the quantity of life and the quality of life. Using a generic outcome measure such as the QALY means that outcomes across different disease areas can be compared. QALYS are weighted using utilities. These represent the preferences of individuals for health states. There are many ways of measuring utility. Irish guidelines recommend that utility is measured indirectly using generic preference-based instruments. The EQ-5D-3L is the instrument recommended by the NCPE. This is composed of two parts. A descriptive system is used to evaluate a person's health state. Then, a utility is generated by applying a previously derived value set associated with that health state. A value of 1, represents perfect health. States equivalent to death are represented by 0. States worse than death are possible.

A standard measure of quantifying the value of an intervention is the incremental cost-effectiveness ratio (ICER). This can be defined as the additional cost per unit of health benefit compared to another intervention (herein ''the comparator'').

The ICER of drug 'D' vs comparator 'S' for is defined as:

$$ICER_D = \frac{C_D - C_S}{Q_D - Q_S}$$

where C_D, C_S and , Q_D , Q_S are the mean costs and QALYs of drug 'D' and standard of care 'S' and the incremental costs and QALYS are greater than 1. A decision rule may be defined such that an intervention is considered cost-effective and reimbursed if the ICER of drug 'D' (versus the standard of care) is below the cost-effectiveness threshold (λ).

In situations where either the incremental costs or QALYS are negative or both are negative. Comparing the ICER and the threshold requires greater effort. The net monetary benefit is easier to interpret in these scenarios. This is defined as the incremental QALYS multiplied by the cost-effectiveness threshold minus the incremental costs. An intervention is cost-effective if the net monetary benefit is greater than zero.

2.1.2 National Centre for Pharmacoeconomics

The NCPE was established in 1996. The mission of the NCPE is to "Facilitate healthcare decisions on the reimbursement of technologies, by applying clinical and scientific evidence in a systematic framework in order to maximise population wellness" (17).

One of the primary functions of the NCPE is to advise the Health Service Executive (HSE) regarding the cost-effectiveness of new pharmaceuticals. In 2013, legislation underpinning decisions on pharmaceutical reimbursement was laid out for the first time with publication of the Health Act (18). The act states that the HSE should have regard to the cost-effectiveness of supplying the item concerned as opposed to providing other services and also to take into account the proposed costs benefits and risks of the item and the of uncertainty in relation to them.

A series of periodic agreements between the Irish Pharmaceutical Healthcare Association (IPHA), the HSE, the Department of Health and the Department of Public Expenditure and Reform lays a framework for the interaction between the pharmaceutical companies and the state when seeking reimbursement of new drugs in Ireland (19).

Marketing authorisation holders must apply to the Corporate Pharmaceutical Unit (CPU) of the HSE when seeking state reimbursement of their product in Ireland. A rapid review must be submitted to the NCPE by the applicant as part of the application. The rapid review outlines information regarding the target population, clinical efficacy and safety data, potential comparators, price of the intervention relative to comparators and five-year estimated budget impact. If a drug has multiple indications, an individual rapid review must be submitted for each. Based on an assessment and critique of the applicant's submission as well as an independent review of the literature, the NCPE make a recommendation to the CPU regarding reimbursement and the need for further assessment. The NCPE endeavour to complete this process within four weeks of receiving the applicant's submission.

The aim of the rapid review process is to categorise submissions and identify those which require a more intensive assessment. HTA has an opportunity cost in terms of delaying access of potentially valuable and cost-effective treatments. Therefore, the rapid review is adopted to direct resources to where an assessment is most valuable (20). Following a rapid review, reimbursement may be recommended at the submitted price or subject to price negotiations. If the NCPE believe that a further, more detailed assessment is warranted to justify the price sought, a HTA may be recommended. Drugs with robust clinical efficacy data with similar efficacy to comparators and a low budget impact may avoid a HTA. Given the potential of a

positive recommendation at rapid review stage to avoid a HTA, empirical analysis of the factors affecting rapid review recommendations have been published in the literature (21). Analyses highlight that the factors driving rapid review outcomes are not clear cut (20,21). All analyses were subject to limitations including the omission of important variables that may be driving the reimbursement decision (20).

The applicant is mandated to complete a full HTA report according to the NCPE applicant template, local NCPE guidelines and national guidelines for the economic evaluation which outline the Irish reference case. The onus is on the applicant to present a sufficient clinical and economic case to support a positive reimbursement recommendation from the HSE. Potential rapid review and full HTA outcomes are presented in Box 2.1 and Box 2.2 respectively.

Box 2.1 Possible rapid review outcomes

- A full HTA is recommended to assess the clinical effectiveness and cost effectiveness of [Medicine] compared with the current standard of care
- A full HTA is recommended to assess the clinical effectiveness and cost effectiveness of [Medicine] compared with the current standard of care, on the basis of the proposed price relative to currently available therapies.
- A full HTA is not recommended. The NCPE recommends that [Medicine] be considered for reimbursement.*
- A full HTA is not recommended. The NCPE recommends that [Medicine] not be considered for reimbursement at the submitted price
- A full HTA is not recommended until additional efficacy and/or safety data is submitted. On the basis of current evidence, the NCPE recommends that [Medicine] not be considered for reimbursement.*

* This recommendation should be considered while also having regard to the criteria specified in the Health (Pricing and Supply of Medical Goods) Act 2013.

Box 2.2 Full HTA recommendations

- The NCPE recommends that [Medicine] be considered for reimbursement*.
- *This recommendation should be considered while also having regard to the criteria specified in the Health (Pricing and Supply of Medical Goods) Act 2013.
- The NCPE recommends that [Medicine] be considered for reimbursement if cost-effectiveness can be improved relative to existing treatments*.
- The NCPE recommends that [Medicine] not be considered for reimbursement unless cost effectiveness can be improved relative to existing treatments*.
- The NCPE recommends that [Medicine] not be considered for reimbursement*.
- This recommendation should be considered while also having regard to the criteria specified in the Health (Pricing and Supply of Medical Goods) Act 2013.

A HTA process has previously been conceptualised as a diagnostic test (22). In this manner the rapid review process could be conceptualised as a screening test to identify drugs and indications which require greater scrutiny. The rapid review process is an efficient way of determining the requirement for a full HTA and targeting resources for those drugs for which there is most value in conducting an HTA (20).

The IPHA agreement does not specify that interventions above defined cost-effectiveness thresholds should not be reimbursed (19). Rather, a decision authority level table is defined such that a cost-effectiveness threshold is combined with budget impact thresholds to define the decision authority level within the HSE for the drug concerned. At cost-effectiveness thresholds above €45,000 per QALY or if the net drug budget impact (as estimated by the NCPE) is greater than >€20 million the decision lies with HSE leadership team. If the ICER is lower than €45,000 per QALY and the net drug budget impact is < €5million or if the ICER is lower than €20,000 per QALY and the net drug budget impact is between €5 million and €20 million non-HSE leadership may make a decision on reimbursement. Drug reimbursement decisions for drugs not meeting these thresholds must be taken by HSE leadership. The HSE leadership team are advised by the HSE drugs group who are in turn informed by NCPE recommendations, National Cancer Control Program for cancer drugs and the HSE Rare Disease Technology Review Group for certain orphan drugs which are referred by the HSE. Views of the applicant company are presented through the CPU. The framework defines the situation where if the HSE cannot fund the drug within existing resources, it may inform the Department of Health of its decision. The framework outlines how the Department of Health may bring a memorandum to Government in relation to the funding implications and request consideration of same (19).

2.1.2.1 Cost-Effectiveness threshold

There is no formal cost-effectiveness threshold in Ireland. Previous incarnations of the HSE-IPHA agreement recommended reimbursement if ICERs calculated by the NCPE indicated that ICERs were below €45,000 per QALY. But NCPE assessments of full HTAs generally report the probability of cost-effectiveness at thresholds of €20,000 and €45,000 per QALY. Given the confidential nature of price negotiations, it is unclear what cost-effectiveness threshold applies in practice. Prior to the current IPHA agreement, O'Mahony and Coughlan highlighted that the current threshold of €45,000 per QALY had no empirical basis and that it was likely too high given UK estimates of approximately £13,000 per QALY (23). They were concerned that as a threshold resembles a price floor than a ceiling, it was weak barrier to cost-ineffective interventions (23).

Chen et al reviewed the literature informing the cost-effectiveness of procedures with the longest waiting lists in the Irish public health system (24). Of the top 20 waiting list procedures, 17 had ICERs lower than €45,000 per QALY, 14 fell below €20,000 per QALY with 10 falling below €10,000 per QALY. This implies resource misallocation and that population health could be improved. It provides indirect evidence that the threshold adopted commonly adopted in Ireland may be too high (24).

2.2 Heterogeneity

It has previously been recognized that population heterogeneity may lead to variations in estimates of the cost-effectiveness of an intervention across subgroups (25).

Variations in cost-effectiveness may be driven by differences in the clinical or economic outcomes. Clinical outcomes may differ because of difference in the absolute or relative size of the treatment effect or because of differences in the underlying baseline risk. Differences in economic outcomes often differ because of drug costs. For example, where the dose and therefore the cost of a drug is based on weight, the cost of therapy can differ substantially by weight and age.

Coyle et al examined the role of stratified cost-effectiveness analysis and measured the value of restricting reimbursement of interventions to the subgroups in whom they are cost-effective (26). Basu and Meltzer examined the extension of this concept to decisions on an individual level (27). Further work by Kim and Basu suggested the incremental evaluation of policy alternatives which affect adoption behaviour in tandem with the assessment of the cost-effectiveness of the intervention (28).

However, the causes of heterogeneity in the cost-effectiveness across the population may not always be known and there is a cost to conducting further research to identify them. Espinoza et al previously differentiated between revealing the factors associated with heterogeneity in costs and outcomes using existing evidence (static value) and the value of conducing further subgroup related evidence generation (dynamic value) (29). The expected value of both forms of uncertainty can inform research and reimbursement decisions (29).

There may be ethical issues associated with restricting access to drugs by subgroups. However, Sculpher has previously highlighted that this may be in part counter-balanced with the opportunity cost of reimbursing drugs in subgroups of the population in whom they are not cost effective (25).

Care should be taken to avoid "data dredging" where spurious findings are used to support increased benefit in certain groups. Clinical trials are rarely powered to support subgroup specific analysis. Therefore, Irish national guidelines recommend that stratified analysis of subgroups is appropriate when there is biological or clinical support for heterogeneity in the target population (8).

2.3 Cardiovascular disease and the role of LDL-C

2.3.1 Hypercholesterolemia

Hypercholesterolemia is a type of dyslipidaemia characterised by elevated levels of total cholesterol or LDL-C in the bloodstream. Atherosclerosis as a form of CVD can be defined as the formation and hardening of fatty plaques on the inner surface of arteries. Over time plaque build-up causes narrowing of arteries. This may cause lesions to become unstable and acute CV events may occur. CV atherosclerosis can be broadly characterised into three main entities based on the location of atherosclerosis: coronary arterial disease (CAD), ischaemic stroke and peripheral arterial disease (PAD). The cause of CVD is multifactorial including patient modifiable and unmodifiable risk factors. One clinical modifiable risk factor is hypercholesterolemia. Clinical trials, epidemiological studies and genetic studies have shown a link between reducing cholesterol levels and a reduction in CVD (1).

In most cases, hypercholesterolemia is due to a combination of factors including dietary and environmental with some genetic influence. However, in familial hypercholesterolemia (FH), elevated LDL-C levels are primarily due to the inheritance of a defective gene. This results in LDL-C levels between 5-15 millimoles per litre (mmol/L). It is an autosomal dominant condition meaning that the inheritance of a defective gene from one parent is enough for LDL-C elevations to occur. Heterozygous FH (HeFH) is the most common form of FH where the gene has been inherited from one parent. Homozygous FH (HoFH) is a very rare more severe form of FH with LDL-C often >13mmol/L due to the inheritance of a defective gene from both parents (1).

2.3.2 PCSK9 inhibitors

As described in Chapter 1, evolocumab and alirocumab are the first drugs licensed in a new class of medicines called PCSK9 inhibitors. Both are fully human, monoclonal, immunoglobulin G1 antibodies. By inhibiting the binding of PCSK9 to LDL-C receptors, PCSK9 inhibitors increase the number of LDL-C receptors available to clear LDL-C, thereby lowering LDL-C levels.

A timeline of the history of PCSK9 inhibitors from drug target discovery to reimbursement is presented in Figure 2.1.

PCSK9 as a target for drug therapy was first identified when a gain of function mutation in the PCSK9 was identified as a cause of hypercholesterolemia in patients in FH (30). While numerous drugs were investigated, only evolocumab and alirocumab have been licensed to date. Both were initially licensed for the treatment of hypercholesterolemia by the EMA in 2015. Both went on to receive a broader label between 2018 and 2019 for the secondary prevention of CVD (4,5). Phase III trials of another PCSK9 inhibitor bocoizumab were stopped early because of the development of anti-drug antibodies (31).

Both evolocumab and alirocumab are self-administered by subcutaneous injection but the dosing regimens differ (4,5). Evolocumab is given at a dose of 140mg every two weeks. A therapeutically equivalent regimen at a dose of 420mg once monthly is also licensed (5). However, it is more expensive given the need to use three injections every four weeks rather than two under the fortnightly dosing schedule. Alirocumab may be initiated at a dose of 75mg or 150mg every two weeks or 300mg once monthly. The dose can be titrated as required to meet the target LDL-C level (4).

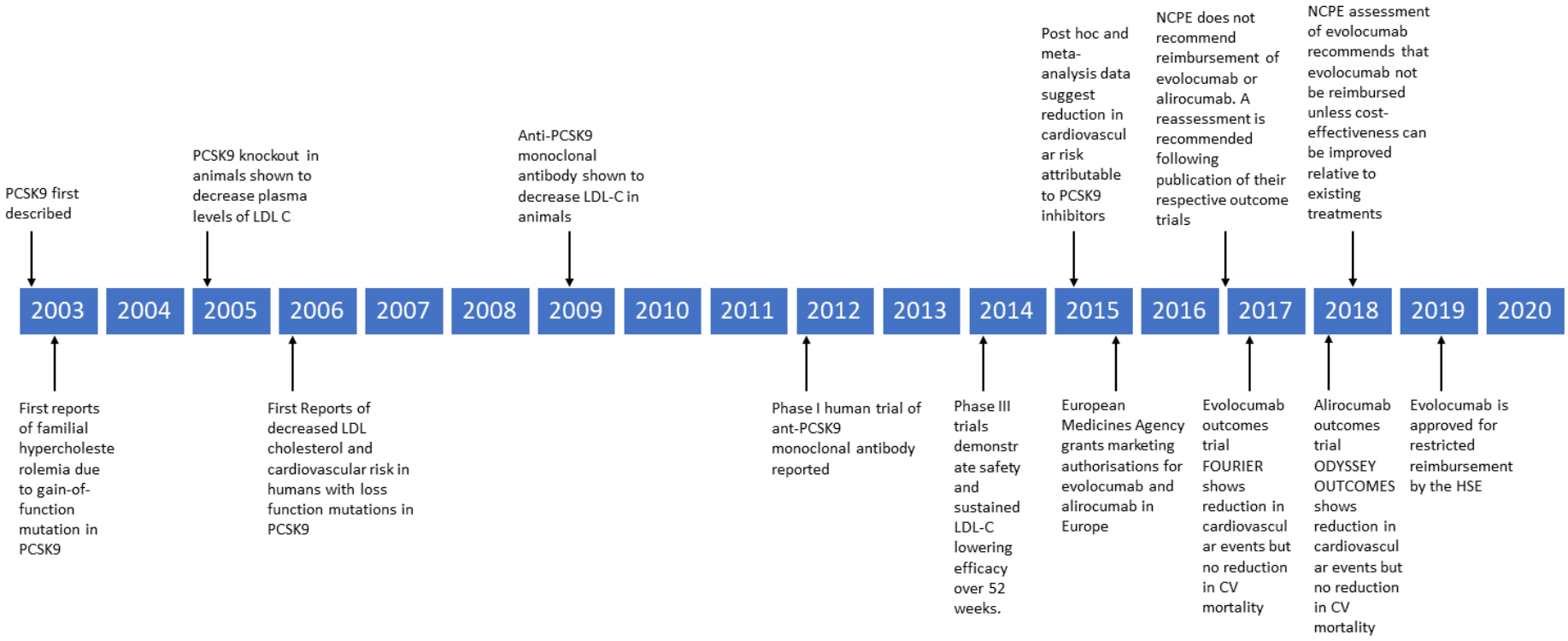


Figure 2.1 Timeline of PCSK9 inhibitor therapy from discovery to the present time.

Based on data by Page and Watts (30).

2.3.2.1 European treatment guidelines

European Society of Cardiology guidelines are followed by many clinicians in Ireland (32). Guidelines assess total CVD risk and stratify patients into four risk groups (low, moderate, high and very high). The target LDL-C reduction varies depending on the patient's risk group and is described in Table 2.2. All secondary prevention patients are classified as very high risk. The LDL-C target for very high-risk patients is <1.4mmol/L. The target was lowered in 2019 following the publication of cardiovascular outcomes trials for PCSK9 inhibitors [1]. The previous target was 1.8mmol/L.

[1] These trials are discussed in detail in Chapter 3.

Table 2.2 ESC/EAS risk categories and LDL-C targets (32)

ESC/EAS Risk Level	Characteristics	LDL-C Treatment Target
Very high risk	• Documented ACVD, either clinical or unequivocal on imaging, previous MI, ACS, stable angina, coronary revascularisation, and other revascularisation procedures, IS, TIA, PAD. • DM with target organ damage or at least three major risk factors or early onset of T1DM of long duration (>20 years) • Severe CKD (eGFR <30/min/1.73m^2) • A calculated 10-year risk SCORE ≥10% for fatal CVD • FH with ASCVD or with another major risk factor.	<1.4mmol/L and a reduction that achieves ≥50% LDL-C reduction from baseline
High Risk	• Markedly elevated single risk factors, in particular TC >8 mmol/L, LDL-C >4.9 mmol/L or BP >180/110 mmHg. • Patients with FH without other major risk factors. • Patients with DM without target organ damage, with DM duration >10 years or another additional risk factor • Moderate CKD (eGFR 30-59ml/ min/1.73m2) • A calculated SCORE ≥ 5% and <10% for 10-year risk for fatal CVD	<1.8mmol/L and a reduction of at least 50% if the baseline is between 2.6 and 5.2 mmol/L
Moderate Risk	• Young patients (T1DM < 35 years; T2DM < 50 years) with DM duration <10 years, without other risk factors. • SCORE ≥1% and <5% 10-year risk for fatal CVD	<2.6mmol/L
Low Risk	• Total Score <1%	<3mmol/L

ACS, Acute Coronary Syndrome; ASCVD, Atherosclerotic Cardiovascular Disease; BP, Blood Pressure; CVD, Cardiovascular Disease; CKD, Chronic Kidney Disease; DM, Diabetes Mellitus; eGFR, Estimated Glomerular Filtration Rate; ESC, European Society of Cardiology; EAS, European Atherosclerosis Society; FH, Familial Hypercholesterolemia; IS, ischaemic Stroke; LDL-C, low density lipoprotein cholesterol; mmHg, millimetre of mercury; mmol/L, millimole/Litre; MI, Myocardial Infarction; PAD, peripheral artery disease; T1DM, Type 1 diabetes mellitus; T2DM, Type 2 diabetes mellitus; TC, Total Cholesterol; TIA, Transient Ischemic Attack.

To achieve the targets, the following treatment pathway is recommended:

1. First line: High intensity Statin
2. Second line: High intensity Statin + Ezetimibe
3. Third Line: High intensity Statin + Ezetimibe + PCSK9 inhibitor (32).

No preference is given to either PCSK9 inhibitor.

2.3.2.2 Reimbursement status in Ireland and other countries

Marketing authorisation holders for both evolocumab and alirocumab submitted reimbursement applications to the HSE in 2015. Full HTAs were recommended by the NCPE following rapid reviews. At the time of the initial assessment, the results of CV outcomes trials were not available (10,11). In both cases ICERs far exceeded a threshold of €45,000 per QALY. Reimbursement was not recommended. However, reassessments were recommended following publication of their respective CV outcomes trials (10,11). In a resubmission for evolocumab, ICERs ranged from €207,000 to €908,000 in the NCPE preferred basecase for the secondary prevention population at an LDL-C of 4mmol/L (9). No further subgroup analysis is reported in the secondary prevention population. Therefore, it is unclear how cost-effectiveness varied across subgroups (9). Reimbursement was not recommended for evolocumab unless cost-effectiveness could be improved relative to existing treatments. Following confidential price negotiations, evolocumab was approved for reimbursement in July 2019. However, reimbursement is restricted to those with a history of MI or coronary artery bypass graft and have an LDL-C persistently greater than 4mmol/L or primary prevention patients with a diagnosis of HeFH at the same threshold. Further criteria ensure that other cholesterol lowering medications are optimised in the first instance. Applications for reimbursement are strictly controlled by the Medicines Management Program of the HSE (12). Alirocumab has not been reimbursed in Ireland to date.

In the UK, both the National Institute for Health and Care Excellence (NICE) and the Scottish Medicines Consortium (SMC) found PCSK9 inhibitors to be cost-effective in subgroups of the licensed population (33–35). However, the reimbursement criteria are broader than the Irish criteria with thresholds of 4mmol/L for the secondary prevention population and a lower threshold of 3.5mmol/L for a recurrent disease population. These UK HTAs were conducted before publication of the PCSK9 inhibitor outcomes trials. The full criteria are outlined in Table 2.3.

Table 2.3 PCSK9 inhibitor reimbursement criteria for NICE and the SMC (33–36)

	LDL-C Threshold for Reimbursement		
	Without CVD	**With CVD**	
		At High Risk of CVD[1]	**At very high risk of CVD** [2]
Primary non-familial hypercholesterolemia or mixed dyslipidaemia	Not recommended at any LDL-C concentration	Recommended only if LDL-C concentration is *persistently above* **4.0** mmol/L	Recommended only if LDL-C concentration is *persistently above* **3.5** mmol/L
Primary heterozygous familial hypercholesterolemia	Recommended only if LDL-C concentration is *persistently above* **5.0** mmol/L.	Recommended only if LDL-C concentration *persistently above* **3.5** mmol/L	

[1] High risk of cardiovascular disease is defined as a history of any of the following: acute coronary syndrome (such as myocardial infarction or unstable angina requiring hospitalisation), coronary or other arterial revascularisation procedures, chronic heart disease, ischaemic stroke, peripheral arterial disease.
[2] Very high risk of cardiovascular disease is defined as recurrent cardiovascular events or cardiovascular events in more than 1 vascular bed (that is, polyvascular disease).
CVD, cardiovascular disease; LDL-C, low-density lipoprotein cholesterol; mmol/L, millimole per Litre.

In the next chapter, a systematic review and evidence synthesis of the clinical evidence for PCSK9 inhibitors is conducted.

Chapter 3. Efficacy of PCSK9 inhibitors: systematic review and evidence synthesis

3.1 Introduction

In order to populate our economic model, estimates of the comparative efficacy of PCSK9 inhibitors and corresponding uncertainty are required. National guidelines recommend that these are derived from high quality randomised controlled trials sourced from a systematic review of the literature (37). Meta-analysis is a statistical technique which can be used to combine evidence from multiple sources to derive a common average effect estimate across trials. Techniques such as network meta-analysis can be used to calculate relative effectiveness estimates between interventions where there is no direct evidence between them or there is a mixture of direct and indirect evidence.

Using these methods, the aim of this chapter is to quantify the efficacy of PCSK9 inhibitors class on the risk of CV events versus placebo and other lipid modifying therapies. Also, to quantify the relative effectiveness of alirocumab and evolocumab for incorporation in an economic evaluation of PCSK9 inhibitors presented in Chapter 7. Results of the systematic review will inform a meta-analysis of the efficacy of PCSK9 inhibitors versus placebo and other lipid lowering therapies. It will also inform a network meta-analysis comparing evolocumab and alirocumab to placebo and other lipid modifying therapies.

3.1.1 Chapter outline

First, a description and critique of existing systematic reviews and evidence syntheses of PCSK9 inhibitors in the literature is provided and their applicability to our research question is examined. Then, the methods and results of our de novo analysis are described before concluding with a discussion of our findings.

3.2 Literature review of previous meta-analysis and network meta-analysis of PCSK9 inhibitors

Previous literature in this area can be broadly classified into two group based on whether they were completed before or after publication of the large cardiovascular outcomes trials, FOURIER and ODYSSEY OUTCOMES (6,7).

3.2.1 Pre-outcomes trials

DESCARTES and ODYSSEY LONG TERM are two examples of some of the first phase III trials conducted (38,39). DESCARTES examined the efficacy and safety of evolocumab versus placebo in 905 patients stratified by cardiovascular risk (39). Patients were followed for 52 weeks. ODYSSEY LONG TERM examined the efficacy and safety of alirocumab versus placebo in 2,341 patients at high risk of CV events over 78 weeks (38). Both trials showed the ability of both

PCSK9 inhibitors to reduce baseline LDL-C but were not powered to detect treatment effects on CV outcomes (38,39). For meta-analyses conducted prior to the publication of CV outcomes trials, the number of CV events was small even after aggregating results across trials. This meant there was substantial uncertainty surrounding effect estimates (40). Initial systematic reviews conducted primarily relied upon data reported as safety outcomes. Therefore, many relevant outcomes were not reported in a form amenable for analysis. For example, many of these early trials reported coronary heart disease (CHD) mortality and did not report CVD mortality (38,39). It is of concern, that when estimating the effectiveness of PCSK9 inhibitors on the CVD mortality endpoint, many systematic reviews included trials reporting the number of deaths due to CHD in each arm instead. The definition of CHD is much narrower than that of CVD (which also encompasses stroke and other causes of vascular disease such as deep vein thrombosis in addition to CHD). This disparity means that many deaths due to CVD (that are not attributable to CHD) are not accounted for in the analyses. The Cholesterol Treatment Trialists Collaboration (CTTC) meta-analysis indicates that CHD mortality accounts for less than 50% of the total number of CV deaths (41). The CTTC estimated a rate ratio per 1 mmol/L reduction in LDL-C of 0.80 for CHD death versus 0.88 for any CVD death (41). Therefore, inclusion of CHD event counts is expected to lead to biased estimates of CVD mortality treatment effects.

I have previously highlighted concern in this area in response to the review conducted by Navarese et al (40,42). The disparity can have a large impact on the results of meta-analyses. For two of the 24 trials analysed by Navarese et al, (ODYSSEY LONG TERM (38) and ODYSSEY COMBO I (43)), the authors inputted the number of deaths due to CHD for each arm rather than the number of deaths due to CVD. Since the total numbers of CVD deaths are not reported or computable from the published trial results referenced, these two trials should not have been included in the analysis of this endpoint. Although the meta-analysis included 24 trials, 49.2% of this endpoint was weighted to the results of these two trials alone. Therefore, their inappropriate inclusion has an impact on the results of this analysis. An odds ratio (OR) of 0.77 (95% Confidence Interval [CI] 0.26 to 2.35) was calculated when these trials are excluded compared to an OR of 0.50 (95% CI 0.23 to 1.10) reported in the Navarese et al analysis (40). It was concluded that the potential beneficial effect of PCSK9 inhibitor therapy on CVD death was overestimated given the data available at the time of the analysis. This is of concern as the results of this meta-analysis has been used to inform cost effectiveness evaluations submitted to national HTA agencies (33,34) Navarese et al is not alone in not differentiating between CHD and CVD deaths (44,45). However, in later analyses when larger outcomes trials are included,

the relative weight given to these trials is smaller, reducing the bias introduced through their inclusion (44,45). (See section 3.2.2 for a description of these reviews).

Overall, Navarese et al concluded that PCSK9 inhibitors reduced all-cause mortality and the rate of MI without an increase in the risk of serious adverse events (40). While the trials included over 10,000 participants, the number of CV events was very small. For example, conclusions regarding all-cause mortality were based on only 25 deaths.

Lipinski et al report the first network meta-analysis comparing PCSK9 inhibitors; ezetimibe; PCSK9 inhibitors in combination with ezetimibe; and placebo (46). Studies were assessed for risk of bias using the original Cochrane Risk of Bias tool. The GRADE appraisal tool was applied, albeit on a study level rather than on the recommended outcome level. The network meta-analysis was limited to biochemical outcomes and only a meta-analysis of PCSK9 inhibitors versus placebo was conducted for other clinical outcomes. Results for CV outcomes were similar to those reported by Navarese et al (40). Lipinski et al also included neurocognitive adverse events as an outcome measure. Data was only available in a subset of studies. The estimates showed increased odds in the PCSK9 inhibitor arm for neurocognitive adverse events (OR 2.34; 95% CI 1.11 to 4.93). However, these results have not been replicated in large trials designed to examine this issue (47). In a large neurocognitive sub-study of FOURIER (entitled EBBINGHAUS), there was no significant difference in the rate of neurocognitive events between trial arms (47). The corresponding neurocognitive study for alirocumab is ongoing (48).

Toth et al conducted a systematic review and network meta-analysis of evolocumab versus alirocumab and other therapies on lipid levels only (49). They calculated a treatment difference of -13.63% (95% CI -22.43 to -5.33) in favour of evolocumab on the extent of LDL-C reduction. However, statistical heterogeneity and population heterogeneity was high, limiting confidence in the results. Adverse events were not included in the network meta-analysis, but the latter was included in meta-analysis of individual therapies versus placebo. CVD outcomes were not assessed.

3.2.2 Post-outcomes trials

The publication of FOURIER, in March 2017, increased the power of meta-analysis to detect treatment effects. Schmidt et al were the first to incorporate these results in meta-analysis in their Cochrane Review (44). Bocoizumab trials (SPIRE 1 and 2) were also included despite the trials being stopped early due to the development of anti-drug antibodies (31). Many evolocumab trials had a 12-week duration and were excluded (the minimum trial duration specified in the systematic literature review inclusion criteria was 24 weeks). They concluded

that the mean percent change from baseline in LDL-C at 6 months compared to placebo was -53.86% (95% CI -58.64 to -49.08; GRADE: moderate). PCSK9 inhibitors decreased the incidence of CVD (OR 0.86; 95% CI 0.80 to 0.92; GRADE: moderate) without affecting all-cause mortality (OR 1.02; 95% CI 0.91 to 1.14; GRADE: moderate) compared to control arms. There was an increase in the risk of any adverse events for PCSK9 inhibitors, but this was driven by trials examining bocoizumab (OR 1.08; 95% CI 1.04 to 1.12; GRADE: low). Extensive subgroup analysis was conducted for biochemical outcomes but there was insufficient information for subgroup analysis on clinical outcomes. No statistically significant subgroup effect was identified.

Subsequent systematic reviews have incorporated ODYSSEY OUTCOMES (the alirocumab CV outcomes trial (45,50,51). The primary outcome of a systematic review and analysis conducted by Du et al was Major Adverse Cardiovascular Events (MACE) (50). MACE was predefined as a composite of CVD death, non-fatal myocardial infarction and non-fatal stroke (50). Where the trial definition was inconsistent, the number of MACE was computed by adding the number of patients who reported each of the three independent outcomes. However, we question this approach as it leads to double counting of patients who experience more than one type of event over the course of the trial. Du et al included bocoizumab and other experimental PCSK9 inhibitors but the weight given to these drugs was very small given the very limited experience with these drugs in trials to date (50). Random-effects meta-analysis was conducted but no further methodological details were described. Their meta-analysis showed that PCSK9 inhibitors were associated with a reduced risk of MACE (Risk Ratio (RR) 0.84; 95% CI 0.79 to 0.89; GRADE: moderate). No significant reduction in risk was observed for CV death or all-cause mortality, but the quality of evidence was low. They concluded that PCSK9 inhibitors reduce the risk of non-fatal MI as well as stroke. No safety endpoints were analysed.

Guedeney et al included trials published up to March 2018 (45). Primary efficacy endpoints included the effect of PCSK9 inhibitor therapy versus control on all cause death, CVD death, MI and stroke. No significant treatment effect was observed for all-cause mortality or CVD deaths, but PCSK9 inhibitor therapy was shown to reduce the risk of MI (RR 0.80; 95% CI 0.74 to 0.86) and ischemic stroke (RR 0.78; 95% CI 0.67-0.89) compared to control. No significant difference was observed in the risk of neurocognitive adverse events. Difference in effects between alirocumab and evolocumab were assessed by conducting subgroup analysis and calculating heterogeneity statistics and interaction p values between both sets of trials. This method can be used to determine if there is a statistically significant difference in treatment effect between trials assessing these drugs. However, it is not an indirect comparison (as the difference in treatment effect between drugs is not quantified and the causes of heterogeneity are not

considered). Using this method, a significant interaction p value was observed for evolocumab versus alirocumab for the all-cause mortality endpoint (p=0.03), but was not observed for CVD death, MI or ischemic stroke.

Khan et al conducted a systematic review and meta-analysis analysing the association of baseline low density LDL-C with total and CVD deaths (51). They conducted a meta-regression examining the association between mean baseline LDL-C, extent of LDL-C reduction and relative risk of CVD deaths and all-cause mortality. After removal of SPIRE trials, they found a RR of 0.97 (95% CI 0.94-0.99) per unit increase in baseline LDL-C expressed as mg/dL. However, the mean baseline LDL-C in most studies was low. This mean that studies that did have higher baseline LDL-C values had a considerable impact on the regression. In scenario analysis, they included the RR and mean baseline LDL-C of the subgroup (from ODYSSEY OUTCOMES) with a baseline LDL-C >100 mg/dL in the meta regression. However, I note a critical error. They assumed the mean baseline LDL-C of this group to be 101mg/dL. However, this value represents the minimum baseline LDL-C of this cohort rather than the mean. The actual mean baseline LDL-C of this group is reported, in the supplementary appendix of Schwartz et al, as approximately 130 mg/dL (6). Correcting this error would lead to a very large shift towards the null. Comparable flaws were evident in their analysis of the all-cause mortality endpoint.

Network meta-analyses have also been conducted. However, these reviews have primarily compared PCSK9 inhibitors versus other classes of other lipid lowering therapy (rather than comparing between individual PCSK9 inhibitors in the same class). Khan et al conducted a Bayesian network meta-analysis (52). The authors state that they included randomised controlled trials (RCTs) which compared PCSK9 inhibitors ezetimibe, statins, placebo or combinations of these in subjects with hypercholesterolemia. However, in the network diagram, there are no trials connecting PCSK9 inhibitors and placebo. Instead the authors have attributed the efficacy of PCSK9 inhibitors versus placebo as the effect versus statins. This is incorrect as statins were maintained as background therapy in both PCSK9 inhibitor and comparator arms in the major PCSK9 inhibitor trials. This error means that the results from this analysis are flawed and cannot be relied upon.

Zhau et al (53) performed an alternative network meta-analysis using a frequentist approach. Nodes included ezetimibe, statins, placebo and PCSK9 inhibitors. There was direct evidence for each comparison except PCSK9 inhibitors versus statins. Homogeneity is required for results of network meta-analysis to be valid, but this was not assessed by the authors. Studies were restricted to those published since 2000 which excludes many relevant statin trials. Risk of bias

was assessed using the Cochrane Risk of Bias tool on a study level. The quality of the evidence was not assessed. Statins had the highest ranking for the probability of reducing the risk of CVD events at 60.6%, followed by PCSK9 inhibitors at 37.1%, but there was no statistically significant difference in efficacy observed between the two classes of drugs (OR 0.83; 95% CI 0.75 to 0.91). In a meta-analysis comparing PCSK9 inhibitors to placebo, a reduction in the risk of CVD death was not observed (OR 0.99; 95% CI 0.87 to 1.13). The point estimate decreased marginally after incorporating indirect evidence through the network meta-analysis, but results did not reach statistical significance (OR 0.94; 95% CI 0.76 to 1.17).

The results of a network meta-analysis comparing alirocumab and evolocumab have been published in abstract form (54). Thirty trials were included. They concluded that alirocumab was associated with a reduction of all-cause death (RR 0.80; 95% CI 0.66 to 0.97) but not CVD death (RR 0.83; 95% CI 0.65 to 1.05). The authors detected no significant differences in the risk of MI or stroke between alirocumab and evolocumab.

Other reviews examined the effect of PCSK9 inhibitors on the risk of incident type II diabetes (55,56) and on circulating hs-CRP levels (57).

3.2.3 Summary of literature review

Numerous systematic reviews and evidence synthesis have been conducted but none of the identified studies have included all relevant outcomes for our study as defined in Section 3.3.1.4. A critical issue here is heterogeneity and consideration of whether relative treatment effects are consistent across relevant subgroups. However, this has not been assessed. Further, none of the published analysis used hazard ratios (HRs) to measure PCSK9 inhibitor treatment effects. Therefore, we considered that a de novo analysis was required to address the research question. The publication of the updated Cochrane handbook in July 2019 will facilitate the incorporation of recent methodological developments and consensus to be incorporated into study (58).

3.3 Methods

The analysis was designed with reference to the most recent edition of the Cochrane Handbook for Systematic Reviews of Interventions (58). The review and analysis is reported in line with the PRISMA statement for systematic reviews and meta-analysis (59).

3.3.1 Criteria for considering studies for this review

3.3.1.1 Types of studies

Double-blind parallel group randomised controlled trials (RCTs) with a median follow up time of at least two years were included. Studies were required to have a minimum of 500 participants.

Studies were included whether published as full text articles, abstracts only and from regulatory sources. Open label studies were excluded.

3.3.1.2 Types of participants

Studies assessing adults 18 years or over, with or without a prior history of CVD, were included. No restriction was placed on lipid levels or form of hypercholesterolemia or co-morbidities. Pre-specified subgroups included examined primary versus secondary prevention, baseline LDL-C and previous history of CVD.

3.3.1.3 Types of interventions and comparisons

Trials were included that examined alirocumab and/or evolocumab (which are the only licensed PCSK9 inhibitors in Europe to date). Other investigatory drugs which have not received regulatory approval (such as bococizumab) were excluded. No restrictions were placed on the form of background lipid lowering therapy or intervention doses. Trials were included that compared the named interventions to other licensed lipid lowering therapy or to placebo.

3.3.1.4 Types of outcome measures

3.3.1.4.1 Primary outcomes

The primary outcomes are those that are required to populate the treatment effectiveness parameters for our economic model.

1. Time to non-fatal MI
2. Time to non-fatal ischemic stroke
3. Time to non-fatal haemorrhagic stroke
4. Time to CV death.

3.3.1.4.2 Secondary outcomes

The following secondary outcomes were defined:

1. Time to all-cause death
2. Time to CHD death.
3. Risk of serious adverse events (as defined in the pivotal trial publication).
4. Health related quality of life (HRQoL) (as defined by validated quality of life measures or instruments used in each trial).

The effectiveness of PCSK9 inhibitors on reducing mortality has previously been identified as a critical parameter of the economic model (9). Therefore, the outcomes all-cause death and CHD death were included to examine outcomes associated with using both the narrower and broader definitions of mortality. The benefits of any drug must always be balanced versus their adverse

effects. Serious adverse events were chosen as this endpoint is most likely to be clinically and economically relevant and reported consistently across trials. HRQoL is an important outcome to inform utility estimates in an economic model and was included for this reason.

3.3.1.4.3 Reporting of outcomes in relevant studies

It was envisaged that outcomes will be reported differently across trials and that all endpoints may not be reported. It was pre-specified that, if the non-fatal event type is not reported, the results for the composite non-fatal and fatal event would be substituted provided that the number of deaths due to the event type is below 10% of the total number of events. For example, if non-fatal MI is not reported, the results for any MI (defined as both non-fatal and fatal MI) could be substituted (provided the total number of deaths due to MI is below 10% of the total number of MI reported). The 10% threshold was chosen arbitrarily.

3.3.2 Search methods for identification of studies

The search strategy was extensive. MEDLINE, EMBASE, Web of Science and CENTRAL database were searched from 2005 to December 3rd, 2019. The search strategy included a mixture of PCSK9 inhibitor/CV search terms and randomised controlled trial search terms. Search terms were sourced from a published Cochrane Review of the efficacy of PCSK9 inhibitors (44) complemented with search terms previously provided by David Mockler (Medical Librarian, Trinity College Dublin). In line with the strategy adopted by Schmidt et al, papers published before 2005 were excluded (44). As, PCSK9 was only identified as a potential target in 2003, it is impossible to find relevant studies before this time. The full search strategy is documented in Appendix 1.1.1. In a post-protocol decision, abstract books of conferences from the following society and associations were searched: European Society of Cardiology, American College of Cardiology, American Heart Association, (US) National Lipid Association, and European Atherosclerosis Society. For this, the terms: PCS, proprotein, evolocumab and alirocumab, were used for the years 2015 to 2019. Regulatory information sources including European product assessment reports from the EMA, and clinical trial reports from ClinicalTrials.Gov (www.clincialtrials.gov) were also searched. References from included studies were screened for relevant studies. Articles were restricted to those published in English.

3.3.3 Citation management

References identified were imported into Endnote® and transferred to Covidence®. Duplicates were both identified manually throughout the review process and systematically searched for using software in Endnote® and Covidence®. Next, the library was screened by title and abstract in order to identify any potentially eligible studies for full text review. The full texts of all

potentially eligible citations were retrieved and their compatibility with the eligibility criteria assessed. For quality assurances purposes, 10% of full-text articles were screened in duplicate by Laura McCullagh.

A data extraction form was developed based on the standard Cochrane data extraction form. Items recorded included information on study design, quality, interventions, participants and outcomes and measures of uncertainty. Outcomes data was checked in duplicate by Laura McCullagh.

3.3.4 Assessment of risk of bias in included studies

Bias has been defined as "a systematic deviation from the effect of intervention that would be observed in a large randomised trial without any flaws" (60). In recent years, the Cochrane risk of bias tool has been recommended. However, several limitations have been identified including potential for confusion because of limited guidance. A further concern was moderate reliability of judgements across reviewers (60,61). This review adopts an updated version of the tool called Cochrane Risk of Bias 2 (ROB 2) which addresses some of these limitations. While traditionally risk of bias was assessed at a study level, ROB 2 assesses bias separately for each relevant outcome. This approach acknowledges that for each study, the risk of bias may be dependent on the outcome. For example, a subjective outcome, such as a self-reported pain score, may be more susceptive to bias than an objective outcome measure such as death (60).

ROB 2 assesses bias across five domains:

1. Bias arising from the randomization process;
2. Bias due to deviations from intended interventions;
3. Bias due to missing outcome data;
4. Bias in measurement of the outcome;
5. Bias in selection of the reported outcome.

An important addition of the tool is a series of signalling questions for each risk of bias domain. An algorithm maps question response to a proposed judgement which may be overruled by the reviewer. These changes increase consistency and adherence to the intended use of the tool. Risk of bias was assessed in duplicate by Laura McCullagh.

3.3.5 Measure of treatment effect

No measure of treatment effect for HRQoL was pre-specified (given the expected limited number of results and the numerous potential methods of reporting). We reported results as hazard ratios for time to event endpoints, relative risks for binary outcomes. Published meta-

analysis, identified to date, have used the relative risk or odds ratio. However, the relative risk of an event over time depends on the length of time observed. For example, a trial is not needed to calculate the relative risk of death of evolocumab versus alirocumab over a lifetime time horizon – the relative risk will always be 1 (as all patients will have died). Hazard ratios account for the differences in the rate of events over time. They rely on the proportional hazard's assumption. This assumes while the hazard rate may change in both arms over the study, the ratio of the hazards is constant over time. Hazard ratios are a measure of relative treatment effect rather than absolute treatment effect. Therefore, even if they are assumed to be constant across subgroups, the absolute benefits of a drug will be greater in subgroups with the highest baseline risk. All time-to-event endpoints were required to be reported as hazard ratios (with the associated 95% confidence intervals) or reported in such a way that these can be derived. For example, if Kaplan-Meier curves are available for both trials' arms, it was planned that hazard ratios would be derived by digitising the curve and deriving the hazard ratio in this way. This is the first meta-analysis of PCSK9 inhibitors to measure treatment effectiveness using the hazard ratio to the best of our knowledge.

3.3.6 Data synthesis

Two forms of data synthesis were used in this study. First, the efficacy of the PCSK9 inhibitor class versus placebo and other lipid lowering therapies was estimated by conducting a meta-analysis combining evidence from all relevant trials using the R package 'meta' (62,63). Next, an adjusted indirect treatment comparison was conducted by comparing the efficacy of alirocumab versus evolocumab. Calculations for adjusted indirect comparisons were conducted in Microsoft Excel® (64).

3.3.6.1 Meta-analysis

3.3.6.1.1 Fixed versus random-effects

There are two main types of meta-analysis – fixed-effects and random-effects. A fixed-effect meta-analysis assumes that all studies in the meta-analysis are estimating the same underlying treatment effect (58). In contrast, the random-effects approach assumes that studies are estimating different treatment effects, even though they are related, through a normal distribution across studies. The additional parameters mean than confidence intervals under the random effects approach are wider compared to the fixed-effects approach. The choice to conduct a fixed-effects versus a random-effects meta-analysis is discussed extensively in the Cochrane handbook (58). Clinical, methodological and statistical heterogeneity should be assessed. For this review, random-effects method was pre-specified given the clinical

heterogeneity in combining treatment effects estimates from different drugs in the PCSK9 inhibitor class and the expected differences in outcomes across trials.

The assumptions underlying the standard random-effects model are outlined by Veroniki et al (65). Briefly, it assumes that the estimated treatment effect size from the ith study y_i is

$$y_i = \theta_i + \varepsilon_i \,,$$

Where the study specific random error ε_i, and the underlying true effect sizes in the individual studies θ_i are normally distributed as $\varepsilon_i \sim N(0, v_i), \theta_i \sim N(\mu, \tau^2)$.

The random-effects estimated overall effect size can be denoted as:

$$\widehat{\mu_{RE}} = \frac{\sum y_i w_{iRE}}{\sum w_{iRE}} \text{and } var(\widehat{\mu_{RE}}) = \frac{1}{\sum w_{iRE}} \text{;with weights } w_{i,RE} = \frac{1}{(v_{i+}\widehat{\tau^2})}.$$

These weights are the inverse of the estimated total study variances.

3.3.6.1.2 Methods of Random-Effects

The most common method of random-effects analysis is known as Der Simonian and Laird inverse variance approach (66). This method adopts the Wald Type normal distribution (to estimate confidence intervals) and the Der Simonian and Laird method (to estimate the between study variance). The advantage of the Wald Type method is that calculations are straightforward. However, it is acknowledged that confidence intervals are slightly too narrow for the estimated degree of heterogeneity. This is because it assumes that we have a true estimate of the between-study heterogeneity when in reality we just have an estimate of this parameter (65). Acknowledging this additional uncertainty, the Cochrane handbook recommends the Hartung and Knapp adjustment, which more accurately preserves the type 1 error rate (the false positive rate). Under this method a variance estimator q is proposed so that the test statistics for the treatment effect t is distributed with k-1 degrees of freedom (which further widens the confidence interval in order to encompass this heterogeneity) (58,67,68). However, when there are only two independent studies, confidence intervals are over-conservative and can be extremely wide. Simulation studies have shown that there is very little power to detect a true positive treatment effect in these cases (67). In a recent review of alternative approaches, Veronicki et al tentatively suggested the Hartung and Knapp adjustment as the standard approach, at least in a meta-analysis with five or more studies (65). Others recommend a Bayesian approach (69). Given the lack of consensus in methods, the Wald Type normal approach is adopted which is considered standard practice whilst bearing in mind its limitations. (65)

There are also different ways to estimate between study variation. While different methods can lead to substantial differences in the estimates of heterogeneity, they rarely have important implications for estimating the summary effects (58). The default method in the R package (*meta*) is the standard Der Simonian and Laird method, but additional estimators including the Paule-Mandel, restricted maximum likelihood, maximum likelihood, Hunter-Schmidt, Sidik-Jonkman, Hedges and the Empirical Bayes estimator are available. Both the Cochrane handbook and the *meta* package recommend the Paule-Mandel and maximum likelihood methods based on the recent simulation studies (58,62,70). Therefore, we adopt the Paule-Mandel approach in our analysis.

Following international best practice, national guidelines recommend that prediction intervals are reported, in addition to confidence bounds, when random-effects models are used (37,71). However, a minimum of three studies are required in each meta-analysis for computation of prediction intervals (72). Further, intervals are difficult to compute when the number of studies is small (<10), therefore, they are not computed in this analysis (58).

3.3.6.2 Network meta-analysis

Network meta-analysis describes meta-analysis across a network of evidence with multiple comparisons. Hoaglin et al suggest that the term network meta-analysis can be used to describe any analysis where "the evidence involves more than two randomised controlled trials and more than two interventions" (73). It should be noted that other authors use the term 'network meta-analysis' regardless of the number of trials (74).

Of interest in this study, is the comparative efficacy of alirocumab versus evolocumab. However, in the absence of large clinical trials comparing the two treatments directly, indirect methods must be used in order to estimate the treatment effect. If a common comparator can be found across relevant trials, it may be possible to conduct an indirect treatment comparison.

Naïve indirect comparisons are possible but are not recommended. In this study we used the Bucher's adjusted indirect comparison method to compare evolocumab versus alirocumab via placebo (75). This method preserves randomisation between treatment groups. Bucher's method can estimate relative treatment effects for star pattern in networks (where interventions are joined together by a common comparator of interest).

In a network containing evolocumab (E), alirocumab (A) and placebo (P), the treatment effect of alirocumab versus evolocumab (T_{AE}) may be calculated as

$$T_{AE} = T_{AP} - T_{EP}$$

where treatment effects and standard errors are calculated as the natural log (ln) of the HR. The standard error (SE) is calculated as:

$$SE(T_{AE}) = \sqrt{SE(T_{AP})^2 + SE(T_{EP})^2}.$$

Results are exponentiated to obtain results on the natural scale.

More complex networks are possible. Networks which have closed loops (i.e. interventions in the network have both direct and indirect comparisons) are deemed mixed treatment comparisons (76). Mixed treatment comparisons are important even if there is direct evidence connecting the interventions of evidence as the inclusion of indirect evidence can lead to a more precise estimate of treatment effects. In a mixed treatment comparison, the transitivity or consistency assumption is required. This means that there should be no discrepancy between direct and indirect comparisons. For example, if alirocumab, evolocumab and placebo were connected by means of a closed loop through a trial directly comparing evolocumab and alirocumab– the indirect comparison of alirocumab and evolocumab through placebo must be consistent with the hypothetical direct trial evaluating the two PCSK9 inhibitors.

The decision on whether to proceed with an adjusted indirect comparison or network meta-analysis should be taken after assessment of the similarity assumption – trials contained in networks should be methodologically and clinically similar (76). But there are inherent differences in all trials, there is no commonly accepted definition which defines what quantifies an acceptable difference (73) At the very least, it requires that all trials in the network are comparable across features which are effect modifiers (76). But this is difficult to assess if the effect modifiers are themselves unknown. Without prior knowledge, conducting multiple subgroup analyses to identify them can lead to multiplicity. Further even if modifiers are known, they may not be measured or reported in trial reports. When treatment effect modifiers are known, meta-regression can be used to adjust for interaction of co-variates to remove bias but it requires a large number of studies or access to patient-level data to be successful (76).

The methods of network meta-analysis described here adopt the frequentist approach (which presents results as point estimates with 95% confidence intervals). This can be interpreted that, with repeated sampling the interval would contain the true population parameter 95% of the time (76). More complex methods are available including Bayesian network meta-analysis. However, with simple networks and few trials, they provide very similar results to the traditional frequentist approach. Therefore, for this analysis, the Bucher method was considered sufficient.

3.3.7 Subgroup analysis and assessment of heterogeneity

Heterogeneity encompasses clinical, methodological and statistical heterogeneity. Before performing a network or standard meta-analysis, heterogeneity should be assessed. Statistical heterogeneity was assessed using the I^2 statistic. However, as its discriminatory power is limited when there are a small number of studies, it was interpreted in the context of the other study characteristics. Pre-specified subgroups in our analysis include primary versus secondary prevention, baseline LDL-C and previous history of CVD.

3.3.8 Assessment of quality and generation of summary of findings tables

GRADE was used to assess the quality of the evidence for the primary outcomes using the methods described in GRADE handbook. The quality of clinical evidence reflects the extent that we are confident that an estimate of treatment effect is correct. GRADE is a systematic approach to retaining the certainty of evidence in systematic reviews and evidence synthesis (77).

Each outcome is assessed for quality independently. First, an initial quality grade is assigned dependent on the study type from which the treatment effect estimate was obtained (trials or observational studies). Next, the evidence is assessed across factors which can reduce or increase the GRADE applied. Factors include risk of bias, inconsistency, indirectness, imprecision, publication bias, magnitude of effect, presence of dose-response gradient and possible confounding. The possible final grades and their definitions are outlined in Table 3.1. Inevitably, there is a degree of subjectivity in decision to upgrade or downgrade. Therefore, the analysis was also conducted in duplicate, by Laura Mc Cullagh, for each of the primary outcomes. Disagreement was resolved by consensus or if required in reference to another reviewer (Michael Barry). Transparency is considered a strength of the GRADE approach where reasons for up or downgrading should be outlined with each quality assessment (77). Summary of findings tables were prepared using GradePRO software (78).

Table 3.1 Quality of evidence grades under the GRADE approach (77)

Grade	Definition
High	We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.
Very Low	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

3.4 Results

After exclusion of duplicates, 2,306 titles and abstracts were located. Results of the electronic search were supplemented with 29 citations located from other sources. A PRISMA flow chart is presented in Figure 3.1. Following screening and review of potentially eligible studies, we included 104 citations describing five studies. Each trial had multiple publications including abstracts, regulatory reports and multiple trial analyses. Some eligible trial results were presented in citations describing analyses of pooled studies. These citations were only included if results of the relevant individual trial were reported in addition to the aggregate results.

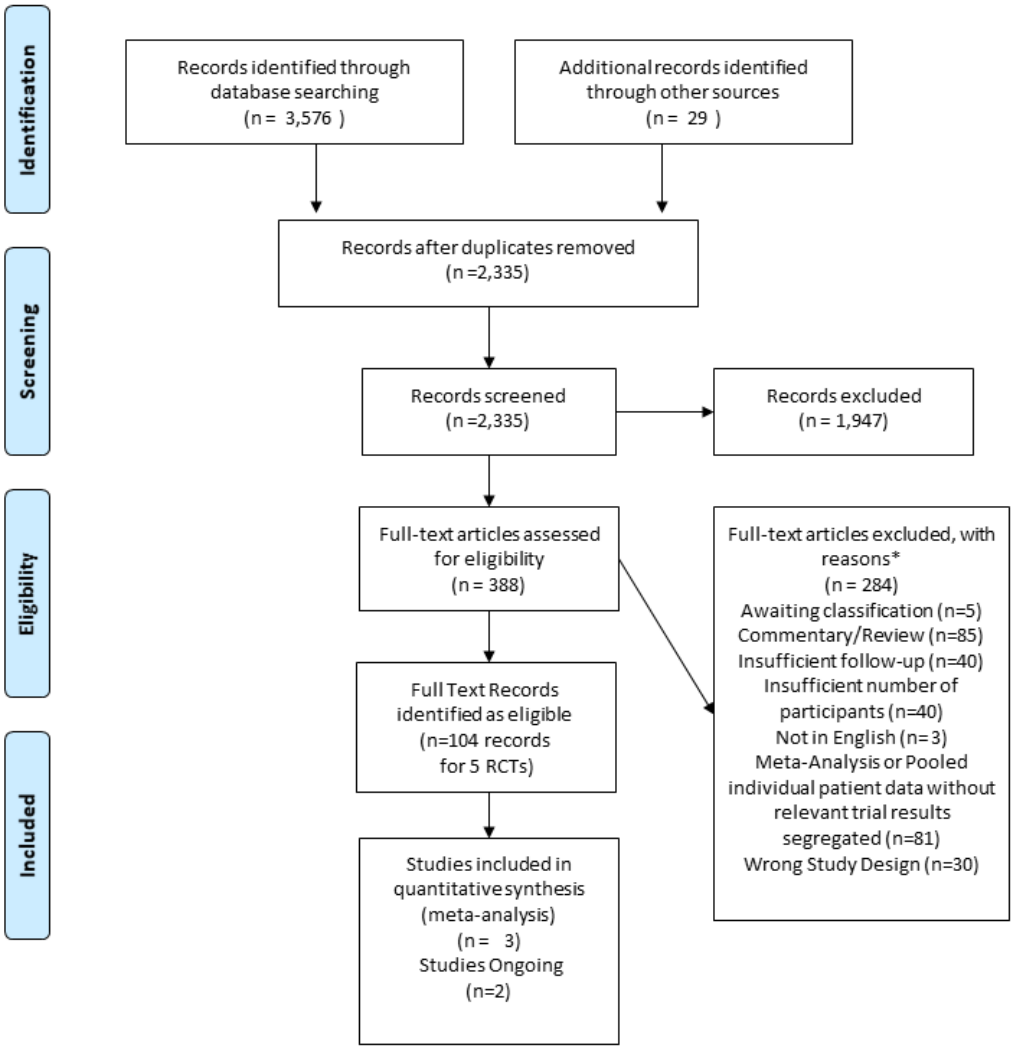


Figure 3.1 PRISMA diagram

3.4.1 Excluded studies

The number of citations excluded at full text stage was 284. Reasons for exclusion for the most relevant studies potentially eligible studies at the full text stage are provided in Appendix 1.1.2 Five citations were deemed "Awaiting Classification". For four of these, the probability of the citations meeting the inclusion criteria was low as the available records indicated they were likely describing reviews. However, as the full text could not be sourced, they were not excluded. There was a moderate probability that a further study (NCT02957682) (48) would meet the eligibility criteria. This study is a randomised double-blind placebo-controlled study to evaluate the effect of alirocumab versus placebo on neurocognitive function. The only citation available was a clinicaltrials.gov record. This limited the ability to compare study characteristics to the review criteria. The actual study enrolment is 2,176 patients (meeting the 500-patient threshold in this review). No relevant outcomes to the review were described as primary or secondary outcomes on its clinicaltrials.gov record. However, as important safety endpoints, it is likely that all-cause mortality and serious adverse events will be assessed in this safety study. The intended duration of follow-up was not stated. Many of the secondary outcomes are assessed at 96 weeks. Hence, a median follow-up of at least two years may be met if the trial duration is marginally longer. For these reasons, it was recorded as "Awaiting Classification" given its eligibility could not be ascertained. The estimated primary completion date of this study is March 2020.

3.4.2 Characteristics of included studies

Three completed studies met the eligibility criteria. – FOURIER (7,47,79–117), ODYSSEY OUTCOMES (6,118–141) and ODYSSEY COMBO II (3,135,142–157). Characteristics of these trials are summarised in Table 3.2. Each trial is described in turn below. No completed studies examined the primary prevention population. Therefore, the remainder of this review focuses on the secondary prevention subgroup only. Ongoing studies identified included VESALIUS-CV (158) and NEWTON-CABG (159).

3.4.2.1 Completed studies

FOURIER was a randomized, multi-centre double blind, placebo-controlled trial. It was designed to evaluate the efficacy and safety of evolocumab in reducing CV morbidity and mortality (7). Eligible patients were required to be aged 40 to 85 years, with a history of MI, stroke or clinically evident PAD as defined in the trial protocol. In addition, patients were required to have one major, or two minor risk factors (as defined in the trial protocol). The narrow trial eligibility criteria limit generalisability to the broader secondary patient population in Ireland. Patients were randomised in a 1:1 fashion to receive evolocumab (140mg every two weeks or 420mg

once monthly) or matching placebo as subcutaneous injections (N=27,564). The median length of follow-up was 2.2 years. Evolocumab significantly reduced the risk of the primary composite endpoint of CV death, MI, stroke, hospitalisation for unstable angina or coronary revascularisation. (HR 0.85; 95% CI 0.79 to 0.92; p <0.001). Evolocumab also reduced the risk of the key secondary composite end point of CV death, MI, or stroke (HR 0.80; 95% CI 0.73 to 0.88; p <0.001). FOURIER results present landmark analysis, of HRs before and after one year, in the supplementary appendix (adjusted for the extent of cholesterol reduction and without). These generally show a numerically increased treatment effect after one year compared to treatment effects derived before one-year follow-up [2]. However, this is very uncertain. Statistical tests for this assumption are not reported.

ODYSSEY OUTCOMES is a randomised multi-centre double blind, placebo-controlled trial, designed to evaluate if alirocumab would improve CV outcomes after a recent acute coronary syndrome (ACS) event, in patients receiving maximum tolerated statin therapy (6). Patients were required to have a recent ACS event (1 to 12 months prior to randomisation). Like FOURIER, the narrow trial eligibility criteria limit the generalisability to the broader secondary prevention population in Ireland. Patients were randomly assigned in a 1:1 fashion to receive alirocumab 75mg or matching placebo subcutaneously every two weeks (n=18,924). The dose of alirocumab was up titrated to 150mg or switched to placebo under blinded conditions to a target LDL-C of 0.6 to 1.3 mmol/L. The median length of follow up was 2.8 years. Alirocumab significantly reduced the risk of the primary endpoint (defined as death from CHD, nonfatal MI, fatal or nonfatal ischemic stroke or unstable angina requiring hospitalisation). (HR 0.85; 95% CI 0.78 to 0.93, p<0.001).

COMBO II was designed to demonstrate the LDL-C lowering ability of alirocumab, as add on therapy to maximum tolerated statin therapy in patients with high CV risk in comparison to ezetimibe (160). Over 88% of patients had a history of CHD. Other patients reported a history of stroke and PAD. Patients classified under primary prevention who met pre-defined eligibility criteria classifying them as very high risk were also eligible. It is not possible to deduce the exact proportion of primary versus secondary prevention from the trial characteristics reported but it is less than 12%. Patients were randomised in a 2:1 ratio to alirocumab 75mg every two weeks

[2] In the original New England Journal of Medicine publication of the FOURIER trial, 99% confidence intervals from the CTTC meta-analysis were mislabelled as 95% confidence labels in the supplementary appendix. This error led to an overestimation of the uncertainty surrounding the CTTC analysis and the extent of overlap with FOURIER (321). Following a letter I wrote to the journal editors, the appendix was corrected. (Letter was not published as the error was not in the main body of the manuscript.)

subcutaneously or ezetimibe 10mg once daily orally (n=721). Alirocumab doses were titrated according to a pre-defined titration regimen. The duration of the trial was 104 weeks. The primary endpoint was percent change in calculated LDL-C from baseline to week 24. The means ± SE difference between groups was 29.8% (95% CI 34.4 to 25.3; p<0.0001).

3.4.2.2 Ongoing studies

VESALIUS CV is randomised double blind trial, comparing evolocumab versus placebo, on the risk of major CV events in adults who are at high risk of a CV event but have not experienced a prior MI or stroke (158). Patients are required to have significant CAD, significant cerebrovascular disease or diabetes. Therefore, the trial will include some primary prevention patients who have diabetes. Patients are also required to have at least 1 high risk feature.[3] The primary endpoint is time to CHD death, MI or ischemic stroke from randomisation. The estimated median duration of follow-up is 4.5 years. The estimated primary completion date is May 2024.

NEWTON-CABG is a randomised double blind trial, comparing evolocumab versus placebo for 24 months, on the risk of CV events in a broad population of patients undergoing coronary artery bypass graft (CABG) surgery (159). The primary outcome is the saphenous vein graft disease rate[4]. Other trial outcomes include pre-defined outcomes. The estimated enrolment is 766, with an estimated primary completion date of December 2023.

[3] The definition of high-risk feature is not provided in the clinicaltrials.gov summary which is the only record identified.

[4] Defined as the proportion of vein grafts with significant stenosis or total occlusion (≥50%) on 64-slice (or greater) cardiac CT angiography (CTA) or clinically indicated coronary angiography.

Table 3.2 Baseline characteristics of trial populations

Trial	Interventions	N	Key Inclusion Criteria	Key Exclusion Criteria	Age Yrs	Male %	LDL-C mg/dL	CHD %	Stroke %	PAD %
FOURIER (7)	Evolocumab	13,784	Age ≥ 40 to ≤ 85 years History of clinically evident CVD as evidenced by ANY of the following: diagnosis of myocardial infarction diagnosis of non-haemorrhagic stroke (TIA does not qualify as stroke for inclusion) symptomatic PAD as defined in protocol. ≥1 major or at ≥2 minor risk factors as described in protocol.	Subject must not be randomized within 4 weeks of their most recent MI or stroke NYHA class III or IV, or last known left ventricular ejection fraction < 30% Known haemorrhagic stroke at any time Planned or expected cardiac surgery or revascularization within 3 months after randomization Uncontrolled hypertension Severe renal dysfunction	62.5	75.4	Median 92	80.9	19.5	13.5
	Placebo	13,780			62.5	75.5	92	81.3	19.2	12.9
ODYSSEY OUTCO-MES (6)	Alirocumab	9,462	Hospitalisation for ACS, defined by symptoms of myocardial ischemia with an unstable pattern, occurring at rest or with minimal exertion, within 72 h of an unscheduled hospital admission due to presumed or proven obstructive coronary disease and at least one of the following: Elevated cardiac biomarkers Resting ECG changes consistent with ischemia or infarction, plus additional evidence of obstructive coronary disease Lipid levels inadequately controlled by atorvastatin 40-80 mg or rosuvastatin 20-40 mg daily or maximum tolerated	Age below 40 years Qualifying index ACS event more than 52 weeks before randomization Not on stable lipid-modifying therapy for ≥2 weeks before randomisation Uncontrolled hypertension visit) New York Heart Association class III or IV CHF History of haemorrhagic stroke Fasting triglycerides >4.52 mmol/L Recurrent ACS event within 2 weeks prior to randomization Coronary revascularization procedure performed within 2 weeks prior to randomization visit or planned after randomization	58.5	74.7	*mean 92	100	3.2	3.9
	Placebo	9,462			58.6	74.9	92	100	3.2	4.1

Trial	Interventions	N	Key Inclusion Criteria	Key Exclusion Criteria	Age Yrs	Male %	LDL-C mg/dL	CHD %	Stroke %	PAD %
			dose of one of these drugs, defined by at least one of the following: - LDL-C ≥ 70 mg/dL - Non–HDL-C ≥ 100 mg/dL - Apolipoprotein B ≥ 80 mg/dL	Creatine kinase x 3 times upper limit of normal; estimated glomerular filtration rate b 30 mL/ (min 1.73 m^2); Use of fibrates other than fenofibrate or fenofibric acid						
COMBO II (151,160)	Alirocumab	479	Patients with hypercholesterolemia and established CHD or CHD risk equivalents with LDL-C poorly controlled with a maximally tolerated statin at stable dose Baseline LDL-C levels required depend on history of documented CVD: LDL-C ≥1.81 mmol/ with a history of documented CVD or LDL-C ≥2.59 mmol/L in patients without history of documented CVD	Age <18 years Fasting serum triglycerides >400 mg/dL (>4.52 mmol/L) during the screening period Currently taking a statin that is not simvastatin, atorvastatin, or rosuvastatin taken daily at a registered dose	61.7	75.2	Median 2.8	91.2	8.4	5.0
	Ezetimibe	241			61.3	70.5	Median 2.7	88.0	8.3	4.6

ACS, Acute Coronary Syndrome; CHD, Coronary Heart Disease; CHF, Congestive Heart Failure; CVD, Cardiovascular Disease; dL, Decilitre; ECG, Electrocardiogram; HDL-C, High Density Lipoprotein Cholesterol; LDL-C, Low Density Lipo-Protein Cholesterol; m, metre; mg, milligram; MI, myocardial infarction; min, minute; mmol, millimole; NYHA, New York Heart Association; PAD, Peripheral Arterial Disease; TIA, Transient Ischaemic Attack.

3.4.3 Data availability

As shown in Table 3.3, data was not available for every outcome across all trials. Pre-defined fatal event thresholds were used when deciding whether to substitute reports of any event into the event outcomes as described in Section 3.3.1.4.1. It was decided to report data on the outcome fatal or non-fatal haemorrhagic stroke as a post-hoc review outcome given the absence of data on non-fatal haemorrhagic stroke. The effect of PCSK9 inhibitor therapy on HRQoL was not reported for any study (despite being measured in two). COMBO II provided evidence on serious adverse events, but no evidence was reported any of the other outcomes.

Table 3.3 Data availability for outcomes of interest across included studies

	FOURIER	ODYSSEY OUTCOMES	ODYSSEY COMBO II	Comments
Time to non-fatal MI				Time to non-fatal MI is not reported in FOURIER, but as <10% of the total number of MI were fatal, fatal or non-fatal MI will be substituted for this endpoint in FOURIER as pre-defined in the review protocol.
Time to fatal or non-fatal MI				
Time to non-fatal ischemic stroke				Time to non-fatal ischemic stroke is not reported in any of the included studies. Time to fatal or non-fatal ischemic stroke is reported in two studies. As the total number of deaths due to ischemic stroke is <10% of the total number of ischemic strokes in both trials, this endpoint may be substituted for time to non-fatal ischemic stroke
Time to fatal or non-fatal ischemic stroke				
Time to non-fatal haemorrhagic stroke				Time to non-fatal haemorrhagic stroke was not measured in any of the trials. The total number of deaths due to haemorrhagic stroke is > 10% of the total number of haemorrhagic strokes in both trials. Therefore, results did not meet the pre-defined outcome
Time to fatal or non-fatal haemorrhagic stroke				
Time to CV death				Results are available for 2/3 studies
Time to CHD Death				Results are available for 1/3 studies
HRQoL				The EQ-5D was a pre-specified outcome measure in ODYSSEY OUTCOMES. The EQ-5D was measured in ODYSSEY COMBO II (No protocol available) but no results by trial arm is reported.
Serious Adverse Events				Results are available for all studies.
KEY				Outcome is reported in this Trial / Outcome is measured but results could not be located. / Results for this outcome could not be located for this trial.

CHD, Coronary Heart Disease; CVD, Cardiovascular disease; HRQoL, Health related quality of life; MI, Myocardial Infarction.

3.4.4 Risk of bias – primary efficacy outcomes

As recommended by the Cochrane Handbook, the risk of bias in each study was assessed separately for each of the primary outcome measures and presented graphically with the corresponding meta-analysis (Figure 3.2) (58). Summary results for each domain of ROB 2 stratified by primary outcome are presented in Supplementary Appendix 1.1.3. As ODYSSEY COMBO II did not contribute any data to the primary outcomes, the its risk of bias was not assessed (151,160).

For the outcome time to non-fatal MI, the overall risk of bias was low across trials. For the remaining outcomes, one or both trials were at risk of selection bias meaning the overall risk of bias for that trial outcome was unclear. FOURIER was deemed to have a high risk of selection bias for non-fatal ischemic stroke and non-fatal or fatal haemorrhagic stroke as these outcomes were not pre-specified (7). This was also the case for time to haemorrhagic stroke in ODYSSEY OUTCOMES (6). CV death was pre-specified in ODYSSEY OUTCOMES (6). However, it was not included in the original trial protocol but was added in the final version of the statistical analysis plan in 2017. As two unpublished pre-specified interim analyses for futility and overwhelming efficacy was conducted by Sanofi/Regeneron in 2016, there is a risk that this endpoint was selectively added at this point after favourable analysis were conducted. Therefore, it was deemed to have an unclear risk of selection bias. Haemorrhagic stroke was not pre-specified in any version of the protocol and trials were considered at unclear risk of selection bias for this reason.

3.4.5 Other potential source of bias

As there was only a maximum of three studies published for each outcome, we were unable to formally test for publication bias. However, we note that despite the EQ-5D being measured in ODYSSEY COMBO II and ODYSSEY OUTCOMES, it was not reported (6,151,160).

3.4.6 Meta-analysis

The results from our meta-analysis of the effects of PCSK9 inhibitors versus placebo on efficacy outcomes are presented below. The effect of PCSK9 inhibitors versus any control on serious adverse events is presented in the Supplementary Appendix Figure A1.

3.4.6.1 Primary outcomes

Forest plots, showing the results of the meta-analysis for each of the primary outcomes, are presented with the corresponding ROB 2 assessments in Figure 3.2. A table outlining the Summary of Findings include GRADE assessment is presented in Table 3.4.

There was moderate quality evidence that PCSK9 inhibitors reduce the time to non-fatal MI (HR 0.79; 95% CI 0.68 to 0.93) and the time to non-fatal ischemic stroke (HR 0.74; 95% CI 0.64 to 0.87) versus placebo. No data was available on the effect of PCSK9 inhibitors on the time to non-fatal haemorrhagic stroke. A post hoc analysis estimated that PCSK9 inhibitors did not affect the time to fatal or non-fatal haemorrhagic stroke, but the evidence was of low quality (HR 1.02; 95% CI 0.67 to 1.56).

There was low quality evidence PCSK9 inhibitors do not have an effect on the time to CV death over the time observed in the clinical trials (HR 0.96, 95% 0.81 to 1.14). But this is very uncertain, meaning that our confidence in this estimate is limited, and the true effect may be substantially different from the true estimate of effect.

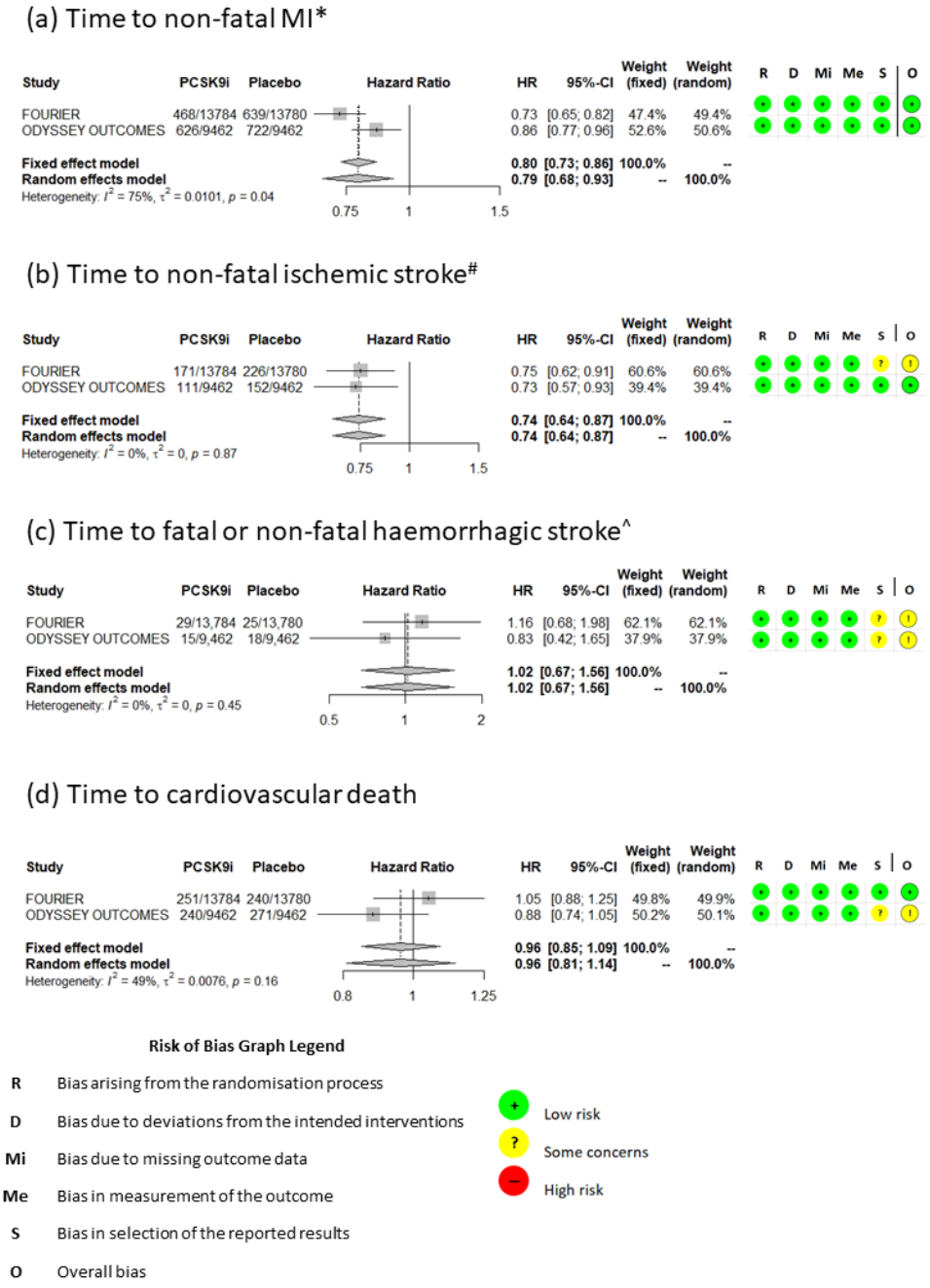


Figure 3.2 Efficacy of PCSK9 inhibitors versus placebo for reducing the risk of the primary outcomes

CI, Confidence Interval; HR, Hazard Ratio; PCSK9i, PCSK9 Inhibitor; RR, Risk Ratio.
* Time to non-fatal MI was not reported in ODYSSEY OUTCOMES. However, ODYSSEY OUTCOMES met pre-specified criteria in the protocol which stated that the number of fatal MI was <10% of the total number of MI, the endpoint fatal or non-fatal MI could be used instead.
# The pre-specified review outcome was time to non-fatal ischemic stroke. But it was not reported in FOURIER or ODYSSEY Outcomes. However, both trials met pre-specified review protocol which stated that if the number of fatal ischemic stroke was <10% of the total number of ischemic strokes, the endpoint fatal or non-fatal ischemic stroke could be substituted.
^Time to non-fatal haemorrhagic stroke is not reported in FOURIER or ODYSSEY OUTCOMES. Neither trial met the pre-specified criteria in the review protocol which stated that the number of fatal haemorrhagic stroke was <10% of the total number of haemorrhagic strokes the endpoint fatal or non-fatal haemorrhagic strokes could be substituted.

Table 3.4 Summary of findings PCSK9 inhibitors compared to placebo for the secondary prevention of cardiovascular disease across primary outcomes

Summary of Findings PCSK9 inhibitors compared to placebo for the secondary prevention of cardiovascular disease

Patient or population: Secondary Prevention

Setting: Ireland

Intervention: PCSK9 inhibitors

Comparison: Placebo

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with placebo	**Risk with PCSK9 inhibitors**				
Time to non-fatal myocardial infarction	**Low**		**HR 0.79** (0.68 to 0.93) [Time to non-fatal myocardial infarction]	46488 (2 RCTs)	⨁⨁⨁◯ MODERATE [a]	
	59 per 1,000	**47 per 1,000** (40 to 55)				
Time to non-fatal ischemic stroke	16 per 1,000	**12 per 1,000** (10 to 14)	**HR 0.74** (0.64 to 0.87)	46488 (2 RCTs)	⨁⨁⨁◯ MODERATE [a]	
Time to non-fatal haemorrhagic stroke - not reported						
Time to fatal or non-fatal haemorrhagic stroke	2 per 1,000	**2 per 1,000** (1 to 3)	**RR 0.86** (0.43 to 1.74)	46458 (2 RCTs)	⨁⨁◯◯ LOW [a,b]	
Time to cardiovascular death	22 per 1,000	**21 per 1,000** (18 to 25)	**HR 0.96** (0.81 to 1.14)	46488 (2 RCTs)	⨁⨁◯◯ LOW [a,b]	

***The risk in the intervention group** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI, Confidence interval; HR, Hazard Ratio; RCT, Randomised Control Trial; RR: Risk ratio.

Explanations

a. ODYSSEY only included patients who had a recent acute coronary syndrome, FOURIER only included patients who had a major vascular event and additional risk factors. There is no direct evidence for other patient groups in the secondary prevention population such as patients with a history of stable angina, transient ischemic attack or no history of cardiovascular disease. Therefore, we downgraded for indirectness.

b. The sample size is not sufficiently powered for this outcome. The 95% confidence intervals cross the line of no effect

3.4.6.3 Secondary outcomes

Forest plots showing the results of meta-analyses examining the effects of PCSK9 inhibitors versus placebo for the secondary outcomes are show in Figure 3.3.

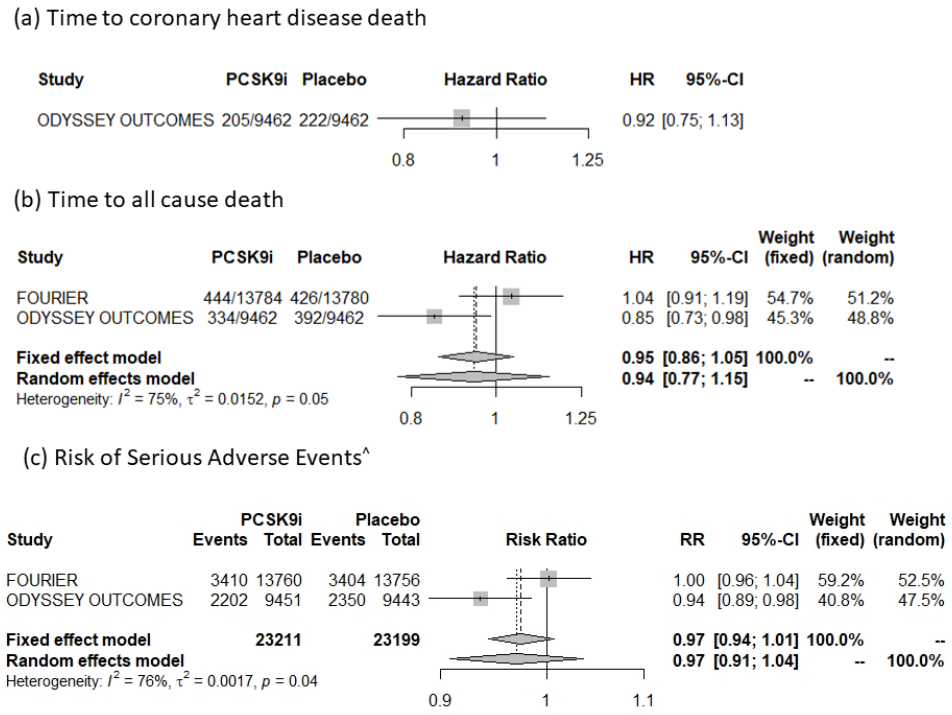


Figure 3.3 Meta-analysis of secondary outcomes for PCSK9 inhibitors versus placebo

Meta-analysis of secondary outcomes for the comparison of PCSK9 inhibitors versus control are presented in the Supplementary Appendix Figure A1. CI, Confidence Interval; HR, Hazard Ratio; PCSK9i, PCSK9 Inhibitor; RR, Risk Ratio.

Only one study reported on the time to CHD death. No statistically significant effect of PCSK9 inhibitors on the time to CHD death was observed (HR 0.92; 95% CI 0.75 to 1.13). Two eligible studies reported on the time to all-cause death. Data synthesis did not reveal any treatment effect of PCSK9 inhibitors on the time to all cause death (HR 0.94; 95% CI 0.77 to 1.15). No increase in the risk of serious adverse events was observed between PCSK9 inhibitors and placebo (RR 0.97; 0.81 to 1.04). Similar results for risk of serious adverse events were observed for the PCSK9 inhibitors versus control comparison (Supplementary Appendix Figure A.1).

No relevant information on the effect of PCSK9 inhibitors on HRQoL was observed. The EQ-5D is a validated generic preference-based HRQoL instrument which is commonly used in clinical trials. It was administered to patients in both ODYSSEY OUTCOMES and ODYSSEY COMBO II. However, available trial reports do not state which version of the instrument was used (3L or 5L), or pre-specify the value set to be used in the analysis. EQ-5D values by history of CVD were

presented in a pooled analysis of ODYSSEY COMBO II trial (161). However, disaggregated results were not presented by study or by treatment arm. The full protocol of ODYSSEY COMBO II was not located in the systematic search; therefore, the planned analysis is unknown. The protocol of ODYSSEY OUTCOMES includes EQ-5D by treatment arm and time as a pre-specified analysis. However, results could not be sourced. An abstract presented at the American College of Cardiology describes EQ-5D values from ODYSSEY OUTCOMES as an economic model input but no further detail is provided (136).

3.4.7 Heterogeneity

An important question before conducting meta-analysis is whether trials are sufficiently similar for comparisons to be valid.

Between study heterogeneity (measured as I^2) in treatment efficacy was observed for many outcomes. I^2 represents the percentage of statistical heterogeneity that is due to variability rather than due to chance. Across primary outcomes, no heterogeneity was observed for either stroke outcome, but significant heterogeneity was observed for time to non-fatal MI (I^2 = 75%; p = 0.04). Heterogeneity was also a feature for secondary outcomes where strong heterogeneity was observed for all-cause mortality (I^2 = 75%; p = 0.05) and serious adverse events (I^2 = 76%; p = 0.04). The low specificity of this statistic and the small number of studies mean that it is difficult to draw firm conclusions, but they support the pre-specified decision to conduct a random-effects meta-analysis.

Assessing heterogeneity between trials for an adjusted indirect comparison or network meta-analysis is more critical. If trials are not sufficiently similar, differences in treatment effects driven by study imbalances in clinical characteristics may be attributed to differences in the efficacy between treatments. An understanding of how treatment varies across subgroups is also important from a clinical and economic perspective.

In this network, deciphering the main drivers of study heterogeneity is difficult given the low number of studies and the numerous potential causes described below.

1. Baseline LDL-C was similar between ODYSSEY OUTCOMES and FOURIER but dose titration regimens differed (6,7). In ODSYSSEY OUTCOMES, the initial dose was 75mg every two weeks. Patients in the alirocumab arm were up titrated to 150mg every two weeks or switched to placebo to achieve a target LDL-C of 25-50mg/dl (0.65-1.29mmol/L) under blinded conditions (6). In contrast, no target range was defined in FOURIER and patients in the evolocumab arm received either 140mg every two weeks

or 420mg once monthly (7). Therefore, heterogeneity could be caused by differences in treatment effect at very low-levels of LDL-C. A priori, this would be expected to lead to a reduced treatment effect for alirocumab compared to evolocumab *ceteris paribus*.

2. The median duration of ODYSSEY OUTCOMES (2.8 years) is longer than FOURIER (2.2 years) (6,7). The relative treatment effect may vary with time as suggested by Sabatine et al (7). This assumption would support an increased treatment effect in ODYSSEY OUTCOMES compared to FOURIER *ceteris paribus*. But would also cast doubt on our choice of treatment effect measure which assumes constant proportional hazards.
3. Both trials randomised patients from a secondary prevention population but the populations in the two trials were very different. While patients in ODYSSEY OUTCOMES had a history of multiple event types, all randomised patients were required to be 1 to 12 months post ACS (6). FOURIER randomised patients post MI, ischemic stroke, or PAD with additional risk factors (7). No difference in relative effectiveness is expected across groups based on observations in statin trials (41,162,163).

While it not possible to assess the first two potential causes of heterogeneity because of the low number of studies, it is possible to examine the potential effect of CV history by examining the consistency of treatment efficacy in subgroups across the clinical trials. If consistent, it will provide reassurance that an adjusted indirect comparison of evolocumab and alirocumab using FOURIER and ODYSSEY OUTCOMES is valid. While baseline LDL-C was balanced between trials, it was pre-defined as an important subgroup given the expected increased treatment benefit from the greater absolute LDL-C reduction in this group.

Therefore, subgroup results for the secondary prevention population are presented by baseline LDL-C and by history of CVD. Given the paucity of data, the subgroup analysis was restricted to the comparison of PCSK9 inhibitor versus placebo for the primary outcomes only. Trials were not powered for subgroup analysis and statistical tests are not adjusted for multiplicity. These issues should be accounted for when interpreting the results.

3.4.7.1 Primary outcomes

3.4.7.1.1 Baseline LDL-C

Baseline LDL-C is associated with increased CV risk and absolute reductions in LDL-C are higher in this group. Therefore, it is hypothesised that the both the absolute and relative treatment effect of PCSK9 inhibitors will be greater in this group. Because of the higher expected benefit, many countries including Ireland have included minimum baseline LDL-C thresholds as part of their reimbursement criteria for PCSK9 inhibitors.

Subgroup analysis by baseline LDL-C was not reported for any of the primary outcomes in FOURIER (7). Subgroup analysis by baseline LDL-C was only reported for the time to CV death outcome in ODYSSEY OUTCOMES (136). Results are presented in Table 3.5. In the subgroup with a baseline LDL-C ≥100 mg/dL, alirocumab reduced the time CV death (HR 0.69; 95% CI 0.52 to 0.92). It is concerning those outcomes for the subgroup <100 mg/dL were not presented. Results of statistical tests to examine an interaction between baseline LDL-C and HRs were not reported. Subgroup analysis for the other primary outcomes specified in this review were not reported across either trial.

Table 3.5 Effect of PCSK9 inhibitors versus placebo on time to CV death by baseline LDL-C (136)

Baseline LDL-C	**Alirocumab** Number with event/Number of patients in arm	**Placebo** Number with event/Number of patients in arm	**HR 95% CI**
ODYSSEY OUTCOMES			
<100mg/dL			Not reported
≥100mg/dL	81/2814	117/2815	0.69 (0.52 – 0.92)

CI, Confidence Interval; dL, decilitre; HR, Hazard Ratio; LDL-C, low density lipo-protein cholesterol; mg, milligram.

3.4.7.1.2 History of CVD

Subgroup results by history of CVD were not publicly available for any of the remaining primary outcomes in this review for any eligible studies.

3.4.7.2 Composite outcomes

Given the paucity of subgroup data for the primary outcomes, studies which reported subgroup analysis of composite outcomes containing any of the reviews primary outcomes were examined. Results are presented in Table 3.6. Results were not presented in a form that would facilitate meta-analysis.

3.4.7.2.1 Baseline LDL-C

No statistically significant interaction in the treatment effects of PCSK9 inhibitors versus placebo was observed between baseline LDL-C and any of outcomes reported. Hazard ratios decreased consistently with increasing LDL-C for stroke, but this relationship was not observed for the ODYSSEY OUTCOMES primary endpoint. For the secondary efficacy endpoint of FOURIER (time to non-fatal MI, non-fatal ischemic stroke and CVD death), hazard ratios by baseline LDL-C quartile point estimates moved in the opposite direction to that which would be expected if HRs decreased with increasing LDL-C. Results for the primary endpoint in FOURIER were comparable. However, for both trials, the range of baseline LDL-C was relatively narrow across quartiles limiting the power to detect differences in treatment effect across baseline LDL-C. There is

insufficient evidence to determine if increasing baseline LDL-C is associated with decreasing hazard ratios.

3.4.7.2.2 History of CVD

Subgroup analysis by history of CVD for available outcomes is presented in Table 3.6. The population of FOURIER had a greater patient diversity as measured by CV history compared to ODYSSEY OUTCOMES (6,7). A statistically significant interaction in the treatment effects of PCSK9 inhibitors versus placebo was only observed for one outcome subgroup combination – FOURIER Primary Endpoint stratified by Time since MI (EMA version) (107). No statistically significant interaction was observed for the FOURIER Secondary Efficacy Endpoint stratified by the same subgroup. No interaction was observed in the equivalent ODYSSEY OUTCOMES stratification albeit, everyone had experienced an ACS event with the previous year (135). Again, it is important to note that p values are not adjusted for multiplicity.

There is little evidence to suggest that a history of CVD has a differential treatment effect on the hazard ratio, but trials were not powered or designed to detect differences in this outcome. Therefore, a differential effect cannot be excluded.

Table 3.6 Review outcomes by baseline LDL-C and history of CVD

Subgroup	Alirocumab				Evolocumab		
	N	Time to CVD death (136)	Stroke (125)	Odyssey Primary Endpoint(6,135)	N	FOURIER PEP (7,107,110)	FOURIER SEP (7,107,110)
Baseline LDL-C mg/dL							
<80	7,164	NR	0.90 (0.61, 1.34)	0.86 (0.74, 1.01)	6,961	0.80 (0.69, 0.93)	0.78 (0.64, 0.95)
80>100	6,128	NR	0.72 (0.46, 1.12)	0.96 (0.82, 1.13)			
80-92					6,886	0.82 (0.71, 0.96)	0.79 (0.65, 0.96)
92-109					6,887	0.89 (0.77, 1.03)	0.79 (0.66, 0.94)
>=100	5,629	0.69 (0.52, 0.92)	0.59 (0.40, 0.86)	0.76 (0.65, 0.87)			
>109					6,829	0.89 (0.77, 1.02)	0.83 (0.70, 0.99)
Interaction p		NR	0.31	0.09		0.69	0.96
History of CVD							
History of CVD FOURIER							
MI alone					19,113	0.88 (0.80, 0.96)	0.80 (0.71, 0.90)
stroke alone					3,366	0.70 (0.54, 0.90)	0.88 (0.58, 1.02)
PAD alone					1,505	0.67 (0.47, 0.96	0.57 (0.38, 0.88)
Polyvascular					3,563	0.88 (0.75, 1.03)	0.86 (0.71, 1.04)
Inter p value						0.19	0.38
Time since MI (EMA)							
No MI					5,213	0.68 (0.56, 0.83)	0.69 (0.55, 0.87)
MI < 1year					5,711	0.81 (0.70, 0.93)	0.75 (0.62, 0.91)
MI >1yr - <2 yrs					2,691	0.81 (0.65, 1.01)	0.79 (0.59, 1.04)
> 2 years					13,918	0.95 (0.85, 1.05)	0.87 (0.76, 0.99)
Interaction p						0.02	NS
Time since MI							
MI < 1 year					5,711	0.81 (0.70, 0.93)	0.75 (0.62, 0.91)

Subgroup	Alirocumab N	Alirocumab Time to CVD death (136)	Alirocumab Stroke (125)	Alirocumab Odyssey Primary Endpoint(6,135)	Evolocumab N	Evolocumab FOURIER PEP (7,107,110)	Evolocumab FOURIER SEP (7,107,110)
MI > 1 year					16,609	0.92 (0.84, 1.01)	0.85 (0.76, 0.96)
Interaction p v						0.13	0.24
Stroke							
Yes			0.90 (0.52-1.56)		5,337	0.85 (0.72, 1.00)	0.89 (0.74, 1.08)
No			0.68 (0.53 – 0.88)		22,227	0.85 (0.79, 0.93)	0.77 (0.70, 0.86)
Interaction p			0.37			NS	NS
PAD							
Yes	3,462					0.79 (0.66, 0.94)	0.73 (0.59, 0.91)
No	23,922					0.86 (0.80, 0.93)	0.81 (0.73, 0.90)
Interaction p						NS	NS
Time between index ACS and Randomisa-tion (weeks)							
<8	5,240			0.85 (0.72, 0.99)			
>=8 to <24	10,113			0.83 (0.74, 0.94)			
>=24	3,571			0.91 (0.73, 1.15)			
Interaction p				0.77			

ACS, acute coronary syndrome; CVD, cardiovascular disease; dL, decilitre; EMA, European Medicines Agency; mg, milligram; MI, Myocardial Infarction. PAD, peripheral arterial disease; PEP, primary endpoint; SEP, secondary endpoint; LDL-C, low density lipo-protein cholesterol; N, number; NS, reported as not significant.

3.4.7.3 Heterogeneity in LDL-C lowering ability of PCSK9 inhibitors

In the absence of direct evidence of the effect of baseline LDL-C on the treatment effect of PCSK9 inhibitors, previous economic evaluations have estimated differences in treatment effects across subgroups through the surrogate endpoint of LDL-C. At higher baseline LDL-C values, greater absolute reductions in LDL-C are expected to lead to more favourable hazard ratios.

Previous economic evaluations have applied mean percentage reductions in baseline LDL-C from clinical trials. However, they have failed to examine if baseline LDL-C is a treatment effect modifier of alirocumab's ability to lower LDL-C in percentage terms (164). The ODYSSEY LONG TERM trials examined the efficacy and safety of alirocumab for LDL-C reduction in patients with a history of CVD (38). All patients in the alirocumab arm received a dose of 150mg every two weeks subcutaneously. A forest plot in the LONG-TERM trial shows a statistically significant difference in percentage LDL-C reductions compared to placebo across subgroups. The least squares mean difference versus placebo for patients with a baseline LDL-C <100mg/dL (2.59mmol/L) is -75% compared to -41.3% in patients with a baseline LDL-C >160mg/dL (4.14mmol/L). A similar treatment effect size was observed in the ODYSSEY HIGH FH trial where everyone had an LDL-C of >4.14 mmol/L (165). This confirmation reduces the likelihood that the ODYSSEY LONG TERM results represent spurious findings.

In the other alirocumab trials it is not possible to elucidate the impact of baseline LDL-C on percentage LDL-C reduction, because of the correlation of baseline LDL-C with the proportion of patients requiring up titration from the 75mg to the 150mg strength. Patients with higher LDL-C are more likely to be up-titrated as they require a greater percentage reduction in LDL to reach treatment targets. Because of the confounding, previous meta-regressions of percentage LDL-C reduction on baseline LDL-c have failed to detect an effect (44).

To our knowledge, no evolocumab trial has published results by subgroup to confirm whether this treatment interaction is also present for evolocumab. However, in a trial of evolocumab in patients in patients with homozygous FH, the percentage LDL-C reduction was only 30% despite a different treatment dose being used in this population (44). Given there is little clinical difference between alirocumab and evolocumab, it is reasonable to conclude that the evolocumab's ability to lower LDL-C in percentage terms also depends on baseline LDL-C.

3.4.8 Adjusted indirect comparison of alirocumab versus evolocumab

The evidence network of clinical trials is presented in Figure 3.4. Evolocumab and alirocumab are connected in the network by their trial comparisons with placebo. Ezetimibe is connected to the network via alirocumab in the ODYSSEY COMBO II trial. As there is no connection between

ezetimibe and evolocumab this trial is not included in the indirect comparison of alirocumab and evolocumab.

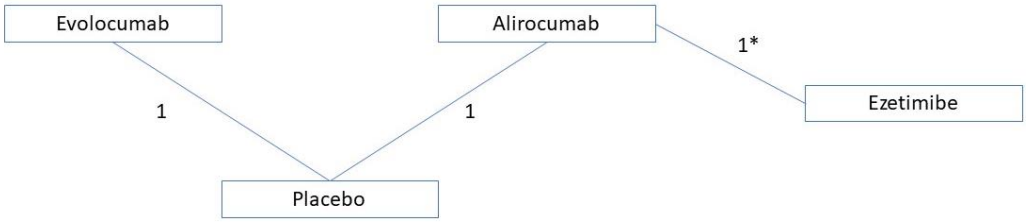


Figure 3.4 Network diagram of RCT evidence

Each node in the network represents a treatment strategy. Complete lines between nodes represent pairwise treatment comparison from randomised controlled trials which are labelled with the number of trials for that link. *Only outcome available for this comparison is the Serious adverse events outcome. RCT, Randomised Controlled Trial.

Before conducting an adjusted indirect comparison, it is important to assess whether the similarity assumption holds. Table 3.2 shows that the population characteristics are quite different between trials. All of patients in ODYSSEY OUTCOMES had a recent ACS event but many patients also had a history of CVD in other vascular beds (6). FOURIER also included patients with a history of MI, but no restriction was placed on the time since clinical event (7). Patients were also eligible for FOURIER if they had a history of stroke and clinically evident PAD. As described in Table 3.2, additional risk factors were also required.

While the populations in the included trials are quite different, an adjusted indirect comparison between alirocumab and evolocumab will still be appropriate if treatment effect modifiers are balanced between the trials. Based on the analysis in Section 3.4.7, there is little evidence to suggest history of CVD has a treatment modifying effect on the hazard ratio. Furthermore, previous meta-analysis analysing statins have not observed a difference in treatment effect across history of CVD after accounting for the extent of absolute LDL-C reduction (162). We concluded that there was insufficient evidence to reject the similarity assumption. Therefore, it was decided to proceed with the adjusted indirect comparison whilst acknowledging its limitations of the evidence available.

Adjusted estimates of the effect of alirocumab versus evolocumab on primary outcomes with 95% confidence intervals are presented in the summary of findings Table 3.7. No statistically significant treatment effect was observed for alirocumab versus evolocumab for the primary outcomes time to non-fatal MI, time to non-fatal ischemic stroke, and time to CVD death. We graded the quality of the evidence as very low. No results were available for the time to non-

fatal haemorrhagic stroke outcome. A post hoc analysis of the time to fatal or non-fatal haemorrhagic stroke showed no statistically significant treatment effect. Again, the quality of evidence graded was very low.

Results for secondary outcomes are presented in Table 3.8. There was no statistically significant difference in the time to all cause death or in the risk of serious adverse events between alirocumab and evolocumab. No data was available for the secondary outcomes time to CHD death or difference in HRQoL.

Table 3.7 Summary of findings table alirocumab compared to evolocumab for prevention of cardiovascular disease

Alirocumab compared to evolocumab for prevention of cardiovascular disease

Patient or population: Secondary Prevention
Setting: Ireland
Intervention: Alirocumab
Comparison: Evolocumab

Outcomes	Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
Time to non-fatal MI	**HR 1.18** (1.00 to 1.38)	(2 RCTs)	⨁◯◯◯ VERY LOW [a,b]	
Time to non-fatal ischemic stroke	**HR 0.96** (0.70 to 1.32)	(2 RCTs)	⨁◯◯◯ VERY LOW [a,b]	
Time to non-fatal haemorrhagic stroke - not reported	-	-	-	
Time to fatal or nonfatal haemorrhagic stroke	**HR 0.72** (0.30 to 1.70)	(2 RCTs)	⨁◯◯◯ VERY LOW [a,b,c]	
Time to CVD death 0	**HR 0.84** (0.65 to 1.07)	(2 RCTs)	⨁◯◯◯ VERY LOW [a,b]	

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI, Confidence interval; CV, Cardiovascular HR, Hazard Ratio; MI, Myocardial Infarction.
a. Downgraded two levels for indirectness. Evidence is from an indirect treatment comparison. Further, the patient populations in both trials were different. FOURIER included patients who had a history of a major vascular event and additional risk factors. ODYSSEY OUTCOMES only included patients with a history of recent acute coronary syndrome (ACS). Exploratory subgroup evidence from FOURIER suggests that the treatment effect of evolocumab depends on the length of time since myocardial infarction in those who have had an MI. However, another analysis suggests that the treatment effect in FOURIER did not differ between the recent ACS population versus the non-recent ACS population.
b. Downgraded one level for imprecision as the 95% confidence interval crosses the line of no effect.
c. Downgraded a further level for imprecision because of the very low number of events.

Table 3.8 Adjusted indirect comparison alirocumab v evolocumab for the primary review outcomes

	Alirocumab v Evolocumab	
Secondary Outcomes	HR	95% CI
Time to CHD Death	Not Reported	
Time to All Cause Death	0.82	(0.67-1)
	RR	95% CI
Risk of Serious Adverse Events	0.93	(0.88-1)
Difference in HRQoL	Not Reported	

CI, Confidence Interval; CHD, Coronary Heart Disease; HRQoL, Health Related Quality of Life; HR, Hazard Ratio; RR, Risk Ratio.

3.5 Discussion

3.5.1 Main findings

Our study shows moderate quality evidence that PCSK9 inhibitors reduce the time to non-fatal MI and non-fatal ischemic stroke compared to placebo in the secondary prevention population. The available evidence indicates that PCSK9 inhibitors do not reduce the time to CVD death over the time period observed in clinical trials but the quality of evidence for this outcome is low. This means that it is possible that the true effect could be substantially different. The results are broadly similar to those reported in other meta-analyses (45,50). However, to our knowledge this is the first meta-analysis of PCSK9 inhibitors to measure treatment effectiveness using the hazard ratio which is required to use these results in our economic model.

It is reasoned in the literature that the duration of trials (that investigate PCSK9 inhibitors for their effect on cardiovascular outcomes) may have been too short to observe expected treatment effects on mortality (166). In a pre-specified analysis of follow-up of ODYSSEY OUTCOMES in patients with at least 3 years of follow up, alirocumab decreased all-cause mortality by 22%, (HR 0.78; 95% CI 0.65 to 0.94) (167). No decrease in all-cause mortality was observed in those patients with less than three years of follow-up but the interaction p value was not significant (HR 0.95; 95% CI 0.76-1.21) (167). However, evidence from FOURIER is not consistent with this hypothesis. In landmark analysis from FOURIER, the HR point estimate for CV death was higher beyond the first year (HR 1.12; 95% CI 0.88 to 1.42) compared to the first year alone (HR 0.96; 0.74 to 1.25). No statistical tests were reported to test significant differences between these groups. The evidence on the effect of PCSK9 inhibitors on the risk of death is inconclusive.

We included HRQoL as an outcome measure on the basis that literature reviews should review the evidence for the most important outcomes rather than just focusing on what is most commonly reported (58). It was also included as an outcome measure in the Cochrane review by Schmidt et al but no results were found at the time that their search was executed (44). We identified references that were published subsequently which outline that EQ-5D was included in both ODYSSEY COMBO II and ODYSSEY OUTCOMES. However, no references presenting results of this endpoint were located. It is concerning that important pre-specified analyses from ODYSSEY OUTCOMES are not released in the public domain. It is well documented in the literature that less favourable results are subject to publication bias (58).

Evidence for the relative treatment efficacy of PCSK9 inhibitors versus placebo by baseline LDL-C and history of CVD was assessed for the primary review outcomes. Alirocumab is reported to

reduce the rate of CVD death in the subgroup ≥100mg/dL by 31% compared to a reduction of 12% in the total population (6,136). Subgroup analysis should be interpreted in the context of supporting evidence available given the possibility of spurious results. The relative treatment effect is hypothesised to increase with increasing baseline LDL-C which adds biological plausibility to this estimate (162). However, it is it is concerning that the results of the subgroup <100mg/dL were not reported. Comparable results were not presented in FOURIER. But the hazard for the CVD death endpoint was >1 for the total trial population, and a relationship between baseline LDL-C and the hazard ratio was not observed for the composite secondary efficacy endpoint of FOURIER (which includes CV death) (7). Therefore, the efficacy of PCSK9 inhibitors, including alirocumab, in patients with high baseline LDL-C is very uncertain. Analysis of individual patient data with baseline LDL-C modelled as a continuous outcome measure would provide increased power to detect an interaction treatment effect but this is not reported across any trials.

This study incorporated most recent methodological developments in this area. As the meta-analyses was conducted in 'R' using the 'R' package 'meta' as opposed to traditional methods such as RevMan, more complex analysis methods recommended by the Cochrane handbook to quantify uncertainty were incorporated into the analysis (58). Of the systematic reviews of PCSK9 identified, this review is the only one to use ROB 2 which recommends assessment on an outcome rather than a study level. Extensive use was made of published trial protocols and it was deemed that the risk of bias was generally unclear because of concerns surrounding selection bias for some outcomes. Other reviews have assessed bias on a study level and classified the trials included here as having low risk of selection bias (45).

To our knowledge, this study represents the first full report of an indirect treatment comparison of alirocumab versus evolocumab. There is no evidence of a difference in treatment effect between both PCSK9 inhibitors. But as shown in Table 3.7, the quality of this evidence is very low. It has been previously highlighted that the internal validity of an indirect treatment comparison is contingent on the appropriate identification of studies that make up the network, the quality of the individual RCTs and the extent of confounding due to violations in similarity assumptions (76). Comparing this study versus these criteria, there was an exhaustive search strategy. However, an indirect treatment comparison from the literature published in abstract form included more studies due to broader eligibility criteria. They identified a significant difference in treatment effect of alirocumab versus evolocumab on the risk of all-cause death (54). A significant concern is whether the similarity assumption holds. There are always differences between trials in networks and there is no consensus on the point at which the

assumption should fail. It is believed that the judgements on similarity should be considered in the context of the conclusions that are being drawn from the analysis. For example, I consider it reasonable that the similarity assumption holds sufficiently to justify the conclusion that there is no evidence of a difference in treatment effect between alirocumab and evolocumab. But it is questionable whether it holds sufficiently to support the conclusions of Guedeney et al. who conclude that there is a difference in all-cause mortality between alirocumab and evolocumab (54). Meta-regression is a technique which can be used to adjust for differences across trials, but the number of eligible studies was insufficient to consider this technique here.

3.5.2 Strengths and limitations

The quality of the study was enriched by a comprehensive search strategy which identified supplementary appendices, conference abstracts and regulatory documents. These references provided important information on heterogeneity and potential risk of bias across studies.

The inclusion criteria were deliberately selective for the large trials. This approach could potentially omit important evidence from small trials which may be considerable when considered in aggregate. However, in a recent meta-analysis (which included most known trials), FOURIER accounted for 98% of the weighting in the evolocumab arm, and ODYSSEY OUTCOMES accounted for 96% of the weighting for the alirocumab trials when assessing the CVD death outcome. Weightings for other outcomes were similar. Further, we are not aware of other trials examining PCSK9 inhibitors that reported results for time to event outcomes. Therefore, we do not expect our conclusions to change substantially with the addition of further trials (168).

Pre-registration of review protocols in databases such as PROSPERO is recommended. But a protocol was prepared and differences in methods stated in the review and implemented in practice are described.

It is best practice that screening and assessing study eligibility are carried out in duplicate as opposed to a single reviewer. Screening and assessing eligibility are labour intensive processes. The single reviewer approach was considered pragmatic; it was deemed that the likelihood of omitting large-scale trials specified in the review was low. Bias was limited by conducting a pre-specified quality assurance check on 10% of full text articles in duplicate.

The limitations of this clinical evidence should be considered. Both ODYSSEY OUTCOMES and FOURIER were selective in their trial eligibility criteria. While it is expected that these results translate to the wider secondary prevention population, there is no direct evidence to support this. The quality of evidence in GRADE was reduced by one level to account for the indirectness.

The trial durations were relatively short for treatments with an expected lifelong treatment duration. The median follow-up time was only between 2.2 and 2.8 years. Therefore, the long-term treatment effect is unclear. Furthermore, the results are primarily based on only two studies, however the number of patients is large.

3.5.3 Potential for further research

It is recommended that this review is updated as the results of ongoing studies become available. However, this evidence alone is unlikely be sufficiently robust to reduce the uncertainty surrounding PCSK9 inhibitors efficacy on CV mortality. There is potential for observational data to provide to provide long term estimates of treatment effectiveness and to provide subgroup specific estimates. Propensity score analysis can be used to derive treatment effectiveness estimates derived from observational data sources adjusted for known confounders. But a limitation with this approach is that unknown confounders may bias the estimate. However, as described in Chapter 2, evolocumab and alirocumab were approved in the UK for a subgroup of the licensed population based on the baseline LDL-C. The LDL-C cut off may act as a form of natural randomisation measure which could be used to further adjust for unknown confounders.

A similar approach was adopted by Myers et al who exploited variance and difficulty in obtaining reimbursement approval for PCSK9 inhibitors in the US (169). It was hypothesised that CV events would be more prevalent in patients with rejected or abandoned PCSK9 inhibitor prescriptions than for those with paid prescriptions. HRs for composite CV events in a propensity matched analysis were 1.10 (95% CI; 1.01 to 1.19; p=0.02) for rejected versus paid and 1.12 (95% CI, 1.01 to 1.24; P=0.03) for abandoned versus paid [5]. The potential treatment effects observed here are modest, but the mean follow up duration across all groups was less than 14 months.

While we briefly reviewed the existing systematic review and network meta-analysis conducted, we did not assess the quality of the reviews formally. AMSTAR2 is a critical appraisal tool for systematic reviews that could be used to evaluate systematic reviews in a standardised fashion (170). Given the questionable nature of some of the decisions taken during analysis, a formal assessment of review quality in this area is an important research question.

[5] In this analysis, a HR > 1 for the “paid” versus the “rejected” or “abandoned” groups, suggests a treatment effect in favour of PCSK9 inhibitors.

3.6 Conclusion

No results from eligible studies were available for the primary prevention population. However, this review has shown that PCSK9 inhibitors are a clinically effective treatment option for patients with a history of CVD. However, it is still unclear how the relative treatment effects estimated here could translate to absolute reductions in CVD events in Ireland. This depends on the baseline risk of CV events and competing risks. Therefore, identification of the patient groups who could benefit most from PCSK9 inhibitors and the respective cost effectiveness is unclear. In Chapter 7, the treatment effectiveness estimates derived here will be applied in an economic model to examine these issues.

Chapter 4. The interaction between price negotiations and heterogeneity

4.1 Introduction

Economic evaluation is an important element of the decision-making process when considering the reimbursement of drugs in Ireland and across the world. Negotiations between decision-makers and pharmaceutical companies also play a pivotal role (171,172). This can take many forms including price reductions and budget caps. Increasingly, reimbursement is limited to subgroups of the licensed population to optimise their value or restricted to a subset of indications where a drug is licensed for multiple indications (33–36,173). One of the roles of the decision-maker is to maximise population health within the constraint of the budget available.

Pharmaceutical companies wish to maximise treatment effectiveness as part of their aim to maximise profits. The price sought by the pharmaceutical company is constrained by the price the decision-maker is willing to pay, which may be informed by the cost-effectiveness of the drug. In a homogenous population, the company will aim to price the drug at the point the drug crosses the cost effectiveness threshold to maximise return.

However, in a heterogeneous population, the situation is more complex. As described in Chapter 1, heterogeneity can lead to variation in the cost-effectiveness of a drug across subgroups (25,174). At any given price, the decision-maker may only be willing to grant reimbursement to the subset of the population in whom the drug is cost effective. Therefore, companies would be required to reduce the price of the drug to gain reimbursement in the entire population. However, if some of these subgroups are very small, companies may be unwilling to reduce prices to the levels required to make the drug cost effective. Similar issues arise when assessing drugs which are licensed for multiple indications (175).

However, to our knowledge, the relationship between price negotiations and heterogeneity has not been considered in the literature to date. We argue that, when both are present, an interaction may be generated. Failure to account for the interaction, may lead to incorrect conclusions being drawn regarding the cost-effectiveness of drugs.

4.1.1 Aims

This chapter has three aims:

1. Highlight the interaction that strategic behaviour, such as price negotiations, can introduce to economic evaluations.
2. Present a framework, to guide the economic evaluation process in these scenarios, given the additional complexity that the interaction introduces.

3. Explore the implications that the interaction has on the results of economic evaluations and drug prices including the consequences of ignoring the interaction.

4.1.2 Chapter outline

First, the terminology used in the chapter is presented by describing economic evaluation and its assessment in the presence of heterogeneity. Next, a description of how interactions can be generated through price negotiations is provided. Also, a framework to describe how drugs could be assessed in these circumstances is described. Two worked examples are presented, to illustrate the framework in more detail, and to compare results across evaluation approaches. The potential impact that strategic behaviour can have on maximum drug reimbursement prices is described. The chapter concludes with a discussion of the findings.

4.2 Economic evaluation with population heterogeneity

4.2.1 Economic evaluation

We examine the scenario where an applicant applies for the reimbursement of a new drug 'D' for the prevention of CV events. The decision maker must decide if 'D' should be reimbursed, or if standard of care 'S' should be maintained. As described in Section 2.1.1, a standard measure of quantifying the value of a drug is the ICER.

The ICER of drug 'D' vs comparator 'S' for the prevention of CV events is defined as:

$$ICER_D = \frac{C_D - C_S}{Q_D - Q_S} \quad (1)$$

where C_D, C_S and , Q_D , Q_S are the mean costs and QALYs of drug 'D' and standard of care 'S; with $C, Q \geq 0$. A decision rule may be defined such that an drug is considered cost-effective and reimbursed if the ICER of drug 'D' (versus the standard of care) is below the cost-effectiveness threshold (λ). However in practice, numerous other factors are also considered in the decision-making process (171).

4.2.2 Economic evaluation and heterogeneity

Heterogeneity can manifest in numerous forms. Drugs may be licensed for multiple indications. Populations defined by multiple different indications for the same drug may be considered an extreme form of population heterogeneity. Heterogeneity can also be considered as subgroups defined by population and clinical characteristics. Here, the term 'subpopulations' is used to represent both (i) subgroups of a population in receipt of a drug licensed for a single indication and (ii) multiple populations in receipt of a drug for different indications. The term 'total licensed population' represents the total eligible population.

4.2.2.1 *Stratified approach*

In instances of heterogeneity, the calculation of a single ICER which represents the cost-effectiveness of the drug in the total licensed population is insufficient. Using the stratified approach as described by Coyle et al (26), ICERs should be calculated for each subpopulation. For example, where the total licensed population could be divided into two distinct subpopulations based on their risk of CV events deemed 'low-risk' and 'high-risk' (j=1, 2), the ICER of drug 'D' versus standard of care 'S' in each subpopulation may be defined as:

$$ICER_{D,j} = \frac{C_{D,j} - C_{S,j}}{Q_{D,j} - Q_{S,j}} \; j = 1,2. \tag{2}$$

Drug 'D' should only be reimbursed for the subpopulations in which the ICER falls below the designated cost-effectiveness threshold.

4.2.2.2 *Weighted average approach*

In some instances, the option of restricting reimbursement to the subpopulation of the total licensed population where the drug is cost effective may not be available. This may be due to operational reasons or because of equity or ethical constraints (174). In these cases, decisions should be based on the mean total licensed population ICER. This may also be represented as a weighted average ICER of the drug in the two subpopulations - high-risk and low-risk (j=1,2):

$$ICER_D(Weighted\ Average) = \frac{\sum_j (C_{D,j} - C_{S,j}) \cdot n_j}{\sum_j (Q_{D,j} - Q_{S,j}) \cdot n_j} \; j = 1,2 \tag{3}$$

where the size of the subpopulation is given as n_j for $j = 1,2$. If the weighted average ICER falls below the cost-effectiveness threshold, the new drug should be reimbursed in both subpopulations. As restricted reimbursement is not available in this case, drug 'D' would not be reimbursed for any subpopulation if the ICER was above the cost-effectiveness threshold.

4.3 Strategic behaviour such as price negotiations can lead to interactions

4.3.1 Generation of the interaction

The cost-effectiveness of a drug and the reimbursement decision are critically dependent on the price of the drug (ρ). However, this may change during negotiations. If reimbursement is not granted in the entire population at the initial price offered ($\rho = full$), the company may offer a discount ($\rho = discount$). In these circumstances, the assessment of the cost-effectiveness of drug 'D' should be repeated at the discounted price using the methods described above.

However, strategic behaviour, adopted by a company, has the potential to add a further layer of complexity to the assessment process. Using the CVD example, instead of offering a simple

discount on a drug, the company may offer a conditional discount which is dependent on obtaining reimbursement in both subpopulations (high-risk and low-risk). If reimbursement is granted for one subpopulation only (high-risk or low risk), the full price would apply.

This offer generates an interaction between the two subpopulations. Interactions have been defined as situations where the absolute incremental costs and health benefits of one drug are affected by whether another drug is given (176). In this case, the price and cost-effectiveness of drug 'D' in one subpopulation is now dependent on the reimbursement decision in the other subpopulation.

Price negotiations and economic evaluations are critical components of the decision-making process in many jurisdictions. However, there has been little recognition of the interaction that is generated when a confidential discount is dependent on gaining reimbursement in multiple subgroups or indications, in the academic literature. The confidential nature of negotiations may have led to the neglect of this topic given there is little information in the public domain. But interactions mean that the common approaches to economic evaluation (including the stratified and the weighted average approaches) in the presence of heterogeneity are insufficient. Adoption of incorrect approaches may lead to incorrect conclusions regarding the cost-effectiveness of drugs.

4.3.2 Introduction of the hybrid approach

It is argued that in these scenarios, a hybrid approach is required which considers both the value of stratified decision making and the interaction between groups. Cost–effectiveness analyses generally take the form of a series of incremental comparisons of mutually exclusive drugs. This thesis proposes that incremental comparisons under the broader concept of mutually exclusive alternatives are required in these cases. Conceptually, this allows the additional consideration of the number of subpopulations treated and the price of drugs to vary across alternatives. This is an extension of work by Dakin et al which proposes joint analyses for combinations of drugs when such interactions are present (176).

Conceptualising the decision in this way, in the presence of an interaction, has important implications for cost-effectiveness analysis. Critically, where drug 'D' is cost-effective in one subpopulation (high-risk) at the full price, the cost savings realised from the discount in one subpopulation (high-risk) can be used to offset the incremental cost of extending reimbursement of drug 'D' to the other subpopulation (low-risk) at the discounted price. This increases the likelihood of a positive reimbursement decision in the second subpopulation

compared to the stratified approach. The ICER associated with extending reimbursement to the low-risk subpopulation using this approach may be represented by:

$$\frac{(C_{D,disccount,2}-C_{S,2})\cdot n_2+(C_{D,discount,1}-C_{D,full,1})\cdot n_1}{(Q_{D,2}-Q_{S,2})\cdot n_2+(Q_{D,1}-Q_{D,1)})\cdot n_1} \quad (4a)$$

Which is equivalent to:

$$\frac{(C_{D,discount,2}-C_{S,2})\cdot n_2+(C_{D,discount,1}-C_{D,full,1})\cdot n_1}{(Q_{D,2}-Q_{S,2})\cdot n_2} \quad (4b)$$

where j=1,2 represents the high-risk and low-risk subpopulations respectively. As there is no change in the expected QALYS in the high-risk population under both options (Equation 4a), the expected QALY in the high-risk group can be removed from the calculations (Equation 4b). Therefore, strategic behaviour, from a pharmaceutical company can directly influence the results of economic evaluations. A more detailed consideration of the hybrid approach under the proposed framework and an examination of the implications of adopting an incorrect approach is presented in two worked examples below.

4.4 Worked examples

The use of fictitious examples is continued given the confidential nature of price reductions. It is assumed that Drug 'D' is offered by a pharmaceutical company at an annual cost of €5,000 per patient (herein the 'full price', $\rho = full$) for the prevention of CV events. The pharmaceutical company also offers a discount, with a price reduction to an annual cost of €3,000 per patient (herein 'the discounted price', $\rho = discount$), if reimbursement is extended to both the high-risk and low-risk subpopulations.

Total costs and QALYs for each alternative are presented in Table 4.1 for the incorrect stratified and weighted average approaches and in Table 4.2 for the correct hybrid approach. Equal patient numbers in both subpopulations are assumed. Incremental analysis is shown graphically on the cost-effectiveness plane in Figure 4.1, Figure 4.2 and Figure 4.3 for the stratified, weighted average and hybrid approach respectively. Cost-effectiveness thresholds of €45,000 per QALY in Example 1 and €20,000 per QALY in Example 2 are assumed. It is shown that failure to adopt the correct assessment approach can lead to incorrect conclusions regarding a drug's cost effectiveness.

4.4.1 Example 1

4.4.1.1 Stratified approach

Under the stratified approach, the decision maker conceptualises the reimbursement decision problem as two distinct decision problems. Incremental comparisons of drug 'D' versus standard of care are conducted stratified by subpopulation and price offered. Analysis on the cost-effectiveness plane is presented in Figure 4.1. ICER calculations are calculated as described in equation (2) above and presented in Table 4.1. At the full price, the ICER of drug 'D' versus standard of care 'S' is €15,000 per QALY for the high-risk subpopulation and €70,000 per QALY for the low-risk subpopulation. As the ICER of drug 'D' for the high-risk subpopulation is below the cost-effectiveness threshold, it is reimbursed for this subpopulation However, even at the discounted price, the ICER for drug 'D' in the low-risk subpopulation is €50,000 per QALY. Therefore, drug 'D' is not reimbursed in the low-risk subpopulation. Savings in the high-risk subpopulation will be realised if drug 'D' is reimbursed in the low-risk subpopulation. However, these are not accounted for in the incremental cost of reimbursing drug 'D' in the low risk subpopulation.

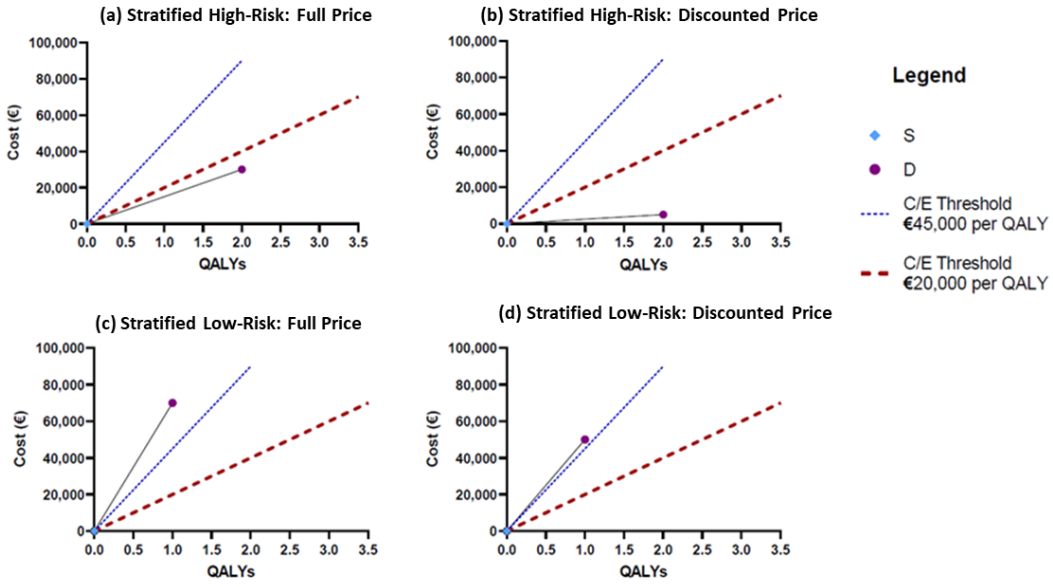


Figure 4.1 Incremental costs and QALYs of drug 'D' v. standard of care 'S' under the stratified approach by subpopulation and price for Examples 1 and 2

Drugs to the right of each cost-effectiveness threshold are considered cost-effective relative to that threshold. C/E, cost-effectiveness threshold; D, new drug; QALY, quality-adjusted life year; S, standard of care.

Table 4.1 Total and incremental costs and QALYS under the stratified and weighted average approaches in Examples 1 and 2 (equal patient numbers are assumed in both subpopulations)

Alternative	Description	Total Costs (€)		Total QALYs	Incremental Costs (€)		Incremental QALYs	ICER (€ per QALY)	
		ρ =Full	ρ =Discount		ρ =Full	ρ =Discount		ρ =Full	ρ =Discount
Stratified: High-Risk Subpopulation (j=1)									
S	Standard of care 'S' in high-risk subpopulation (j=1)	20,000	20,000	2	-	-	-	-	-
D	New drug 'D' for high-risk subpopulation (j=1)	50,000	25,000	4	30,000	5,000	2	15,000	2,500
Stratified: Low-Risk Subpopulation (j=2)									
S	Standard of care 'S' for low-risk subpopulation (j=2)	60,000	60,000	2	-	-	-	-	-
D	New drug 'D' for low-risk subpopulation (j=2)	130,000	110,000	3	70,000	50,000	1	70,000	50,000
Weighted Average: High-risk and Low-risk Subpopulations									
SS	Standard of care 'S' in high-risk and low-risk subpopulations	80,000	80,000	4	-	-	-	-	-
DD	New Drug 'D' for high-risk and low-risk subpopulations	180,000	135,000	7	100,000	55,000	3	33,333	18,333

ρ,price of Drug 'D', QALY, Quality Adjusted Life Year; ICER, Incremental Cost-Effectiveness Ratio; S, Standard of Care, - Not applicable.

4.4.1.2 Weighted average approach

Acknowledging the connections between the decisions in both subpopulations, the weighted average approach has also been presented. The decision maker conceptualises the decision problem as a single decision problem. This approach is presented graphically on the cost-effectiveness plane in Figure 4.2. Total costs and QALYs and corresponding ICERs are presented in Table 4.1. In this example, the ICER of drug 'D' versus standard of care in both subpopulations at the discounted price is €18,333 per QALY. Under this approach, drug 'D' would be reimbursed for both risk subpopulations. However, this approach does fully not account for the value of stratified decision making.

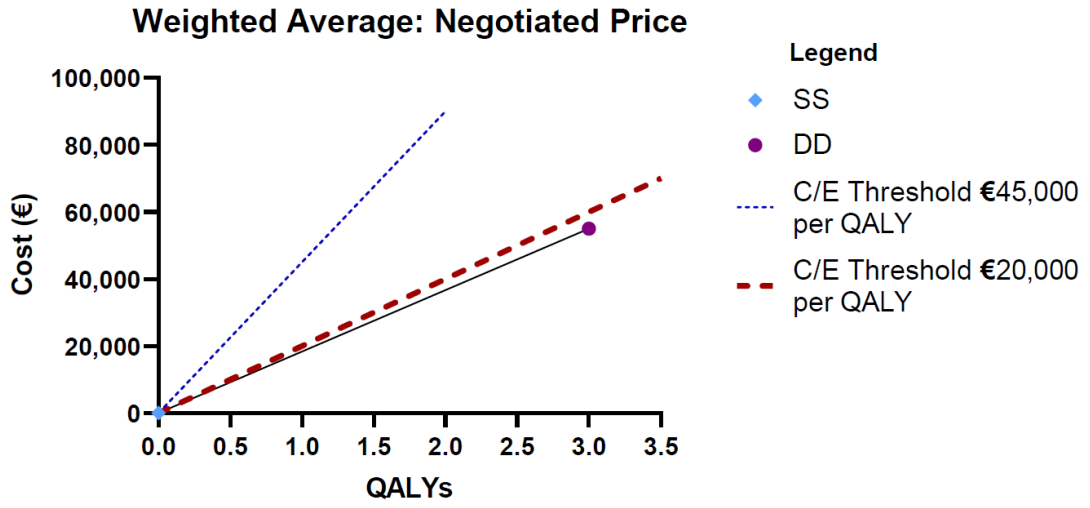


Figure 4.2 Incremental costs and QALYs of drug 'D' v. standard of care under the weighted average approach for the total licensed population at the negotiated price for Examples 1 and 2

Drugs to the right of each cost-effectiveness threshold are considered cost-effective relative to that threshold.
C/E, cost-effectiveness threshold; DD, new drug D reimbursed in both high-risk and low-risk subpopulations at the discounted price; QALY, quality-adjusted life year; SS, standard of care maintained in both high-risk and low-risk subpopulations.

4.4.1.3 Framework for the hybrid approach

Having described common approaches to assessing cost-effectiveness in the presence of heterogeneity, a framework for assessment under the hybrid approach is provided.

4.4.1.3.1 Framework step 1: Frame the decision problem

The decision-maker faces two decisions regarding drug 'D':

1. to reimburse the new drug 'D' or maintain the standard of care 'S' for the high-risk subpopulation(j=1),
2. to reimburse drug 'D' or maintain the standard of care 'S' for low-risk subpopulation(j=2).

The decisions are not independent if a price reduction contingent on reimbursement in both subpopulations is offered. Therefore, the decision problem must be reframed into all mutually exclusive alternatives as outlined in Table 4.2. The decision maker may decide to not reimburse drug 'D' in any subpopulation (SS), reimburse drug 'D' in the high-risk subpopulation only (DS), reimburse drug 'D' in the low-risk subpopulation only (SD), or reimburse drug 'D' for both subpopulations (DD). The discounted price only applies in the final alternative (DD). The four alternatives considered may be plotted on the cost-effectiveness plane as shown in Figure 4.3.

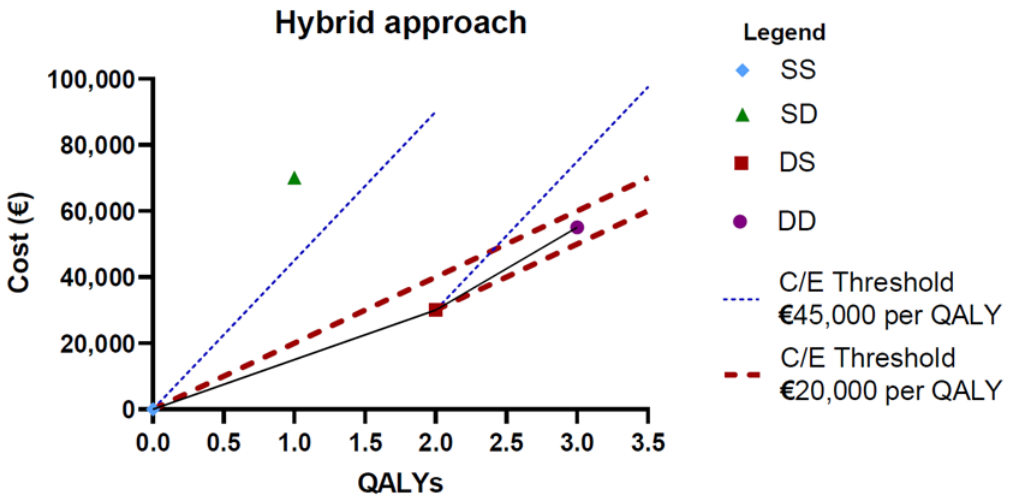


Figure 4.3 Incremental analysis of alternatives under the hybrid approach for Examples 1 and 2.

Drugs to the right of each cost-effectiveness threshold are considered cost-effective relative to that threshold. C/E, cost-effectiveness threshold; DD, new drug D reimbursed in both the high-risk and low-risk subpopulations at the discounted price; DS, drug D reimbursed in high-risk subpopulation and standard of care maintained in low-risk subpopulation; QALY, quality-adjusted life year; SD, standard of care maintained in high-risk subpopulation and new drug D reimbursed in low-risk subpopulation; SS, standard of care maintained in both high-risk and low-risk subpopulations

4.4.1.3.2 Framework step 2: Perform an incremental comparison of alternatives

Using standard decision rules, an incremental comparison of all mutually exclusive reimbursement alternatives is conducted. Alternatives are placed in order of increasing effectiveness. This is presented in Table 4.2. In this example, alternative SD is dominated by DD. Therefore, alternative SD is excluded from further consideration. ICERs are calculated for remaining alternatives. The ICER associated with reimbursing drug 'D' in the high-risk subpopulation only is €15,000 per QALY (DS vs SS). As drug 'D' is cost-effective in the high-risk subpopulation, the ICER associated with extending reimbursement to both subpopulations is now examined (DD vs DS). As the ICER of €25,000 per QALY is below the cost-effectiveness threshold of €45,000 per QALY, the final decision is to reimburse drug 'D' in both subpopulations at the discounted price. The ICER includes the incremental savings in the high-risk subpopulation that are realised by reimbursement of drug 'D' in the high-risk subpopulation at the discounted price compared to the full price. While in this example, the reimbursement decision is consistent with the weighted approach, this will not always be the case. Further, an approach, whereby the stratified approach is conducted first followed by the weighted average approach, is also insufficient. This scenario is explored in Example 2.

Table 4.2 Total and incremental costs and QALYS for Examples 1 and 2 under the hybrid approach

Abbreviation	Description	Price of drug 'D' (ρ)	Total Costs (€)	Total QALYs	Incremental Costs (€)	Incremental QALYs	ICER (€ per QALY)
SS	Drug 'D' is not reimbursed for low-risk or high-risk subpopulations. Standard of care 'S' is maintained the entire population.	N/A Drug 'D' is not reimbursed.	80,000	4	-	-	-
SD	Drug 'D' is reimbursed for low-risk subpopulation only (j=2). Standard of care 'S' is maintained in high-risk subpopulation (j=1)	$\rho = full$	150,000	5	Dominated	Dominated	Dominated
DS	Drug 'D' is reimbursed for high-risk subpopulation (j=1) only. Standard of care 'S' is maintained in low-risk subpopulation (j=2).	$\rho = full$	110,000	6	30,000	2	15,000
DD	Drug 'D' is reimbursed in both low risk and high-risk subpopulations (j=1,2).	$\rho = discount$	135,000	7	25,000	1	25,000

Equal patient numbers are assumed in both subpopulations.
D, New Drug; ICER, Incremental Cost-Effectiveness Ratio; ρ,price of Drug 'D', QALY, Quality Adjusted Life Year; S, Standard of Care.

4.4.2 Example 2

In this example, a cost-effectiveness threshold of €20,000 per QALY is assumed. All other parameters are defined as Example 1. Therefore, ICER calculations presented in Table 4.1 and Table 4.2 are still relevant here. Under the stratified approach, drug 'D' is cost-effective in the high-risk risk subpopulation but not cost-effective in the low-risk subpopulation at both the full and discounted prices. If a weighted average approach is taken by comparing DD (reimbursement in both subpopulations) to SS (maintaining standard of care in both subpopulations), drug 'D' should be reimbursed in both subpopulations at the discounted price. In this example, the decision outcome, under both the weighted average and stratified approaches at a cost-effectiveness threshold of €20,000 per QALY, is identical to that at a threshold of €45,000. If both approaches were taken consecutively, i.e. the stratified approach followed by the weighted average approach, the reimbursement outcome in this case would be to reimburse in both subpopulations.

However, decision outcomes under the hybrid approach differ. As the ICER of DS relative to SS is €15,000 per QALY, alternative DS is cost-effective and drug 'D' should be reimbursed in the high-risk subpopulation. But the ICER associated with extending reimbursement to the low-risk subpopulation (DD vs DS) is €25,000 per QALY. As this is above the cost-effectiveness threshold of €20,000 per QALY, the final reimbursement decision should be DS – reimbursement in the high-risk subpopulation only at the full price. When interactions are present between groups, adoption of the hybrid approach is necessary to correctly classify cost-effective and ineffective drugs.

4.4.3 Further examples

In the previous examples, drug 'D' was always cost-effective in one subpopulation at the full price. However, this may not always be the case. Figure 4.4 presents two examples of incremental analysis on the cost-effectiveness plane using the hybrid approach where a company offers a conditional discount dependent on reimbursement in both the high-risk and low-risk subpopulations. In Figure 4.4 (a) drug 'D' is not cost-effective in either subpopulation at the full price. If reimbursement is granted to both subpopulations, drug 'D' is still not cost effective despite a discount. The final reimbursement decision is SS. Standard of care remains the treatment of choice in both the high-risk and low-risk subpopulations. In Figure 4.4 (b), the discount offered was sufficient to enable reimbursement of drug 'D' in both subpopulations despite being not-cost-effective in either subpopulation at the full price. It is important to note that the ICERs calculated under the hybrid approach represent the ICERs associated with treatment alternatives in the total eligible population. This is not equivalent to the ICER of drug

'D' in the subpopulation of interest. If an irreversible decision has already been made to reimburse drug 'D' for one subpopulation e.g. high-risk (j=1), the ICER for drug 'D' in the low-risk subpopulation (j=2) can be calculated by comparing DD vs DS.

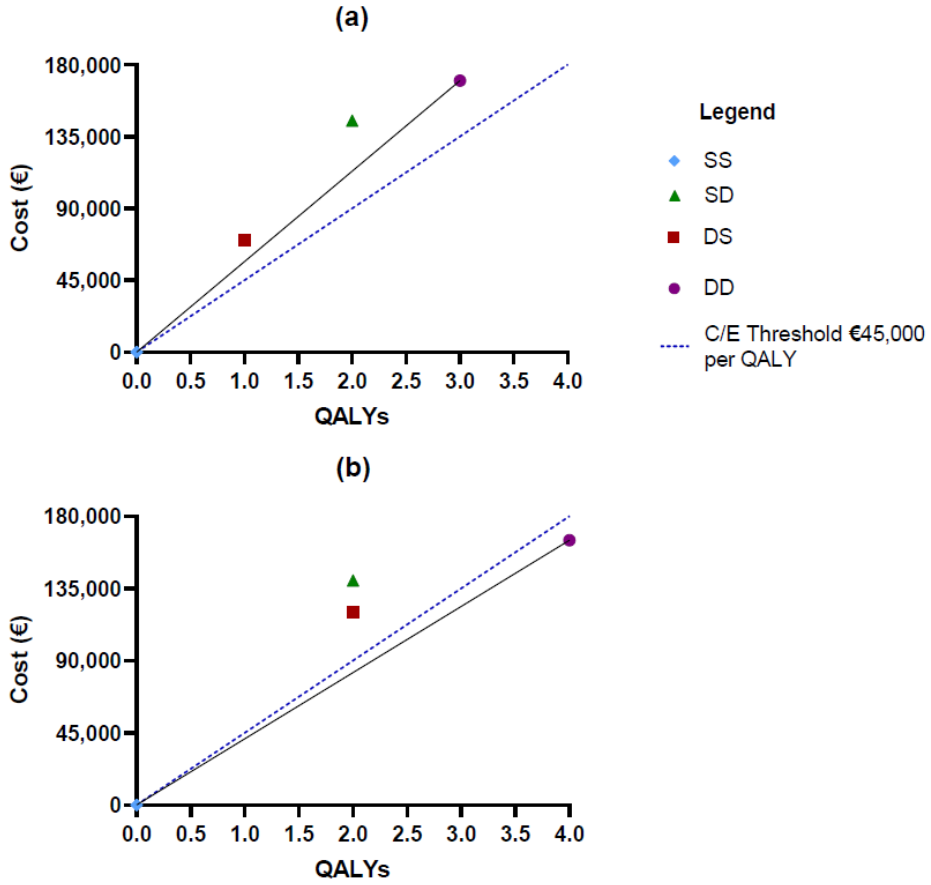


Figure 4.4 Further examples (a) and (b) of potential reimbursement outcomes under the hybrid approach.

In Figure 4.4 (a) drug D is still not cost effective despite a discount. In Figure 4.4 (b) the discount offered was sufficient to enable reimbursement of drug 'D' in both subpopulations despite being not-cost-effective in either subpopulation at the full price.
Drugs to the right of each cost-effectiveness threshold are considered cost-effective relative to that threshold.
C/E, Cost-effectiveness threshold; DD, new drug D reimbursed in both the high-risk and low-risk subpopulations at the discounted price; DS, drug D reimbursed in the high risk subpopulation with standard of care maintained in the low-risk subpopulation; QALY, quality-adjusted life year; SD, standard of care in the high risk subpopulation and drug D reimbursed in the low-risk subpopulation; SS, standard of care maintained in both the high-risk and low-risk subpopulations.

4.5 Strategic behaviour can increase the reimbursement price of new drugs

The results show that strategic behaviour can impact the results of economic evaluation and therefore influence the maximum price that companies can achieve for their drugs. In Example 1, the company achieved reimbursement for both subpopulations at a price of €3,000 per patient per year when adopting the strategic approach. However, if the company offered a price of €3,000 per patient per year in the first instance without any conditions, this would not have been achieved. In this scenario, there is no interaction between subpopulations and the stratified approach is the correct assessment approach. As the ICER of drug 'D' at €3,000 per patient per annum in the low-risk subpopulation is above the cost-effectiveness threshold in this scenario, reimbursement would only be granted to the high-risk subpopulation. Further price reductions below €3,000 would be required to gain reimbursement in the low-risk subpopulation.

The impact that strategic price negotiations can have on the final reimbursement price depends on several factors including the relative size of the two subpopulations, the sensitivity of the ICERs in both subpopulations to drug costs and the difference in ICERs between the two subpopulations.

4.6 Discussion

4.6.1 Main findings

It has been identified that interactions can be created through strategic price negotiations and a framework for the quantification of the cost-effectiveness of drugs in these scenarios has been presented. While this approach will not be applicable to every cost-effectiveness evaluation, the issues raised here are relevant given the critical role that price negotiations and economic evaluations play in decision making for drugs.

4.6.1.1 Interactions

Dakin and Gray have previously highlighted the importance of considering interactions in economic evaluations (176). Their acknowledgement of interactions in different patient groups was limited to patients treated in the same healthcare facilities. It is acknowledged that naturally occurring interactions between different patient groups are rare. However, given the intrinsic link between price negotiations and economic evaluations, interactions between groups are likely to be substantially more common than that considered in the literature to date. Like Dakin and Gray, it is argued that interactions should not be ignored in practice given their ability to change conclusions regarding the cost-effectiveness of an drug (176).

4.6.1.2 Consideration of future price reductions

Grimm et al examined the impact of future price reductions in cost-effectiveness analysis (177). They state that when assessing the cost-effectiveness of interventions such as medical devices, potential future price reductions brought about through increased use of the drug should be considered. They argue that such price reductions may not occur if the drug is not reimbursed and that re-evaluating at a future time point will not address the deficiency. They did not consider the analysis to be relevant to pharmaceuticals as prices are generally not related to production volumes. This work differs in that the price reductions considered are certain if the drug is reimbursed. In this analysis, price reductions are used to reduce incremental costs in a distinct subpopulation as opposed to reducing incremental costs in future incident cohorts in the same subpopulation.

The examples presented here consider the scenario where a price reduction in one subpopulation would reduce the incremental costs of extending reimbursement to another. Theoretically, the incremental cost could also be reduced, if the company proposes increasing the price in a subpopulation for which it is already reimbursed, if reimbursement is not extended. However, given that a decision-maker would be reticent to accept a price increase and that it would reopen a decision to reimburse the drug at all; this scenario is unlikely in practice. Therefore, we do not consider it further. In theory, the price reduction could be offered for an entirely separate drug if the producer is also able to control its price. However, in practice it may not be feasible given the additional uncertainty that would arise. An additional cost-effectiveness evaluation would be required.

4.6.1.3 Extended dominance

The standard decision rules of any cost-effectiveness analysis, where extended dominance is observed, rely on the indivisibility property (178). This means that cost-effectiveness is not affected by the proportion of the eligible patient population receiving that treatment. The indivisibility property fails in this case as the total costs depend on the proportion of the total population receiving the drug. An example of extended dominance is observed in Figure 4b, where alternative DS is extendedly dominated by DD. As the indivisibility property fails, the incremental costs and effects of any alternatives comprising of a combination of SS and DD does not lie on the line joining them. This is the case for those which are explicitly modelled (such as DS and SD) or considered implicitly (i.e. a combination between DS and DD). It has been previously recognised that standard decision rules do not strictly apply where the indivisibility property fails but that standard decision rules are still very good approximations and differences are of minimal practical significance (179). This is because the problematic issues only arise

when the ICER of the chosen alternative is at the cost-effectiveness threshold and the budget available is only enough to allow the implementation of the chosen alternative in a portion of the population. This scenario is unlikely to be realised in practice, as the costs of most individual programs are small relative to total budgets (179). Therefore, given the complexity of the alternative, we consider the rules here sufficient and we do not consider these issues further.

4.6.1.4 Multi-indication pricing

It is acknowledged that decision making may be challenging where there is significant patient heterogeneity. If a drug is cost-effective in a large indication but not cost-effective in a small indication with high unmet clinical need, there may be pressure to reimburse in both despite the lack of cost effectiveness. Producers may be unwilling to reduce the price of the drug sufficiently to make the drug cost-effective under the stratified approach due to the profit losses that would be incurred in other larger indications (180). Application of the weighted average method in this case ignores the value of stratified decision making. Increased awareness of the implications of interactions may encourage price negotiations which may facilitate reimbursement in the indication associated with the smaller population. Multi-indication pricing is another method proposed to expand access to smaller populations where the drug is not cost-effective. But this method may be difficult to implement in practice and requires additional resources to ensure that the agreed prices are paid for each indication. However, these methods are advantageous to producers who may extract all possible economic surplus from provision of the drug (180). Adoption of the hybrid approach may allow some surplus to remain with the decision-maker.

4.6.1.5 Effect on population health.

Adoption of the framework is expected to increase population health through the increased recognition of cost-effective drugs. However, this conclusion does not consider the counterfactual that providers may present higher prices for drugs which would result in increased opportunity cost and a reduction in overall population health. Therefore, the net impact on population health is uncertain.

4.6.1.6 Strategic behaviour and game theory

To our knowledge, this is the one of the first pieces of research to examine the role of strategic behaviour in economic evaluations. Johannessen et al previously characterised the design of the appraisal process of pharmaceuticals as an interactive game when where appraisal organisation need to consider companies reactions to the design and operation of the appraisal process if pricing is endogenous to the appraisal process (22). Paulden has previously examined strategic

behaviour in relation to the cost-effectiveness threshold (181). He proposed that if the decision maker's primary concern is consumer surplus, they should adopt a strategy where cost-effectiveness thresholds are lower than conventionally specified in theory. His work accounts for company strategic behaviour in setting drug prices and placing drugs on the market. Kim and Basu did not use the term strategic behaviour in their paper. However, their proposed use of patient co-pays to alter uptake and therefore the indication weights to calculate an ICER could also be considered as an examination of the effect of strategic behaviour on the assessment of cost-effectiveness (28).

It is proposed that there may be a role for game theory for future study in this area. Game theory is concerned with the analysis of strategies for dealing with competitive situations. Its application was previously described in the context of deriving a conceptual framework for the design of the cost-effectiveness appraisal process of new drugs (22). While it could always have been applied to negotiations between pharmaceutical companies and decision-makers, this work adds to the growing literature that the assessment of cost-effectiveness itself is affected by strategies adopted. Hence, companies now have additional considerations when devising reimbursement strategies.

4.6.2 Limitations

There are several limitations associated with the hybrid approach and its presentation here. Adoption of the hybrid framework marginally increases assessment complexity. However, similar data is also required for the weighted average approach. Here, the simplifying assumption is made that relative size of the two subpopulations is constant. But, the relative size of the subpopulations is dynamic as the rate of uptake and discontinuation in each group changes. In these examples, only two subpopulations are considered with one negotiating condition. More complex scenarios are possible. Furthermore, the costs of stratification and restricting reimbursement are not counted (174). For simplicity, it is assumed that the only criterion for decision making is cost-effectiveness. But decision-makers involve more criteria such as the strength of the clinical evidence, uncertainty, budget impact and unmet need (171). However, none of these limitations materially impact the premise of the framework or the implications presented here.

4.6.3 Future Research

A recent systematic review found that guidance in the specific application of methods for assessing cost-effectiveness in the presence of heterogeneity was scarce (182). Economic guidelines should be updated to account for the use of this framework and the broader

consideration of treatment alternatives that this requires. Further research, to examine the prevalence of scenarios where the hybrid approach can be applied in practice and to examine the impact of this method on reimbursement decisions would be beneficial. Further, the role of strategic behaviour and game theory in economic evaluation should be explored.

4.7 Conclusion

It is proposed that this framework should be employed when decision making is linked to price negotiations and when decision-makers are able to restrict reimbursement by subgroup or indication. By adopting this framework, cost-effective drugs are identified that may have been previously misclassified as not being cost-effective and vice versa. Further, recognition of the interaction by pharmaceutical companies may influence the forms of discounts offered to decision makers. Therefore, this research may have far-reaching effects on medical decision making.

Chapter 5. Bridging the data gap: Mapping TILDA to the EQ-5D-3L

5.1 Introduction

Economic evaluations in healthcare require an assessment of the comparative effectiveness of interventions. The QALY is one of the most widely used outcome measures. QALYs combine both length of life and HRQoL into a single outcome. There are numerous methods of quantifying HRQoL, however, in line with many international guidelines, Irish guidelines for economic evaluation recommend a generic preference-based instrument (8). The EQ-5D-3L is preferred by many international HTA agencies including the NCPE in Ireland (183–185).

In order to perform an economic evaluation of PCSK9 inhibitors, utilities for relevant health states must be sourced. It may be possible to use utilities directly from a relevant clinical study if the EQ-5D-3L or another appropriate preference-based instrument was used. As stated in Chapter 3 the EQ-5D was included as outcome in the ODYSSEY OUTCOMES study, however, results have not been reported (6,7). Therefore, trial data from PCSK9 inhibitor trials is unavailable to answer this research question.

Where an alternative outcome measure was included, it may be possible to map this data to an appropriate instrument by applying a mapping model (186,187). The model is derived by using regression techniques in an alternative sample where both measures were used. A large number of mapping models are available (188). However, no other patient reported outcomes measures are reported for the FOURIER or ODYSSEY OUTCOMES trials (6,7). An alternative approach is to source summary utilities from the literature (189,190).

However, there are limitations associated with all approaches. The study population (in which the target utility was obtained) may not be representative of the population under consideration. Baseline health levels may differ between countries. Data for age, sex or other important subgroups may not be reported. There are numerous specific difficulties in estimating utility values in cardiovascular disease. Causal utility decrements after cardiovascular events are rarely reported. It is unclear how utility should be adjusted after multiple cardiovascular events. People live with the long term effects of cardiovascular disease for many years, but there is little information to describe how utility should change over time (191). It can be difficult to account for these issues without access to patient-level data. All of these issues increase uncertainty in economic evaluations.

Like many other jurisdictions, there are few studies collecting local health preference data in Ireland (192–194). There is a need for a local database of utility estimates that could be analysed for economic evaluation. The ideal database would be large, longitudinal, nationally representative and would capture a wide range of sociodemographic and clinical variables.

TILDA fulfils many of these requirements and is one of the most comprehensive research studies of its kind both in Europe and internationally. As a nationally representative, longitudinal cohort study, it collects detailed data on the social, economic and health characteristics of over 8,000 community dwelling adults age 50 years and older (195,196). While only designed to capture information on those over 50 years, this is the group which faces the greatest ill health burden. Since 2009, TILDA researchers have invited participants to complete a face-to-face interview and a self-completion questionnaire every two years and an objective health assessment every three to four years. To date, five waves of data have been collected from the same cohort of participants and more waves are planned. The anonymised datasets are publicly available for research purposes (197).

The EQ-5D-3L was not collected in TILDA. However, detailed information was collected on areas that are reflective of the EQ-5D-3L domains including mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Therefore, to reduce the data gap, we propose the use of mapping from national studies such as TILDA to the EQ-5D-3L. We hypothesise that, in the presence of strong conceptual overlap between variables, mapping can generate utilities in the survey population of interest.

5.1.1 Aim

The aim of this chapter to derive a mapping model between questions administered in TILDA and the EQ-5D-3L. To our knowledge, this is the first time a mapping study from a national longitudinal survey to a generic preference-based instrument to address an important data gap has been performed. Derivation of the mapping model would unlock the key to a longitudinal patient-level database of predicted utility values connected to a rich supply of social, economic and health related variables.

5.1.2 Chapter outline

First, we briefly outline the use of mapping in the literature. Next, general methodological considerations for mapping are described coupled with the methods employed in this study. Then, the results of this mapping study are presented. In the discussion, we describe the main limitations of our study and opportunities for future research.

5.2 Mapping in the literature

Mapping is the development and use of a statistical model to predict heath-state utility values using data on other indicators or measures of health (198). It is acknowledged that mapping is not a replacement for direct utility measurement as it results in information loss and increased

uncertainty (199). However, it can provide useful information when direct information is unavailable.

Indirect information could include physical outcomes measures or patient reported outcome measures (PROMs) (200). Physical outcomes are those which can be measured objectively and do not require interpretation by the patient. Preference-based PROMs have an associated utility score for the calculation of QALYs. While non-preference-based PROMs provide information on the health state of an individual, they do not have an attached scoring system. Both PROMs and physical outcomes measures can be disease specific or generic (200).

National guidelines for economic evaluation state that mapping may be performed in the absence of utility data from generic instruments (8). Detailed guidance is not provided but detail is provided regarding the need for clear description of the regression model and justification for use of explanatory variables (8). There are numerous mapping guidelines in the literature including NICE Decision Support Unit guidelines (198), Mapping onto Preference-based measures reporting Standards (MAPS) reporting guidelines (201,202) and an International Society for Pharmacoeconomics and Outcomes Research (ISPOR) taskforce report (200). The ISPOR report includes recommendations for all stages in the mapping process including pre-modelling, modelling and data-analysis and reporting. The increased use of the EQ-5D-5L and valuation studies means that the importance of mapping between descriptive systems and value sets is likely to increase going forward.

Box 5.1 Summary of pre-modelling recommendations by Wailoo et al (200)

1. Consider the use or potential uses of the mapping:
    a. Is it for use in a cohort decision model, patient-level model, or trial-based cost-effectiveness analysis?
    b. What are the health states that require utility estimates from the mapping and how do they relate to the PBM?
    c. What is the range of disease severity for which utility values are required?
2. Provide a descriptive account of the clinical explanatory variable, the dependent PBM, and the extent to which they overlap.
3. Assess if a regression-based mapping is required.
    a. How many health states require estimates of utility?
    b. Are there additional covariates of importance?
    c. Are there sufficient observations within each category?
4. Identify if more than one data set is potentially available for estimation. Compare the characteristics of candidate data sets.
5. To what extent does the distribution of patient characteristics in the sample data sets reflect those that are the subject of the cost-effectiveness analysis? In particular, are all extremes of disease severity represented?
6. Is the type of treatment a patient receives likely to influence the relationship between health utility and clinical outcome measures?

PBM, Preference-based measure.

Box 5.2 Summary of modelling and data analysis recommendations by Wailoo et al (200)

1. Consider whether the cost-effectiveness analysis requires a formal regression-based mapping model approach, or if it is suitable to take the mean value for subsamples of patients.
2. If regression is required, then model selection should be based on the following:
    a. Consideration of the most straightforward statistical model type whose assumptions are compatible with the target utility instrument. Use a plot of the distribution of the utility data to help inform that choice.
    b. Existing empirical evidence of the performance of different methods. There is no reason for this to be restricted to evidence from any specific disease area.
    c. The type of cost-effectiveness analysis in which the mapping will be used and the extent to which biased estimates will affect the results.
3. For response mapping, models that respect the ordered nature of the categorical data in the descriptive system should be selected. Expected values should be calculated analytically.
4. Selection of the preferred mapping model is an iterative process that should conform to good practice common to all regression analyses.
5. Covariates should be theoretically justified a priori. Exclusion of covariates, even if they are not to be used in the cost-effectiveness model, risks misspecification.

Box 5.3 Summary of reporting of mapping studies recommendations by Wailoo et al (200)

1. Describe relevant differences between data sets that are candidates for mapping estimation.
2. Give full details of the selected data set. Describe how the study was run and patients were sampled. Provide baseline and follow-up characteristics including the distribution of patients' disease severity. Missingness in the longitudinal pattern of responses should be described.
3. Plot the distribution of the utility data.
4. Justify the type of model(s) selected with reference to the characteristics of the target utility distribution and the proposed use of the mapping function.
5. Compare the dimensions of health covered by the target utility instrument and those covered by the explanatory clinical measure(s).
6. Describe the approach to determining the final model. Include tests conducted and judgments made.
7. Summary measures of fit are of limited value for the total sample. Provide information on fit conditional on disease severity as measured by the clinical outcome measure(s). A plot of mean predicted versus mean observed utility conditional on the clinical variable(s) should be included.
8. Coefficient values, error term(s) distributions(s), variances, and covariances are required.
9. Provide an example predicted value for some sets of covariates. Consider providing a program that calculates predictions for user-defined inputs.
10. Parameter uncertainty in a mapping regression should be reflected using standard methods for PSA. Assessment of model suitability for use in cost-effectiveness analysis should also consider the distribution of utility values for PSA, with particular focus on whether these lie outside the feasible utility range for the PBM.
11. When imputing data from a mapping function, individual-level variability should be incorporated using simulation methods and information about the distribution of the error term(s). These simulated data can be compared with the raw observed data, including an assessment of the range of values compared with the feasible range for the PBM.
12. Re-estimation of mapping results in a separate data set or other forms of validation are not routinely required

PBM: Preference-based measure; PSA, Probabilistic Sensitivity Analysis.

A systematic review of mapping studies between non-preference-based measures to generic preference-based measures was published by Brazier et al (203) and updated by Mukuria et al in 2019 (188). Studies mapping between preference-based measures were excluded. The aim of the review was to assess the reliability and validity of mapping approaches. Brazier et al located 30 studies describing 119 models (203). Sample sizes in mapping studies ranged from 68 to over 23,000. Most mapping functions were estimated using ordinary least squares (OLS) regression. Brazier et al highlight that the purpose of mapping functions is to predict health state utility values in other datasets (203). Therefore, the model should be assessed using the accuracy of its predictions rather than in terms of its explanatory power (for example using R^2). But this was not considered appreciated sufficiently in most of the studies identified and the performance of mapping functions was variable.

The updated review by Mukuria et al located 180 papers describing 233 models (188). The authors reported that there was some evidence in advancement of methods since the previous review. The most common model types used to map to EQ-5D based value sets include OLS (89%), Censored Least Absolute Deviation (CLAD) (24%), Tobit (22%), and response mapping (22%)[6]. As the standard regression method, it is not surprising that OLS is the most common model type investigated. However, no information was provided by Mukuria et al on which model type performed best or was most preferred by modellers. Information on these issues would be more informative (188). Across studies, predictive availability was primarily assessed using mean absolute error (MAE) or root mean squared error (RMSE). But less than 50% of studies assessed performance across disease severity and only 52% plotted predictions (which is considered best practice) (200). Mukuria et al included studies published over the last 10 years (188). It is expected that the availability of ISPOR guidelines may improve methods going forward (200).

A database of mapping studies is maintained by Health Economics Research Centre in Oxford (199,204,205). Unlike the review inclusion criteria defined by Brazier et al and Mukuria et al (188,203), the database is restricted to studies mapping to the EQ-5D. But studies mapping between preference-based instruments are included. As part of this work, Dakin et al conducted a critical appraisal of studies in the database published before December 2016, versus the MAPS statement (199). The MAPS statement is a checklist of 23 items which aims to promote complete and transparent reporting of mapping studies (201,202). It is not an assessment of study quality.

[6] Proportions (%) add to more than 100% as multiple estimation methods were used in 61% of studies.

Mapping studies reporting 190 models from 110 source instruments to the EQ-5D were included in the review. Seventeen papers published in 2016 were assessed for reporting quality using the full MAPS checklist. There was substantial variation in the quality of reporting between papers. A large proportion of papers reported most criteria, but only one paper fulfilled all the criteria. An assessment of the quality of reporting in study abstracts across all study years showed that the mean assessment score for abstracts increased over time. Studies in the musculoskeletal disease area accounted for the greatest number of mapping studies at 28%; studies in patients with cardiovascular disease only accounted for 7%.

The mapping database is updated annually. There were thirteen entries for cardiovascular disease in the most recent version 7.0, which is based on search results conducted in January 2019 (204). The results are summarised in Table 5.1 below. A variety of outcome measures were used as explanatory variables including disease specific quality of life instruments such as the Seattle Angina Questionnaire and disease specific physical outcome measures such as the modified Rankin Scale for stroke. Generic HRQoL measures such as the SF-36 are also included. OLS was the most common model type investigated.

We are not aware of any study which has designed a mapping study to address a data gap in a national survey. Ghatnekar et al is the closest conceptually; here mapping health outcome measures from the national Swedish stroke registry to EQ-5D weights was examined (206). Variables conceptually related to the EQ-5D from the registry questionnaire were considered as independent variables in the model including requiring assistance with toileting or dressing, restricted mobility, general health and mood. The available data was split equally into an estimation set and a validation set (n=272 for each). Numerous model types were investigated including OLS, Tobit and CLAD. The preferred model type was not specified. They concluded that it was possible to map non-validated health outcome measures from a stroke register into a preference-based utility. They anticipated that it would be useful to study the development of stroke care over time and to compare utility across conditions. Mapping is normally conducted in response to a known data gap. But in this case, the EQ-5D was included in the stroke registry. Therefore, the primary value of the mapping model is speculative.

Next, we describe the methods used in our mapping study.

Table 5.1 Extract of mapping studies to the EQ-5D-3L, in CVD, from the Health Economics Research Centre mapping database (199,204,205)

Citation details	Quality of Life Measures	Disease or patient group	No. observations in estimation sample	Mapping models investigated
Ara et al (2014) (207)	General health, acute sickness and demographic variables	Cardiovascular disease	7,998	OLS, Response Mapping
Calvert et al (2005) (208)	Minnesota Living with Heart Failure Questionnaire	Heart failure	813 patients	Linear mixed models
Chen et al (2015) (209)	MacNew Heart Disease Quality of Life Questionnaire (instrument)	Coronary heart disease	943	OLS; GLM; robust MM-estimator
Edlin et al (2002) (210)	Minnesota Living with Heart Failure Questionnaire	NYHA Class II-IV heart failure patients	22,931	OLS; response mapping
Ghatnekar et al (2013) (206)	Stroke outcome measures not restricted to validated instruments	Stroke	272	OLS; CLAD; Tobit
Goldsmith et al (2010) (211)	Clinical outcome measures and demographic variables, including Seattle Angina Questionnaire	Cardiovascular disease	2,855	OLS
Longworth L et al (2005) (212) and Longworth (2007) (213)	Breathlessness Grade and Canadian Cardiovascular Society classification of angina and number of drug classes used	Coronary artery disease	503	OLS; Tobit; response mapping
Longworth, L. (2007) (213)	SF-36	Coronary artery disease	423	Response mapping
Oddershede et al (2014) (214)	5 visual analogue scales measuring patients' self-reported mobility, self-care, ability to perform usual activities, pain, and anxiety or depression	Coronary artery bypass graft recipients	233	CLAD; Tobit; GLS with random intercepts
Rivero-Arias et al (2010) (215)	Modified Rankin Scale (mRS)	Stroke and TIA	2,425	OLS; response mapping
van Exel et al (2004) (216)	Barthel index	Stroke	710	OLS
Whynes et al (2013) (217)	Modified Rankin Scale (mRS)	Stroke	1462 patients	OLS
	Modified Rankin Scale (mRS), Barthel Index and Zung Depression	Stroke	1462 patients	OLS
Wijeysundera et al (2011) (218)	Seattle Angina Questionnaire	Coronary artery disease	1,555	OLS; Tobit

CLAD, Censored Least Absolute Deviation; GLM, generalised linear model; GLS, Generalised Least Squares; OLS, Ordinary Least Squares; mRS, modified Rankin Scale; NYHA, New York Heart Association; TIA, Transient Ischemic Attack.

5.3 Methods

5.3.1 Overview of study design

Mapping methods were carried out and reported in accordance with international guidelines described in Section 5.2 (200–202). Mapping requires a dataset which captures the target variable (EQ-5D-3L) and the source variables (TILDA questions) in the same population. Therefore, this data was collected in a new sample (herein the 'Independent Sample'). Participants were asked to complete a subset of TILDA questions with likely conceptual overlap to the EQ-5D-3L (herein the 'Question Subset') and the EQ-5D-3L. Using this data, regression methods were applied to generate a mapping model which quantifies the relationship between them. The mapping model was then applied to Wave 1 of the national TILDA dataset to generate predicted EQ-5D-3L utilities for the population. Using survey data analysis methods, individual level data from TILDA (including the predicted EQ-5D-3L utility values) was analysed to predict utilities for a range of cardiovascular conditions. A pictorial representation of the study design is presented in Figure 5.1.

Ethical approval was obtained from the Tallaght University Hospital / St James's Hospital Joint Research Ethics Committee for derivation of the mapping model. All participants provided written informed consent to take part in study. Ethical approval for the entire TILDA study is given by Faculty of Health Sciences Research Ethics Committee of Trinity College Dublin for TILDA.

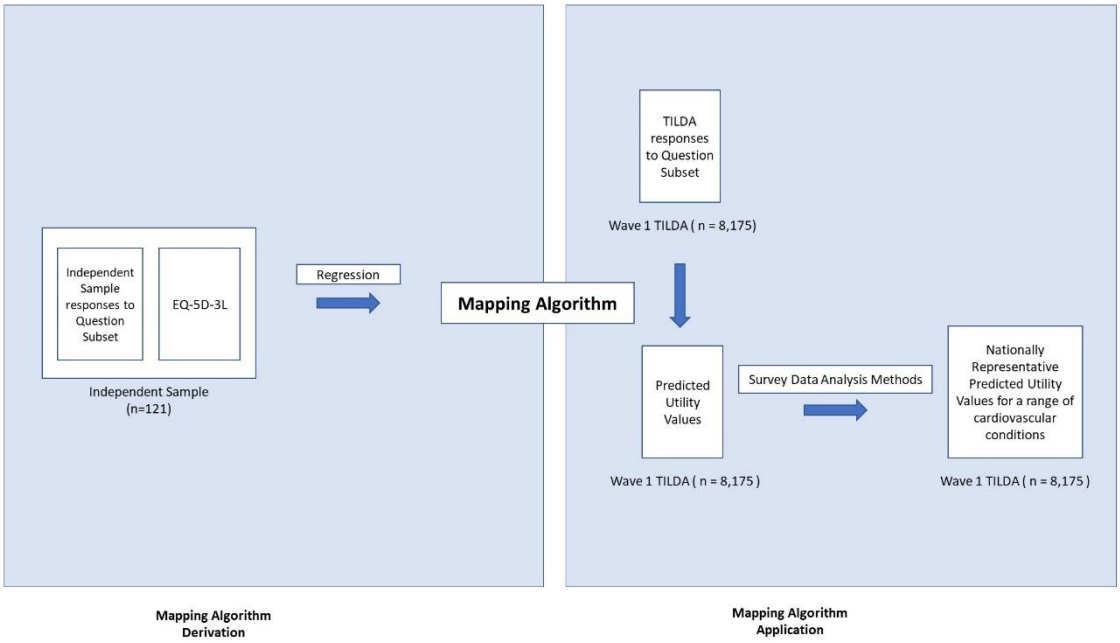


Figure 5.1 Overview of mapping study design
n, number; Question Subset, Subset of TILDA questions considered to have likely conceptual overlap; TILDA, The Irish Longitudinal Study on Aging.

5.3.2 Independent Sample

From 09/08/2017 to 20/04/2018, participants were recruited from cardiovascular outpatient clinics in a tertiary teaching hospital. Potentially eligible patients were identified prior to clinic through chart review. Participants were required to have a history of atherosclerotic CVD and be 50 years or older. The full eligibility criteria are outlined in Box 5.4. Eligible patients were invited to take part in the study and completed both the 'Question Subset' (via face-to-face interview) and the EQ-5D-3L while waiting for their appointment.

Box 5.4 Mapping study inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
•Patients must have been referred to a cardiovascular outpatient clinic •Patients must have a history of atheroscloerotic cardiovascular disease •Patients must be able to self report their health related quality of life. •Patients must be 50 years or older •Patients musts be able to give informed consent.	•Patients who are not in a position to self-report their quality of life •Patients attending the consultant's out-patient clinic for the first time.

5.3.3 Outcome measures

The order in which respondents received the EQ-5D-3L and 'Question Subset' was randomised (1:1). Prior to respondent recruitment, a random number generator in Microsoft Excel® was used to generate a random list of 1 and 2. This list was attached to a list ID numbers (1-200) which indicated the order the patient was to receive the EQ-5D-3L and the 'Question Subset'. Respondents were assigned consecutive ID numbers in the order they consented to take part in the study.

5.3.3.1 EQ-5D-3L

5.3.3.1.1 Description

The EQ-5D is composed of two parts, a descriptive system (used to describe health states) and a value set (which provides a valuation for each state). These are combined to form a health state preference weight (also known as a utility value). Two forms of the EQ-5D descriptive system are available – 5L and 3L. The 5L version of the EQ-5D has been shown to have greater discriminatory capacity and less ceiling effects than 3L (219). An Irish value set for this version was not available at the time this study was designed but has been published in the interim (220). The 3L version of the EQ-5D, which has three response levels (no problem, moderate problems, extreme problems) for each domain, was used for this study as it is recommended by Irish and international HTA agencies (183,184,221). An Irish value set is in development but has not yet been published (222). Therefore, responses were valued using the UK value set (223).

5.3.3.1.2 EQ-5D-3L distribution

The distribution of EQ-5D-3L utility is important for the choice of model. The data generally has several non-normal characteristics. Distributions are often skewed, multimodal, have large spikes, and gaps in the range of feasible values (224).

There is debate in the literature on whether utility is either censored or bounded at 1. Bounded or limited means that it is impossible for a utility to exceed 1 conceptually. In contrast, censored means that while it is possible for an individual's utility to exceed 1, values above 1 are not observed and the person is reported as having a utility of 1. Sullivan believes that utilities are censored (225). He states that while utilities are anchored at 0 (death) and 1 (full health), full health can be exceeded. He argues that the concept of full health is ill defined and will vary across descriptive systems. Hence, he concludes that the descriptive state that corresponds to 1 for full health will differ depending on the instrument used. He proposes that no matter how you define it, there will always be people who exceed full health. He gives the comparative example of people who can run marathons have more perfect health that someone who cannot but who has no problems walking about as measured by the EQ-5D-3L. To support his argument, he states that future descriptive systems should incorporate potential values that exceed full health as is currently done for death.

Sullivan's arguments in favour of censoring are overcome by Pullenayegum et al who believe that utilities are bounded (226). Pullenayegum et al state that utilities cannot exceed 1 because you cannot do better than full health, which by definition is valued at 1. They consider the presence of a large peak at 1 as a measurement issue rather than a censoring issue (226). Given the need to make questionnaires that are usable, they argue the measurements are approximate only and mean that valuations models are insensitive to small departures from perfect health which can lead to a large peak at perfect health. They highlight that in contrast, HRQoL is a separate and more abstract construct without a bound. This means that different instruments can be developed over many different scales and score boundaries for this construct (226).

Theoretically, censoring and bounding are also an issue for the lower end of utility distribution. But given the number of people reporting health is this range is typically low, it is of minimal concern (224).

Another issue relates to the classification of utility data as discrete or continuous. Discrete data can only take certain values. In contrast, continuous data can occupy any value over a specified range. But how data is measured can affect its properties. For example, age as defined by length

of time since birth, is continuous. But people typically round down their age to the nearest year when reporting it. Therefore, in this form it could be classified as discrete data. It can be argued that limitations in measurement accuracy mean that any measured data are, in practice, discrete (227). Therefore, the decision to treat data as continuous or discrete depends on the extent of measurement accuracy and the requirements of the analyst.

By definition, utility is continuous. But the measurement properties of the EQ-5D-3L mean that a large number of people have a full peak at 1 and that there can be large gaps in the range of feasible values. For example, in the EQ-5D-3L UK value set, there is a large gap between 0.883 and 1. But, there are also other gaps in the distribution. Some argue that the extent of discreteness in EQ-5D utilities means that any modelling of the utility distribution parametrically is likely to be extremely difficult (227). However, given the size of other gaps are much smaller, most analysts are happy to treat the rest of the distribution as continuous (224).

The implications of the distribution of EQ-5D utility for the choice of model are considered in Section 5.3.4.3 below.

5.3.3.2 Selection of the TILDA 'Question Subset'

Given the amount of data collected, it was not pragmatic to administer the full TILDA interview to the 'Independent Sample'. Therefore, the national TILDA questionnaire was reviewed, and questions considered likely to have high conceptual overlap with the EQ-5D-3L domains were selected for inclusion in the 'Question Subset'.

Patients in the 'Independent Sample' were also asked further questions from TILDA that relate to sociodemographics and history of CVD. To maximise comparability with TILDA, the 'Question Subset' was administered in person, in the same order as the original questionnaire. Phrasing of questions in the subset were identical to those posed in TILDA, except for two minor modifications for medical history. PAD was added to a list of cardiovascular conditions and a question on non-cardiovascular medical history was simplified. The full 'Question Subset' administered is reproduced in the supplementary Appendix 1.2.1.

The only quality of life instrument in TILDA was CASP-19 which is based on a model that conceptualises quality of life as needs satisfaction (228). Domains include control, autonomy, self-realisation and pleasure. These appear to relate to different constructs than those assessed in HRQoL instruments such as the EQ-5D-3L. Therefore, attempting mapping between them is unlikely to be successful. Therefore, it was not included in the 'Question Subset'. Examples of questions included in the 'Question Subset' for each EQ-5D domain are described below.

5.3.3.2.1 Mobility

The national TILDA questionnaire included multiple questions on mobility. Examples include difficulties running, walking 100 metres, climbing stairs and walking across a room. Participants who reported difficulties were also asked about appliances required such as wheelchairs and walking sticks.

5.3.3.2.2 Self-care

TILDA asked multiple questions on activities of daily living, including reported level of difficulty for both washing and dressing. Participants were also asked about appliances required and if they needed help from carers for these activities.

5.3.3.2.3 Usual activities

Participants who reported a long-term health problem were asked if their activities were limited due to illness. Participants who responded 'yes' were asked further questions regarding the severity of their limitations.

5.3.3.2.4 Pain/discomfort

Participants were asked if they often suffered from pain. Those who responded 'yes', were asked if the pain was mild, moderate or severe.

5.3.3.2.5 Anxiety/depression

TILDA included screening tools for both anxiety and depression. The Centre for Epidemiological Studies Depression Scale (CES-D) (229) was administered in the interview section in TILDA while the Hospital Anxiety Depression Scale - Anxiety subscale (HADS-A) was included in the self-completion component of the evaluation (230). However, a decision was made not to include these measures in our study as it would substantially increase the length of time taken to complete the face-to-face interviews (and thus potentially impact on patient participation). In addition, ethical issues would arise regarding patients who may require clinical intervention having scored highly on screening tools for anxiety and depression. Instead, we included a TILDA question which asks participants to rate their emotional or mental health. In the TILDA population this outcome was correlated to the CES-D score, providing some support for this approach.

5.3.3.3 Important differences between outcome measures

While it appears that there are strong similarities in the constructs measured by the 'Question Subset' and the EQ-5D-3L, it is also important to note important differences between them.

5.3.3.3.1 Recollection period

Across all domains, the EQ-5D asks respondents to describe their own health state today. Whereas, the recollection period for corresponding questions in TILDA differ by domain. For example, questions relating to activities of daily living ask participants to refer to their abilities over the last 6 months. A further difference is that participants are asked to exclude the effect of impairments which are expected to last less than three months. No specific time period is specified for questions that pertain to pain in TILDA. However, the participant is required to give a more general consideration of pain, rather than considering only their pain today. That is, the participant is asked "*Are you often troubled with pain*?" and "*How bad is it most of the time?*". The differing time-recollection periods may affect the extent of overlap between measures.

5.3.3.3.2 Description of domains

While there is high conceptual overlap between domains, there is no perfect overlap. For example, the EQ-5D-3L asks respondents to rank their level of pain or discomfort. Comparable questions from TILDA do not refer to the discomfort aspect.

Given the length of the survey, TILDA was designed to minimise respondent burden. Many routing questions were asked, where participants were only asked certain sets of questions depending on their response to the previous question. In the TILDA questionnaire, participants were asked "*Do you have any long-term health problems, illness, disability or infirmity?*". Participants were only asked about how their activities were limited due to illness if they answered 'Yes' to this question. Many patients in the 'Independent Sample' stated that they had no long-term health problem despite having a confirmed diagnosis of ischemic heart disease. Therefore, these patients were not asked any follow up questions regarding limits to their activities. A comparison of these patients' responses to the TILDA 'Question Subset' and the EQ-5D indicates that some of these patients did indeed have limits to their daily activities.

5.3.3.3.3 Mode of administration

As in TILDA, questions were administered to the 'Independent Sample' by an interviewer who recorded the patient's responses. In contrast, the EQ-5D-3L questionnaire was completed by patients themselves. Responses are known to differ by mode of administration (231).

5.3.4 Statistics and econometrics

5.3.4.1 Correlation

Conceptual overlap between the source ('Question Subset') and the target variables (EQ-5D-3L) was assessed using Spearman rank correlation tests. Interpretation of the strength of the coefficients was based on category thresholds adopted by Davison et al (232). These were

defined as very weak (0-0.19), weak (0.20-0.39), moderate (0.40-0.59), strong (0.60-0.79) and very strong (0.80-1). To facilitate statistical analysis, participant responses to the 'Question Subset' were rearranged into sets of ordinal variables corresponding to each EQ-5D-3L domain.

5.3.4.2 Choice of explanatory variables

Variables corresponding to each EQ-5D-3L domain were included as explanatory variables for each model. For some domains, there was a choice of more than one variable from the 'Question Subset.' In these cases, the variable with the highest Spearman correlation coefficient for the relevant domain was included. Where there were marginal differences in the coefficients, face validity was considered. Where appropriate, variables were collapsed into two or three levels.

5.3.4.3 Potential model choices

The choice of model should be informed by the characteristics of the EQ-5D-3L utility distributions described in Section 5.3.3.1.2 above. Heteroskedasticity of residuals can lead to biased standard errors. Approaches to address this include bootstrapping (226) or the use of models whose estimation is not affected by heteroskedasticity such as CLAD (233–235). In this study we used robust standard errors using a sandwich estimator for all mapping models, as recommended by Pullenayegum et al (226).

Statistical analysis was conducted in Stata version 15.1 (StataCorp, College Station, Texas, USA) and R version 3.5.1 (R Core Team, Vienna, Austria).

5.3.4.3.1 Ordinary least squares

OLS is the most common model used in mapping. The model for utility for a given individual i (y_i) may be written as:

$$y_i = x_i'\beta + \varepsilon_i \quad \text{where } \varepsilon_i \sim N(0, \sigma^2) \tag{1}$$

where x_i represents a vector of explanatory variables, β represents the corresponding beta coefficients and ε_i represents the error term. Errors (ε_i) are assumed to be normally distributed, with mean 0 and variance σ^2.

This model has several limitations (199). It does not account for the bounded nature of the utility distribution. Therefore, it is possible to generate utilities outside the utility range. Further, it does not account for multimodality or the discrete nature of the data (232). The advantage of this method is that the estimated beta coefficients ($\hat{\beta}$) are, given the model assumptions, the best linear unbiased estimator (BLUE) (227).

Pullenayegum et al state that while the linear model does not fit utility well, given the complexity of the utility distribution, no parametric model will ever fit the data well (227). But if one is only interested in the conditional mean of the distribution as a function of co-variates, parametric modelling is unnecessary. They advocate the use of semi-parametric models in this case where

$$E(Y|X\ ) = X\beta. \tag{2}$$

In this case, we do not assume that the residuals are normally distributed. This means that $\hat{\beta}$ is no longer guaranteed to be unbiased or normal. However, under the assumptions of the central limit theorem, $\hat{\beta}$ will still be unbiased and asymptotically normal provided the sample size is sufficiently large. In simulation studies, Pullenayegum et al showed that confidence intervals based on robust standard errors were acceptable for sample sizes of 100 and larger (236). The estimation procedure for equations (1) and (2) are identical and use the ordinary least squares method.

For our research questions, we are not interested in the conditional mean on covariates but are concerned with modelling the whole of the distribution. Therefore, modelling using OLS is unlikely to correctly specify our model.

5.3.4.3.2 Tobit based models

5.3.4.3.2.1 Tobit model

The Tobit model accounts for the fact that it is impossible for the dependent variable to exceed 1 as values are censored or limited. A latent variable y_i^* is modelled such that:

$$y_i = \min\{y_i^*, 1\} \tag{3a}$$

$$y_i^* = x_i'\beta + \varepsilon_i \tag{3b}$$

In some contexts, the latent variable $y_i^*$is given some meaning, but here no interpretation of the latent variable is required (224).

5.3.4.3.2.2 Adjusted limited dependent variable model

The adjusted limited dependent variable (ALDV) model is a modification of the Tobit model which takes account of the gap between full and the next feasible value (224). For example, in the EQ-5D-3L UK value set it is impossible to directly estimate a utility value between 0.883 and 1. It is modelled by:

$$y_i = \begin{cases} 1 \text{ if } y_i^* > \tau \\ y_i^* \text{ otherwise} \\ y_i^* = x_i'\beta + \varepsilon_i \end{cases} \quad (4)$$

Where τ is next feasible value below 1. For the UK EQ-5D-3L value set used in this case τ = 0.883.

5.3.4.3.2.3 Censored least absolute deviation models

CLAD is variation of the Tobit model. Like the Tobit, it assumes that utility values greater than 1 are censored. It assumes that the median is a linear combination of the co-variates rather than the mean. The remainder of the distribution is left unspecified (226).

5.3.4.3.2.4 Discussion of Tobit based models

The performance of OLS, Tobit and CLAD models has been compared by Pullenayegum et al (226). In a simulation study, they showed that when the true utility is bounded at 1, the Tobit and CLAD estimators were both biased. This was in contradiction to conclusions on the performance of Tobit models by Austin et al (237). Pullenayegum et al note that Austin et al simulated the utility data to exceed 1 and then censored; this explains the divergence of results (226). Therefore, the question on whether Tobit and CLAD models are appropriate for estimating conditional means depend on the analysts' concept of utility and whether or not they believe it can exceed 1. In any case, the CLAD model is a poor modelling choice as the median is modelled when generally the mean is of primary concern (226).

For our purposes, we are interested in the accuracy of the models in predicting individual values rather than estimating the mean conditional values. Hernandez-Alava et al considered the Tobit and ALDV model as artificially censoring utility (224). While they considered utility to be conceptually bounded at 1, they exploited the peak at 1 generated by assuming censoring to more accurately describe the utility distribution. While Hernandez-Alava's et al's introduction to the ALDV model primarily focused on the performance of an ALDV mixture model, one component linear, Tobit and ALDV models were also compared (224). The RMSE and MAE were broadly comparable across all three models although the MAE was marginally lower for the linear model. Theoretically, the ALDV model is more appropriate than the Tobit model because of the closer alignment to the EQ-5D-3L utility distribution through the introduction of a gap between full health and the next possible utility value.

5.3.4.3.3 Beta based models

5.3.4.3.3.1 Beta model

The beta model accounts for the bounded nature of the EQ-5D-3L at both ends of the distribution. Its properties are well documented (238). In beta models, the dependent variable

is defined only in the open interval (0,1). Therefore, we transformed the dependent variable which has range of -0.596 to 1 into the [0,1] interval using a logit transformation defined as

$$y \quad = \frac{y-a}{b-a} \tag{5}$$

for the interval $[a, b]$, where a is the minimum value (-0.596) and b is the maximum value (1) of the distribution. As a value of 1 is not supported by the beta distribution, a value of -1x e^{-6} is added to those observations (239)[7].

The probability density function h with mean μ and precision parameter ∅ may be represented as

$$h(y_i|x_i) = f(y_i|x_i, \beta, \emptyset, a, b) \tag{6}$$

and $x'$is a vector of coefficients (238).

After estimation, the expected value is transformed back into its original scale to obtain the predicted value.

5.3.4.3.3.2 Inflated beta model

An inflated beta model may also be estimated. Again, the dependent variable is transformed to the open interval (0,1). The model is estimated on the transformed values and has two parts. First, a logit model is used to model a mass of observations at 1 (full health)[8]. Then a beta model is estimated for the remaining portion of the distribution. Due to sample size limitations, we estimated the logit model with only a constant probability mass at full health, which limits its flexibility. The probability density function may be represented as

$$g(y_i|x_i) = \begin{cases} P(y_i = 1), y_i = 1 \\ [1 - \ P(y_i = 1)]h(y_i|x_i), y_i \ \in (a\ , b) \end{cases} \tag{7}$$

where$h\big(y_i\big|x_i \ \big) = f(y_i|x_i, \beta, \emptyset, a, b)$ is defined in line with standard beta model (equation 6)(238).

[7] The transformation of observed utility to the 0 to 1 closed interval would transform the minimum value of the EQ-5D-3L UK dataset from -0.596 to 0 in the beta distribution. There were no minimum values in our dataset. However, if there were, the '*betamix'* command would add 1 x e^{-6} to this observation prior to estimation as (like 1), 0 is on the boundary of the beta distribution and is not supported.

[8] Because a logit model is used to estimate the mass of values of 1, this value is accounted for in the model and there is no need adjust these values.

5.3.4.3.3.3 Truncated inflated beta model

We also estimated a truncated inflated beta model which extends the inflated beta model to take account of the gap between 1 and the next feasible value (τ). For the EQ-5D-3L UK value set τ = 0.883. The probability density function may be represented as

$$g(y_i|x_i) = \begin{cases} P(y_i = 1), y_i = 1 \\ [1 - P(y_i = 1)]h(y_i|x_i)\ y_i \in (a\ , \tau\ ) \end{cases} \quad (8)$$

where $h(y_i|x_i) = f(y_i|x_i, \beta, \emptyset, a, \tau)$ is defined in line with standard beta model (equation 6) (238). As for the inflated beta model, only a constant probability mass at full health was estimated which limits flexibility.

5.3.4.3.3.4 Discussion of beta-based models

Conceptually, beta-based models are an attractive option for modelling utility as they are bounded at both ends. Modifications described by Gray et al further account for the unique features of the EQ-5D-3L distribution including the probability mass at 1 and the gap (238).

5.3.4.3.4 Mixture models

Mixture models are a flexible semi-parametric method that can account for the multi-modality often observed in utility distributions and a wide variety of unusual distributions. Multi-modality can be a result of unobserved heterogeneity in the participant sample. Each group can be considered under C latent classes (224). Mixture models have been used to estimate a number of model types including normal, beta and ALDV.

For example, in an ALDV mixture model, as proposed by Hernandez Alava et al, a multinomial logit model for the probability of latent class membership can be assumed:

$$P(C_i = c|\omega_i) = \frac{\exp\ (\omega'_i \partial_c)}{\sum_{S=1}^{P} (\omega'_i \partial_S)} \quad (9)$$

where ω_i is a vector of variables that affect the probability of class membership, ∂_c is the vector of corresponding coefficients and P is the number of classes. A number of ALDV models are estimated (n=C) conditional on an observation belonging to each latent class (224).

Judgement is required when determining the number of latent classes as fitting mixture models is complex (224). A mixture model of n-classes where n equals the sample size would be a fully non-parametric model. There is a risk of over-extracting classes in the presence of outliers or non-normality. Statistical issues encountered in model estimation include the presence of several local maxima in the likelihood function and unbounded likelihood functions. Methods of dealing with these difficulties have been described (240).

Mixture models are an attractive method of accounting for the multimodal nature of the utility distribution. As they can be used with many model types including beta inflated truncated and ALDV models, they are also very flexible. There are many papers in the literature providing empirical evidence supporting the validity and flexibility of mixture models (224,238,241,242).

5.3.4.3.5 Response models

Indirect or "response" mapping was also considered. In this method, mapping is performed in two parts. First using regression, responses to the target measure (i.e. levels of the EQ-5D-3L descriptive system) are predicted. Five separate models are estimated corresponding to each domain of the EQ-5D-3L. Utility values are generated by applying the respective value set to the probabilities of each predicted response. Therefore, an advantage of the method is that once a model is estimated, numerous value sets can be applied without having to re-estimate the mapping model.

As for direct mapping, several model types can be used including the multinomial logit and standard ordered probit or logit. Hernandez-Alava et al highlight that the standard ordered models are not flexible enough, as they assume the same coefficients for the explanatory variable across the different categories (parallel line assumption) which led researchers to use the multinomial logit model instead (243). While this approach, relaxes the parallel line assumption, the ordinal nature of the dependent variable is ignored. To overcome this issue, Hernandez Alava et al proposed the use of generalised ordered probit model which accounts for both the ordinal nature of the dependent variable and relaxation of the parallel line assumption (243). However, the ALDV mixture model has been shown to perform better than this model. When both models were compared in a large database of respondents with rheumatoid arthritis (N=100,398), there was limited data from respondents in the most severe health states meaning there was still relatively few data to estimate the generalised ordered probit model limiting its performance (243).

Further work by Hernandez-Alava et al developed mapping models (between the EQ-5D-3L and EQ-5D-5L) which accounted for the background correlation between the responses in the five health domains while still allowing the nature of the relationship between EQ-5D-3L and EQ-5D-5L responses to be quite different in each health domain (244). Five separate bivariate copulas were estimated for each domain and connected via a latent factor to represent common influences on the participant's responses. Previous work using the generalised ordered probit models estimated five unrelated models for each of the five health domains (243).

Response models are an attractive choice for modelling given that the difficulties in modelling the EQ-5D-3L distribution parametrically are avoided. However, they require a very large amount of data and, as described above, they present their own estimation difficulties.

5.3.4.3.6 Models chosen for this study

Six models were estimated for this analysis including OLS, Tobit, ALDV, beta, inflated beta and truncated inflated beta.

OLS is not a suitable parametric model for modelling the entire EQ-5D-3L distribution. But it was included given its ubiquity in the modelling literature and to facilitate comparison versus other model types. CLAD was not included given that it models median when we are interested in the mean values. Response models were not estimated, as it requires participant responses in each response category for each domain of the EQ-5D-3L. In our sample, there were no respondents with level 3 values for the EQ-5D for the mobility or the anxiety/depression domains. This was not unexpected. In a large UK study of 7,998 respondents with CVD, only 0.3% and 2.8% stated that they had extreme problems with these domains respectively (207). Mixture models were not estimated as the sample size available was insufficient to estimate them reliably.

OLS models were fit using the '*reg*' command in Stata. Tobit and ALDV models were fitted using the '*aldvmm*' command in Stata® (240). Beta based models were fitted using the '*betamix*' command in Stata® setting the number of components to one as mixture models were not computed (239).

5.3.5 Validation

5.3.5.1 Goodness of fit

To compare model goodness-of-fit, RMSE and MAE were calculated for the full 'Independent Sample' and for subgroups (based on observed EQ-5D-3L utility) to examine model performance at different levels of health. For models where utility was transformed prior to estimation, all errors were calculated after transforming predicted values back on the original scale. Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) are statistics commonly used to assess relative goodness-of-fit. However, they are not computed here as the transformation used in some models mean these statistics cannot be compared across them.

5.3.5.2 K-fold cross validation

Validation of the model is important to confirm the reliability of the results. K-fold cross-validation was employed using the Stata command '*crossfold*'. Under this method, the estimation dataset is randomly partitioned into 'K' parts or 'folds'. K-1 folds are used for

estimation and the remaining fold is employed for validation. This is repeated 'K' times and measures of model performance are averaged across the 'K' models. There is no established method for choosing the number of folds, but values of five or ten are commonly used. In this study, fivefold (K=5) cross-validation was chosen to compare the predictions of each mapping model with actual utilities. RMSE and MAE were calculated for each model. An advantage of this method is that it avoids sample splitting and thus allows all observations in the dataset to be used for estimation and validation (245).

5.3.6 Selection of the chosen model

A rule-based approach to model selection based on overall fit is discouraged (200), therefore detailed criteria were not pre-specified. The chosen model was selected based on a combination of empirical evidence supporting their validity in the mapping literature, model fit, validation results and plausibility (200).

5.4 Results

5.4.1 Independent Sample characteristics

Complete case analysis of the 'Independent Sample' was conducted (n=121). A patient flow-diagram is presented Figure 5.2. Table 5.2 summarises respondents' characteristics. Age ranged from 50 to 88 years (mean of 69.3 years); 77.7% were male. Summaries of responses to each outcome measure are shown in the Table A.6 and Table A.7 in the Supplementary Appendix. Plots of response categories are shown in Figure 5.4. Distribution of the utilities was left skewed with a mass point of 49 (40.5%) at the upper limit of 1 (Figure 5.3).

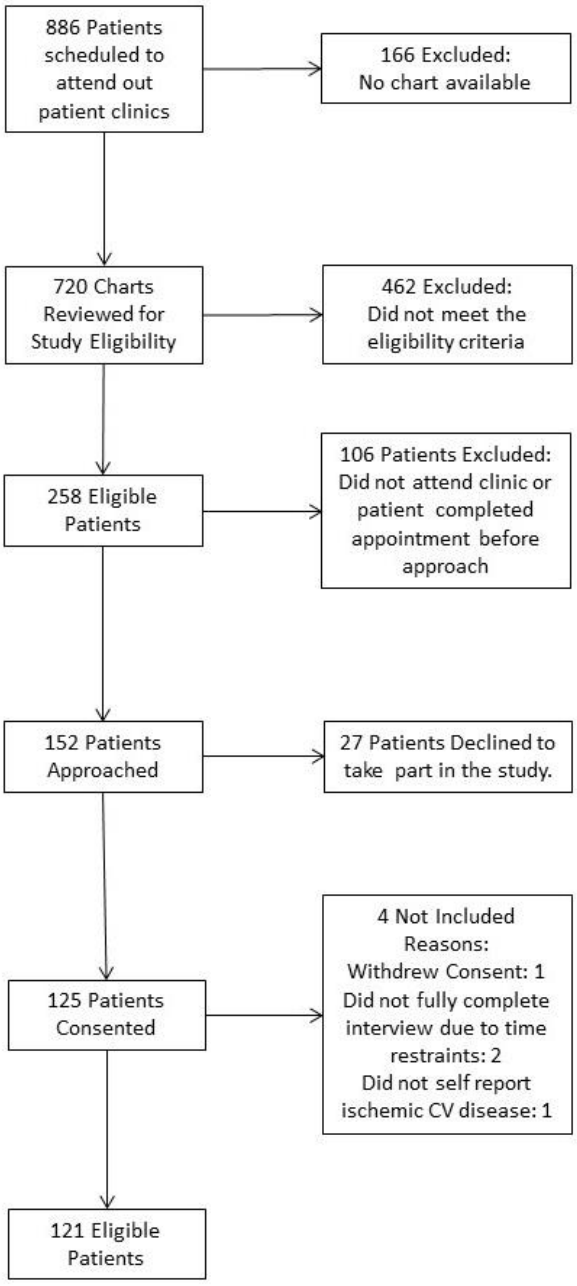


Figure 5.2 'Independent Sample' participant flow diagram

Table 5.2 'Independent Sample' characteristics

Characteristic	Value
Sample Size	121
Age (years), mean ±SD	69.3 ± 9.6
Male, n (%)	94 (77.7)
Angina, n (%)	44 (36)
Myocardial Infarction, n (%)	65 (54)
Congestive Heart Failure, n (%)	12 (10)
Stroke, n (%)	7 (6)
Mini-Stroke or Transient Ischemic Attack, n (%)	19 (16)
Peripheral Arterial Disease (PAD), n (%)	13(11)
Other Heart Trouble, n (%)	46 (38)
EQ-5D-3L Utility, mean ±SD	0.78±0.26
EQ-5D-3L Utility, Interquartile Range	0.66-1

n, Number; SD, Standard Deviation.

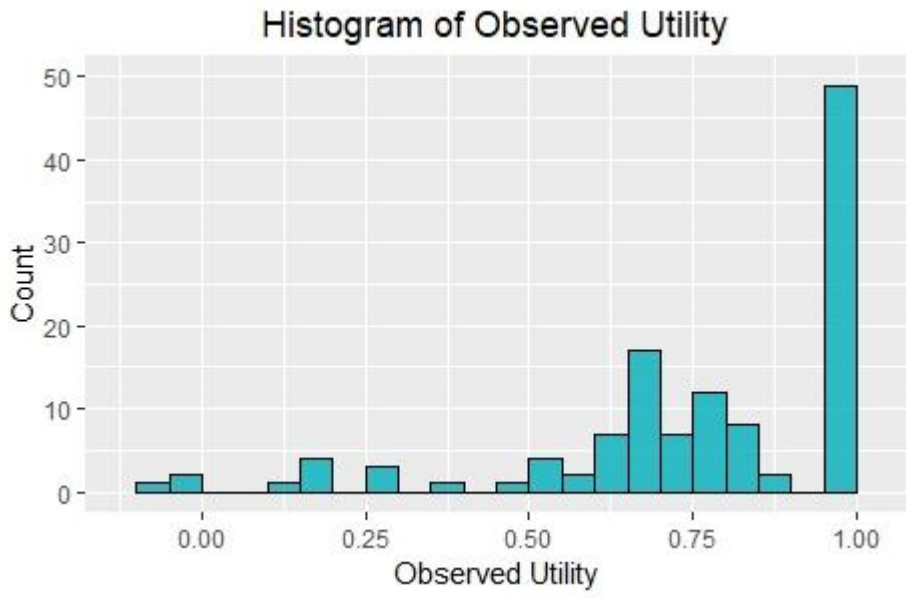


Figure 5.3 Distribution of the EQ-5D-3L in the Independent Sample

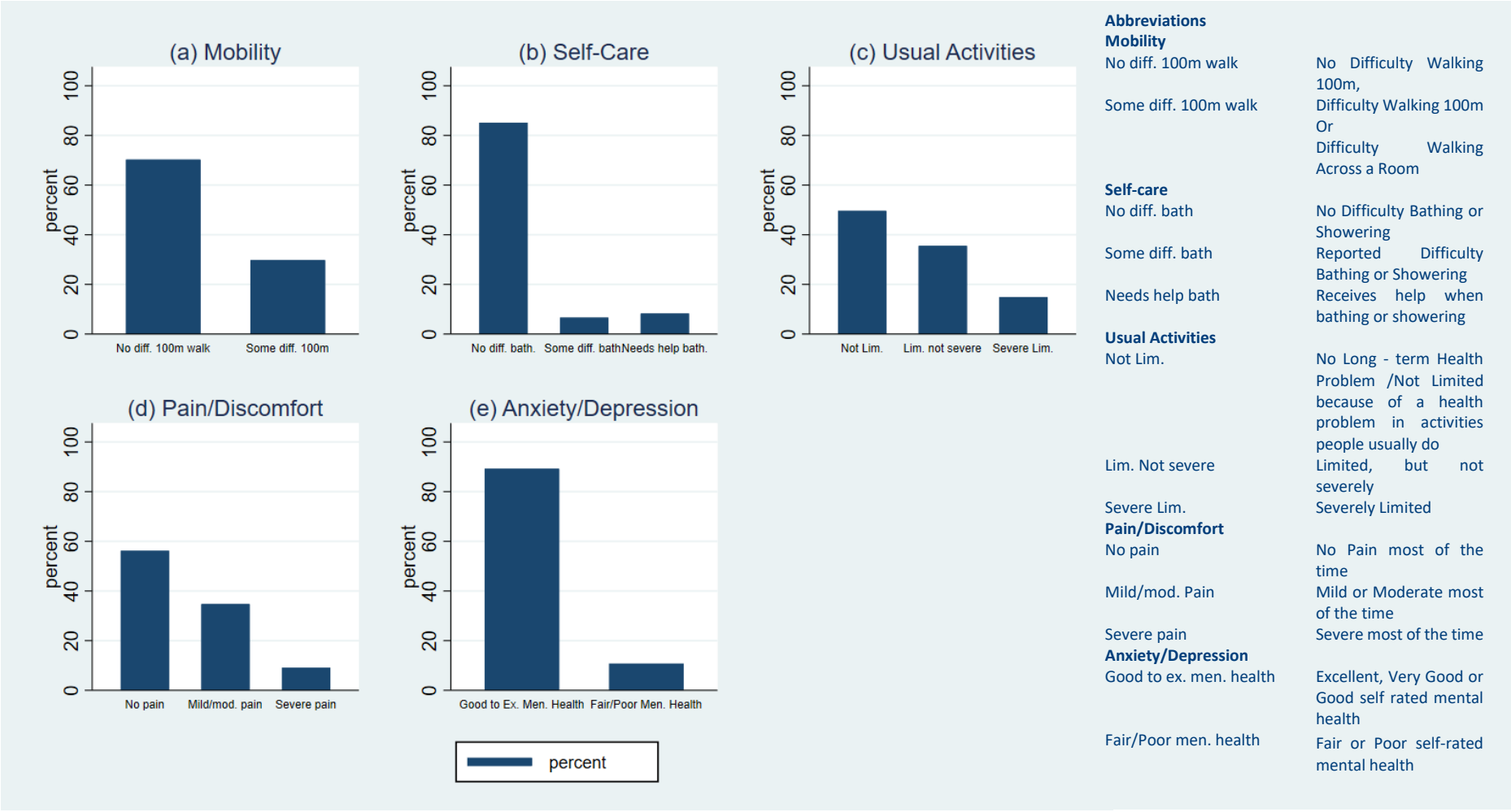


Figure 5.4 Independent Sample responses to the Question Subset variables included in the mapping model.

5.4.2 Correlation

Spearman rank correlation coefficients for the 'Question Subset' variables included in the regression model are presented in Table 5.3. Coefficients for all variables considered are presented in the Supplementary Appendix Table A.8. The results showed moderate correlation between the anxiety/depression domain and the 'Question Subset' and strong or very strong correlation for remaining domains. These results suggest that a mapping function would have sufficient conceptual overlap to be successful.

Table 5.3 Spearman correlation coefficients between EQ-5D-3L domains and comparable TILDA variables included in the mapping model

EQ-5D-3L		Comparable TILDA variables included in the Mapping Model	Spearman Correlation Coefficient
Mobility			
1.	I have no problems in walking about	No Difficulty Walking 100m	0.67***
2.	I have some problems in walking about	Difficulty Walking 100m Or Difficulty Walking Across a Room	
3.	I am confined to bed		
Self-Care			
1.	I have no problems with self-care	No Difficulty Bathing or Showering	0.80***
2.	I have some problems washing or dressing myself	Reported Difficulty Bathing or Showering	
3.	I am unable to wash or dress myself	Receives help when bathing or showering	
Usual Activities			
1.	I have no problems with performing my usual activities	No Long-term Health Problem /Not Limited because of a health problem in activities people usually do	0.62***
2.	I have some problems with performing my usual activities	Limited, but not severely	
3.	I am unable to perform my usual activities	Severely Limited	
Pain/Discomfort			
1.	I have no pain or discomfort	No Pain most of the time	0.67***
2.	I have moderate pain or discomfort	Mild or Moderate most of the time	
3.	I have extreme pain or discomfort	Severe most of the time	
Anxiety/Depression			
		How would you rate your emotional or mental health?	
1.	I am not anxious or depressed	Excellent, Very Good or Good	0.46***
2.	I am moderately anxious or depressed	Fair or Poor	
3.	I am extremely anxious or depressed		

*p<0.05; **p<0.01, ***p<0.001
Very weak (0-0.19), weak (0.20-0.39), moderate (0.40-0.59), strong (0.60-0.79) and very strong (0.80-1).

5.4.3 Estimation

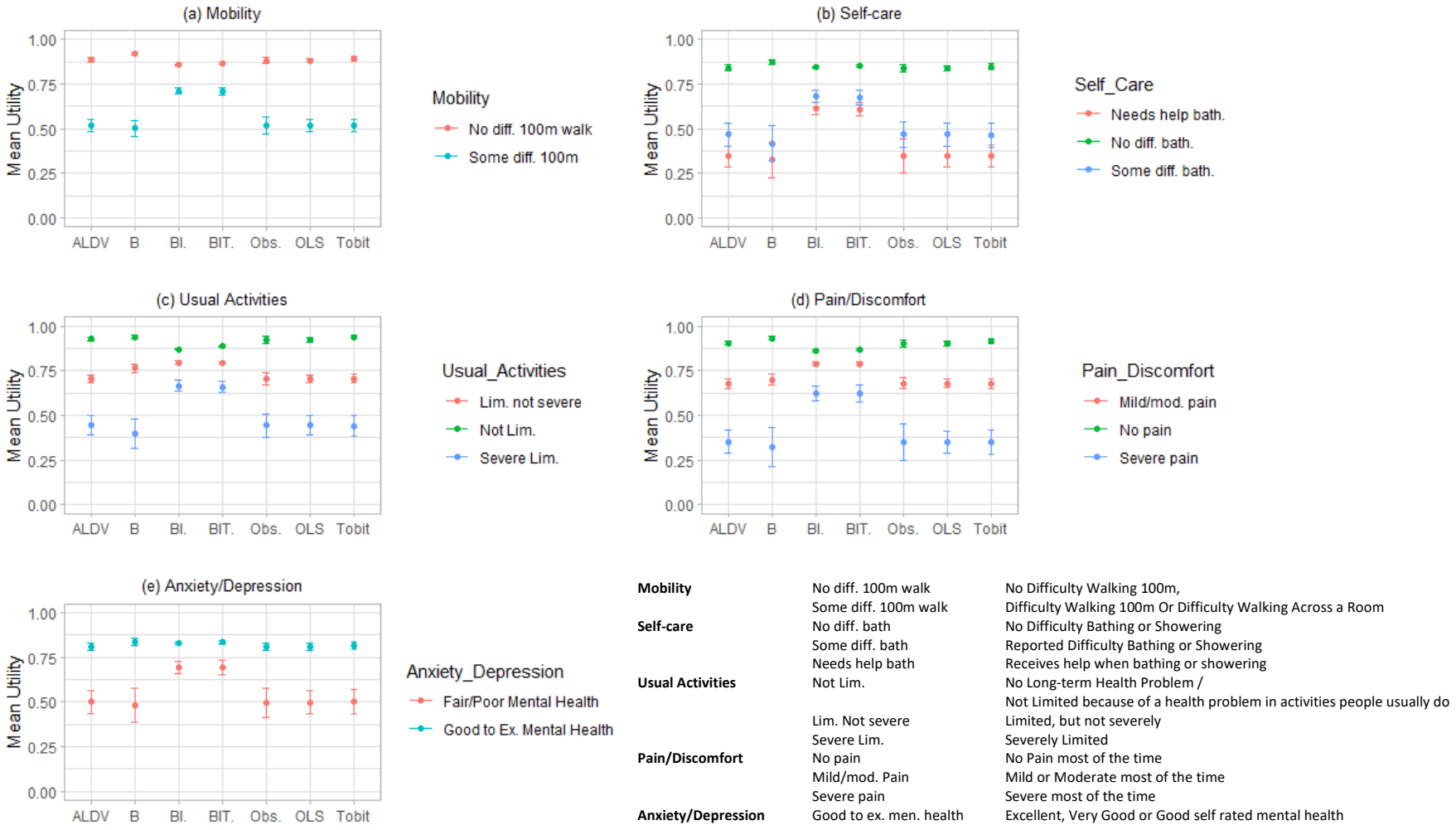
The same independent variables were included in each model. The estimated coefficients for the six models are presented in the Supplementary Appendix Table A.9. As expected, all coefficients in each model were negative and increasing severity levels predicted increasing utility decrements. In exploratory analysis, a dummy coefficient that indicated whether the EQ-5D was administered first or second was added to the regression. As it was not significant, it was removed from the final model.

5.4.4 Prediction

The mean predicted and observed utility values stratified by the independent variables are presented in Figure 5.5 for each of the six mapping models estimated. Visual inspection for OLS, Tobit, ALDV and beta models indicates that the mapping model produced predicted mean values consistent with the observed mean values. Fewer observations in the most severe health states increased the standard error for these analyses. The inflated beta and truncated inflated beta models were a poor fit to the data because of the constant probability mass at 1 and resulted in poorer estimated values.

Goodness-of-fit results are presented in Table 5.4. No model performed best across all measures. The OLS model had the lowest MAE and RMSE. In line with the results of the visual inspection, the inflated beta and truncated inflated beta -models tended to perform less well across all measures.

Error estimates (RMSE and MAE) across observed EQ-5D-3L subgroups are presented in the Supplementary Appendix Table A.10. The number of observations at lower EQ-5D-3L ranges was low. The beta model performed best at the lowest EQ-5D-3L range (observed EQ-5D-3L <0.33), while the OLS model performed best for the remaining groups (observed EQ-5D-3L 0.33< 0.66 and 0.66 ≤ 1).



Mobility	No diff. 100m walk	No Difficulty Walking 100m,
	Some diff. 100m walk	Difficulty Walking 100m Or Difficulty Walking Across a Room
Self-care	No diff. bath	No Difficulty Bathing or Showering
	Some diff. bath	Reported Difficulty Bathing or Showering
	Needs help bath	Receives help when bathing or showering
Usual Activities	Not Lim.	No Long-term Health Problem / Not Limited because of a health problem in activities people usually do
	Lim. Not severe	Limited, but not severely
	Severe Lim.	Severely Limited
Pain/Discomfort	No pain	No Pain most of the time
	Mild/mod. Pain	Mild or Moderate most of the time
	Severe pain	Severe most of the time
Anxiety/Depression	Good to ex. men. health	Excellent, Very Good or Good self rated mental health
	Fair/Poor men. health	Fair or Poor self-rated mental health

Figure 5.5 Mean observed and predicted utility values (± standard error) for the six mapping models estimated across independent variables.
ALDV, Adjusted Limited Dependent Variable; B, Beta; BI, Beta Inflated; BIT, Beta Inflated Truncated; Observed, OLS, Ordinary Least Squares

Table 5.4 Predicted model characteristics and performance

	Model Characteristics						K-Fold Cross Validation	
Model Name	**Mean**	**SD**	**Min**	**Max**	**MAE**	**RMSE**	**MAE**	**RMSE**
Observed	0.7749	0.2570	-0.0740	1.0000				
OLS	0.7749	0.2208	0.0263	0.9646	0.1014	0.1311	0.1110	0.1467
Tobit	0.7822	0.2288	0.0108	0.9699	0.1034	0.1344	0.1120	0.1481
ALDV	0.7763	0.2220	0.0261	0.9600	0.1046	0.1326	0.1136	0.1507
Beta	0.7959	0.2506	-0.1902	0.9645	0.1102	0.1483	0.1246	0.1812
Beta Inflated	0.8136	0.0949	0.4157	0.8810	0.1452	0.1872	0.1472	0.1902
Beta Inflated Truncated	0.8195	0.1022	0.3931	0.8939	0.1426	0.1848	0.1501	0.1945

ALDV, Adjusted Limited Dependent Variable Model; OLS, Ordinary Least Squares; MAE, Mean Absolute Error; Max, Maximum; Min, Minimum; RMSE, Root Mean Squared Error.

5.4.5 Validation

Results of the K-fold cross validation are presented in Table 5.4. The OLS model had the lowest RMSE and MAE errors (0.1467 and 0.111 respectively). Tobit, ALDV and beta models produced slightly higher errors. Inflated-beta and truncated-inflated-beta generated the highest errors.

5.4.6 Selection of the best fitting model

Selection of the best performing model requires judgement as no model dominated goodness-of-fit results. The OLS model performed best across most goodness-of-fit characteristics. However, it fails to account for the bounded nature of the EQ-5D-3L distribution. While it did not produce predictions outside the feasible model range in our sample, it may occur when coefficients are varied in a probabilistic sensitivity analysis (200). Therefore, it was excluded. ALDV and beta models have empirical evidence to support model performance. While the Tobit model had lower prediction errors than the ALDV model, the ALDV is designed to account for the gap in values just below full health unique to the EQ-5D distribution. For these reasons and because the ALDV model is bounded at 1, it was chosen as the preferred model. A variance co-variance matrix for the ALDV model is presented in the Supplementary Appendix Table A.11 so that uncertainty in the mapping model can be incorporated in probabilistic sensitivity analysis where the mapping model is used to derive utilities.

5.5 Discussion

5.5.1 Main findings

While directly observed utilities are preferable for economic evaluations, they are not always available. This study has shown that provided there is strong correlation between sets of variables, mapping questions (relating to HRQoL from a national survey) to the EQ-5D-3L is a feasible method of generating utilities when a preference-based outcome measure was not included in the original survey. As previously mentioned, Ghatnekar et al's mapping (health outcome measures to the EQ-5D-3L) was speculative given that both measures were included in the registry (206). Our study instead illustrates how EQ 5D values were generated from a dataset where the preference-based measure was not collected. To our knowledge, this is the first time that a mapping study of this kind has been performed. Here we have applied this approach to address this data gap in a national survey.

5.5.2 Limitations

Our study has several limitations. Our sample size for generation of the mapping model was modest. This increases uncertainty but the model still performed well given the strong correlation between measures. Mapping guidelines do not give minimum sample size requirements for mapping studies (200). Mapped utilities have more uncertainty that directly observed values. Variance co-variance matrices presented in the Supplementary Appendix should be used to quantify this uncertainty through probabilistic sensitivity analysis where these utilities are used in an economic model.

Mapping models may be sensitive to the population from which they are derived (246). Given the strong correlation and overlap observed between the 'Question Subset', and the EQ-5D-3L, the bias may be low if the relationship between the 'Question Subset' and the EQ-5D-3L is independent of baseline health. The variables derived from the 'Question Subset' are not from validated outcome measures. However, TILDA was established through an extensive design process which referenced many international longitudinal aging studies (195,247–249).

Bias may also be introduced through our choice of mapping model. More complex models such as mixture models offer increased flexibility (224,250). However, the modest size of our 'Independent Sample' was insufficient to consider these. It also limited our ability to include co-variates to model the peak at full health and to exploit the full flexibility of inflated and truncated beta models. Across all goodness-of-fit measures, no single model dominated model choice increasing uncertainty. Goodness of fit statistics and K-fold validation results were reasonable.

There is a tension between picking the model that best fits the theoretical features of the utility distribution and the model which gives the best predictions as measured by K-fold validation. Theoretically, the beta model is attractive as it is bounded at both ends. But it provided a relatively poor fit to the data. The modest size of our 'Independent Sample' limited our ability to include co-variates to model the peak at full health to exploit the full flexibility of beta inflated and truncated models which would capture the utility distribution more accurately. The ALDV model was chosen as the preferred model. As we are not using the beta coefficients directly in an economic model we are less concerned about bias in these parameters. We exploit the fact that it censors the data at 1 to produce a peak at full health which means that individual values can be more accurately predicted.

As seen elsewhere, models often over predicted at lower levels of health and under predicted at higher levels. There were few observations at lower levels of health. There were no level three responses for anxiety/depression or mobility EQ-5D-3L domains in our 'Independent Sample'. Other studies have also observed low frequencies of this response (246). These limitations increase the risk of bias. The model has not been externally validated which increases uncertainty.

5.5.3 Potential for future research

5.5.3.1 Future application

Mapped utilities have more uncertainty that directly observed values. Variance co-variance matrices presented in the Supplementary Appendix should be used to quantify this uncertainty through probabilistic sensitivity analysis where these utilities are used in an economic model.

Anonymised versions of TILDA dataset are publicly available to researchers on request from the Irish Social Science Data Archive. It is envisioned that researchers will apply the mapping model to the TILDA dataset and generate predicted utilities values for each participant using the regression coefficients included in the Supplementary Appendix. Researchers can then conduct their own bespoke analysis with the wealth of data in TILDA including socioeconomic, medical history and objective measured health variables tailored to the aim of their own analyses. For example, some analysts may wish to include control variables for socioeconomic factors which are associated with both myocardial infarction and utility (including income, education and socioeconomic status). Others may wish to include controls for cholesterol values or blood pressure measurements. TILDA data is also linked to national medicine dispensing databases for a proportion of participants of TILDA. Detailed information was also collected on informal care required by participants themselves and given by them to their own parents and partners. This

work may facilitate the examination of family spill-over effects of illness on caregiver utility (which has previously been highlighted as an important research question) (251,252). The longitudinal nature of TILDA will facilitate panel data methods for these research questions.

It is important to note that predicted utility values generated from the application of the mapping model to TILDA are predictions means rather than data (200). Any analysis of predicted utility of TILDA participants should incorporate both the uncertainty of the TILDA analysis and that of the mapping model. One option is to generate simulated data using multiple imputation techniques and then conduct the second stage of analysis of TILDA using the simulated dataset(s). It can be argued that directly observed utilities are themselves estimates or predictions given that they are also estimated from regression equations derived from health state valuations on characteristics of the EQ-5D-3L descriptive system (224).

5.5.3.2 Generation of an improved mapping model

Derivation of a mapping model in a larger, broader, patient population with a greater number of patients in poor health states would reduce uncertainty and ameliorate some of the known limitations with the mapping model.

5.5.3.3 Estimation with other versions of the EQ-5D

An Irish EQ-5D-3L value set is in development. From the data collected, it will be possible to re-estimate the model when the Irish EQ-5D-3L value set is published (222). This will be especially valuable given the dearth of utility values in the literature derived from this value set available on publication. An EQ-5D-5L valuation set for Ireland has been published since the initiation of this study (220). A de novo research study examining mapping to the EQ-5D-5L would be beneficial to those wishing to use this dataset.

5.5.4 Lessons for a future mapping study

After reflecting on the features of this mapping study, a number of potential opportunities for improvement were noted for a future mapping study.

5.5.4.1 Patient population

The eligible patient population for our study was relatively narrow. Only patients with ischemic CVD were eligible. This reduced the ability to recruit participants to the mapping study. Broadening the eligibility criteria would increase the generalisability of our mapping model and facilitate the accrual of a larger sample size. Greater consideration should also be given to the sampling frame used to recruit participants.

5.5.4.2 Defining the TILDA Question Subset

In our TILDA 'Question Subset', we asked participants numerous questions about their history of CVD and the timing of their events. In future studies, the TILDA 'Question Subset' should be refined to include only broad medical history question. This would substantially shorten the time taken to complete interviews and make it easier to accrue a larger sample size.

In our sample, only moderate correlation was observed between the TILDA 'Question Subset' and the Anxiety/Depression domain of the EQ-5D-3L. We excluded asking the CES-D and HADS-A for the reasons described in Section 5.3.3.2.5 above. However, in hindsight, it may be possible to balance the interview duration and ethics consideration by asking a subset of questions from CES-D rather than the whole questionnaire. For example, frequency of participants reporting "I felt depressed" and "I felt sad" could be ascertained rather than the whole of the instrument. A limitation of this approach is that asking a subset of questions only may affect the participant's response. It may also be possible to administer the full CES-D questionnaire after appropriate clinical and ethical consideration. HADS-A was included in the self-completion questionnaire as opposed to the main interview component. This means that the extent of missing data is greater for the HADS-A than the CES-D. This should be taken into account if the consideration is given to including the HADS-A in future studies.

5.6 Conclusion

While this research is primarily of benefit to researchers in Ireland, the methodological approach adopted may be beneficial for researchers who face similar data gaps in their own jurisdictions. Using the mapping model to derive utilities for economic evaluation generates a rich longitudinal data source that can be used to examine the complex relationship between utility and other variables, which may also be of interest internationally. It provides a pragmatic solution to the paucity of utility data in Ireland and ensures that the wealth of data in TILDA can be extracted. However, there are limitations. As in any economic evaluation, the comparative strengths and limitations of all utility sources should be considered before using the model to predict utilities for an economic model. Analysis of uncertainty including sensitivity and scenario analyses play a pivotal role.

In the next chapter, the chosen mapping model to the national TILDA population to generate utility values for cardiovascular disease representative of the Irish health that can be used to inform our economic model in Chapter 7.

Chapter 6. Utility values for an economic model in CVD

6.1 Introduction

A decision analytic model is required to conduct an economic evaluation of PCSK9 inhibitors in Ireland. The economic model by Asaria et al (on which our economic evaluation is based) is described in depth in the next chapter (15). In brief, a population of patients with established CVD are modelled to experience further non-fatal CV events, CV death or non-CV death. Utility values required include:

- Baseline (or starting) value;
- Decrements to account for declines with age and natural disease progression;
- Decrements to account for reduction in health after stroke and MI;
- Treatment associated disutility.

As outlined in Chapter 3, EQ-5D outputs from the ODYSSEY OUTCOMES trial have not been published to date. Therefore, required utility values must be gathered from other sources. A mapping model was derived in the previous chapter so that the depth of data in TILDA could be harnessed to fill this data gap.

Analysis is required to convert the patient-level data in TILDA into health state utility values. The recommendations of international guidelines should be considered (189,253–255). Importantly, given the greater uncertainty associated with mapped values, the comparative strengths and limitations of different sources of utility values should be considered.

6.1.1 Aim

The aim of this chapter is to derive utility values for our economic evaluation of PCSK9 inhibitors in Chapter 7.

6.1.2 Chapter outline

First, the methodological considerations that should be considered when populating an economic model with utilities are presented. Next, a brief overview of CV utilities in the literature is given and approaches used to derive utility in other economic models of CVD are examined. The process of deriving utility values from TILDA is described. The utility values that are applied in the base case and scenario analysis in the economic model in Chapter 7 are presented. The chapter concludes with a discussion of the main findings.

6.2 Considerations when deriving utility values for cardiovascular disease

Compared to a number of disease areas, there is additional complexity when deriving health state utility values in CVD given the extent of co-morbidities, the length of time spent in health

states and the acute and chronic nature of CV events. Important issues to consider when populating an economic model in CVD are described below.

6.2.1 Irish population health and preferences

As described in Chapter 2, Irish national guidelines for economic evaluation recommends the use of generic preference-based measures. For consistency, NCPE the recommend the EQ-5D-3L. As an Irish EQ-5D-3L value set has not been published, utilities should be valued using the EQ-5D-3L UK value set (223,256).

Ideally, the utility values should capture health as measured by the EQ-5D-3L descriptive system in an Irish population. However, utility data from the UK is often employed (given the depth of research in this area in the UK and the geographical proximity to Ireland). However, this may introduce bias. Ireland has the highest self-perceived health in the EU with 84.1% rating their health as very good or good compared to only 73.2% in the UK (257). But these figures are not adjusted for age, sex or other confounding factors.

It is generally difficult to compare utilities across countries due to factors that can influence differences between studies. However, EUROASPIRE studies are standardised across European countries enhancing the comparability (258,259). The EUROASPIRE III study is a cross sectional survey of patients with established CHD in 22 European countries including Ireland (258). Patients were required to be 80 years or less, and hospitalised for coronary artery bypass grafting, percutaneous coronary intervention, acute MI or myocardial ischemia. Eligible patients had events between 2006 and 2007 and were interviewed between six months and three years after the initial event. Out of a total number of 8,745 respondents, 385 (4.4%) were from Ireland and 322 (3.7%) were from the UK. Age and sex balance were similar between countries.[9] There was a substantial difference in mean EQ-5D-3L utilities; 0.71 in the UK and 0.81 in Ireland. The variation between values across all 22 countries was statistically significant.

Further evidence of differences in CV utility between countries is provided by the EUROAPSIRE IV study in which the EQ-5D-5L was administered (259). The eligibility criteria were consistent with the EUROASPIRE III study. In total, 201 patients were recruited from Ireland and 247 from the UK. An age standardised comparison of reported problems for Ireland and the UK is presented in Table 6.1. Across all domains, the percentage of patients in Ireland reporting

[9] The mean age was 62.5 years in Ireland versus 63 years in the UK. The proportion who were male was 76.9% in Ireland compared to 73.6% in the UK.

problems was lower than in the UK. These results would translate to into higher utilities in the population in Ireland. Thus, using health state utility values from other jurisdictions may introduce bias into our economic evaluation.

Table 6.1 Age standardised country-specific reported problems in EQ-5D-5L in populations with CHD

	Mobility (%)	Self-care (%)	Usual Activities (%)	Pain (%)	Anxiety/Depression (%)
Ireland	38.3	10.1	32.4	55.1	34.0
UK	43..9	23.3	46.1	60.3	48.9

CHD, Coronary Heart Disease; UK, United Kingdom.

6.2.2 Baseline health

The baseline health has been described as the utility a person would have if they did not have a particular health condition in a model (260). Baseline values of perfect health should not be used; almost every population has co-morbidities and the prevention of an event will not lead to perfect health (253,261,262). As the starting population in the model has established CVD, the baseline values in the model should reflect the decrement in health that reflects this.

6.2.3 Joint health conditions

6.2.3.1 Theoretical considerations

The starting population in our economic model has a history of CVD. Many, but not all, will have a history of MI or stroke. Therefore, the appropriate disutility for subjects with a history of more than one health condition is important. This is referred to as joint health condition and has been the subject of a number of studies (260,263–267).

A related issue is the appropriate disutility associated with experiencing the same event more than once. For example, quantifying the appropriate decrements associated with experiencing multiple MIs or multiple strokes. However, we are not aware of any study which has specifically examined this issue within economic evaluations.

The effect of joint health conditions on utility can be characterised in one of three ways (268).

- **Additive:** The disutility associated with the joint health condition equals the sum of the disutilities associated with each independent condition.
- **Subtractive:** The disutility associated with the joint health condition is lower than the sum of the disutilities associated with each independent condition.
- **Synergistic:** The disutility associated with the joint health condition is greater than the sum of the disutilities associated with each independent condition.

In general, a subtractive relationship is assumed for joint health conditions (269); the structural assumption that each individual conditions acts independently on utility is unrealistic (264). This will not be the case for all combinations of conditions. For example, the effect on HRQoL of losing mobility in two hands compared to one hand is likely to be synergistic. Thompson et al also observed evidence there could be a synergistic effect between depression and other health conditions (264). However, in most cases the nature of the relationship between the presence of multiple diseases and the associated disutility value is too complex to be predict.

Thompson et al highlight that the effect of joint health conditions on utility may be dependent on the structural assumptions of the value set used (264). For example, patients with joint health conditions are more likely to select level 3 of the EQ-5D (extreme problems). But the N3 term in the UK value set (which places additional disutility for extreme problems across any domain) is not a feature in every value set (223,263).

6.2.3.2 Methods of adjustment when mean utility values are available

Ara and Wailoo have performed a review and critique of the evidence for estimating health state utility values for joint health conditions (with a focus on studies that use mean utilities from cohorts with single health conditions) (260). Five approaches were found. These are described below using the example of a joint health condition of MI and stroke where U_{NC} represents the utility of a population without any condition, U_M represents the utility of the only MI population, U_S represents the utility of the only stroke population and $U_{M,S}$ represents the utility for a population with a history of MI and stroke. Thompson has described the assumptions that each approach has on the underlying theoretical relationship between joint health conditions (263).

6.2.3.2.1 Additive

Under the additive approach the utility of someone with a history of MI and stroke may be represented as:

$$U_{M,S} = U_{NC} - [(U_{NC} - U_M) + (U_{NC} - U_S)]$$

This approach assumes that there is no overlap in the attributes that lead to disutility in MI and stroke (260,263).

6.2.3.2.2 Multiplicative

The multiplicative approach assumes that for each added condition, the disutility is proportional to the disutility without the condition. Therefore, the absolute disutility for each additional condition decreases at a diminishing rate. This is represented by:

$$U_{M,S} = U_{NC} \cdot \frac{U_M}{U_{NC}} \cdot \frac{U_S}{U_{NC}}$$

6.2.3.2.3 Minimum

Under the minimum approach, the minimum utility is set to the lowest of the utilities for the events (260). Under this approach, it is assumed that the utility value associated with the worst condition contains all the attributes that lead to disutility across both conditions (263). This is represented as:

$$U_{M,S} = \min(U_M, U_S)$$

6.2.3.2.4 Adjusted decrement estimator.

The adjusted decrement estimator approach, initially proposed by Hu et al (260,270), is a variance of the minimum approach. The minimum value of the single condition utility represents an upper bound for the joint health condition utility and further decrements are a function of the minimum utility. This is represented as:

$$U_{M,S} = \min(U_M, U_S) - \min(U_M, U_S).(U_{NC} - U_M).(1 - U_S)$$

6.2.3.2.5 Linear Index

The linear index was introduced by Basu et al (266).It is sometimes known as the combination model (260) or a statistical regression model (265). Here, the utility value of the joint health condition is assumed to be a linear function of the additive, minimum and maximum and multiplicative decrements. The linear index weights their respective importance. The parameters of the linear function ($\alpha_0, \alpha_1, \alpha_2, \alpha_3$) obtained by a regression within the dataset itself or from an external dataset (260). It is represented by:

$$U_{M,S} = \alpha_0 + \alpha_0 \max(U_{NC} - U_S, U_{NC} - U_M) + \alpha_2 \min((U_{NC} - U_S, U_{NC} - U_M) + \alpha_3 \left(U_{NC} \cdot \frac{U_M}{U_{NC}} \cdot \frac{U_S}{U_{NC}}\right) + \varepsilon \text{ Where } \varepsilon \sim N(0, \sigma^2)$$

An advantage of the linear index is that it is based on a theoretical framework. This proposes that the value function is convex for losses with a marginal rate of decrement in value for increasing losses and other work on multi-attribute utility functions (260,266).

6.2.3.3 Comparison of approaches

An assessment of which approach is the most appropriate is impeded by the inconsistent methods used across studies (263). For example, the definition of “no condition” utility varies across studies (253,261,263,265).

Often, the data required to calculate a no-condition utility variable is not available. Ara and Brazier have previously concluded that while condition specific baseline is preferable, a general population baseline may be appropriate to use in some cases (261). An exploratory analysis showed that decrements on utility are not constant across age. Some conditions showed an increasing utility decrement with age and others a decreasing trend (which could be due to the presence of co-morbidities).

The conclusion of the review by Ara and Wailoo, on the effect of methods to adjust utilities for joint health conditions, was that no method gave accurate results across the full range of health state utility values (260). Within the confines of the evidence base available they recommended the multiplicative method when only mean values are available. While the linear index method performed well, they stated they it required external validation before it could be recommended.

In the analysis by Thompson et al (264), the linear index performed best followed by the multiplicative method. But no index could be recommended for more than two joint health conditions. An external validation of the linear index was later performed by Thompson where a linear index derived from the UK General Practice Patient Survey was applied to Health Survey for England data (263). Four scenarios were compared. The author concluded that the linear index performed well in the external dataset where the assumptions matched the assumptions used in its derivation. But such assumptions would generally require patient-level data which is often unavailable. The multiplicative approach had highest predictive accuracy in scenarios where patients with co-morbidities were not included in the no condition utility and where a general population baseline utility value was used. The biggest driver in the accuracy of results was the baseline value used. It was not possible to extend the validation study to cases of more than two joint health conditions.

In most approaches tested to date, it is assumed that utility data is available for the subject who only has the condition of interest (and no other co-morbidities). The approach by Thompson et al required very specific analytic results on the presence or absence of co-morbidities (264). Ara et al compares to no condition baseline of a similar age, but again very specific analysis of the data or patient-level data is required to derive this utility (261). A further limitation is that, in estimating the joint health condition using only single conditions, a large proportion of the available data is not used.

The optimal method to adjust utilities for joint health conditions, when only mean data is available, is unclear.

6.2.3.4 *Interaction factors and regression*

Much of the research on utility values for joint health conditions assumes that patient-level data is not available. When patient-level data is available, regression can be used to estimate the disutility (provided that the number of joint health conditions to be estimated is reasonable for the data sample size).

For example, in the case of MI and stroke, a variable is defined which is set to 1 for a person who has a history of MI and Stroke and 0 for others. The decrement associated with a history of MI and stroke is equivalent to the coefficient for MI, stroke and the interaction coefficient MI and stoke. If the relationship between the joint health condition is additive, the coefficient will be zero. Under this approach, data from all participants is used. However, this approach can lead to combinatorial explosion where the number of joint health conditions is large. Ara and Wailoo's highlighted that the focus of their review on joint health conditions was for approaches when patient-level data is not available. However, there is little literature regarding which approaches to take when patient-level data is available.

In an analysis by Fu et al, 25% of mean health state utility values estimated using the minimum method were smaller than the actual mean health state utility values for the joint health condition (270). This is only possible if one of the co-morbidities improved utility. But this is not compatible with accepted theory. Ara and Wailoo acknowledge that a certain amount of anomalies is inevitable because of random error and suggest that the outputs may reflect differences in disease severity between groups (260). An alternative hypothesis is that it may reflect a survivor bias if there is a risk of mortality associated with both events. The survivor bias may still exist even if including an interaction variable in a patient-level regression analysis. Therefore, even a regression-based approach may not be optimal. In any case, the availability of patient-level data is critical to inform the analysis.

6.2.4 Causal health decrements

Disutilities associated with CV events are often calculated by comparing the utility of a population who have experienced the event of interest (such as MI) to a population who has not (255). However, a comparison versus age adjusted baseline values in a population who do not have CVD is not sufficient. This is because CVD events (such as MI) are often associated with other factors which lower utility, including co-morbidities (such as diabetes and chronic obstructive pulmonary disease) and socioeconomic factors (such as income and education). Therefore, the difference between the no CVD population and the MI population also incorporates the utility decrement associated with these confounders. The potential bias that

this could introduce can be elucidated in an analysis of the VALIANT study (271). All patients had a history of CVD at baseline. But as expected there were still substantial differences in patient characteristics between patients who went on to have a subsequent event and those who did not. The mean baseline utility across both groups is 0.77. But the baseline utility for people who went on to have a further CV event was 0.70 compared to 0.80 in patients who did not. By comparing the event group to the total CVD group, a decrement of 0.07 (0.77-0.70) would be inappropriately attributed to the CVD event. These results show that failing to estimate causal utility decrements and ignoring confounders could potentially lead to biased cost-effectiveness results.

The effect of co-morbidities on utility decrements has been previously recognised by Ara and Brazier (261). In their work, they compared age-adjusted utility decrements for subgroups. One subgroup had a health condition of interest and no co-morbidities (Subgroup One). The other subgroup had the same condition and any type of co-morbidity (Subgroup Two). It was found that utility decrements were more than doubled in Subgroup One. These differences imply that ignoring the effects of co-morbidities can introduce significant bias into economic evaluations. No specific guidance was provided about how to address this, except to state that its implications should be considered on an individual basis.

Longitudinal analysis of data is required to robustly calculate causal utility decrements associated with acute events in the CVD setting. But even in the absence of longitudinal data, some adjustments can be made to adjust for confounders. But in many cases, the rich dataset required to do this is not available. Sullivan et al was one of the few studies to acknowledge and make such adjustments in several catalogues of utility values (233–235). These studies are described in 6.3.1 below. The depth of data available in TILDA will allow adjustments for a number of confounders.

6.2.5 Account for acute and chronic effects of events

CVD events (such as MI and stroke) can have a large impact on HRQoL directly after an acute event. However, with time and rehabilitation, many patients regain some of this lost HRQoL. The difficulty in capturing acute decrements in utility has been previously highlighted; patients may not have completed a questionnaire at the time of the acute event. In the absence of utility data collected during or immediately after an acute event, plausible estimates of the QALY loss per event should be used in sensitivity analyses (189).

Utility values should also be estimated for the chronic health state. Most analyses assume the decrement for the chronic health state to be constant after adjusting for age and other modelled

events. But damage to the heart following an MI increases the risk of heart failure at the time of the event and in future years. The appropriate extrapolation of costs and utilities over time across all evaluations has been previously highlighted as an important research question (191). Ideally, longitudinal analysis is required in order to estimate this reliably.

6.2.6 Adverse event associated with treatment

Utility decrements due to adverse events associated with treatment should be incorporated into the economic model (189).

6.2.7 Sensitivity and scenario analyses

Guidelines recommend that one-way and multi-way sensitivity analyses of utility values should be conducted to investigate uncertainty in an economic model (189).

6.2.8 Summary

There are many factors which must be considered when deriving utilities for an economic model. There is some evidence to suggest that health in the population of Ireland, as measured using the EQ-5D-3L, may be systematically different to that in the UK. This may introduce bias into evaluations in Ireland when UK data is used. There are numerous areas where ambiguity in the literature remains regarding the best methods to use. This includes issues pertaining to the estimation of joint condition health state utility values.

6.3 Cardiovascular utilities derived from the EQ-5D in the literature

As described in Chapter 5, we are not aware of any national study that estimates CVD utilities using the EQ-5D-3L in Ireland. Therefore, a description of international studies is provided below.

6.3.1 Utility catalogues estimated by Sullivan et al.

6.3.1.1 Study overview

Three utility catalogues (which include utilities associated with CVD events) have been developed by Sullivan et al (233–235). An analysis of the approach taken in these studies is important because:

1. Utilities adopted in the original Asaria et al economic model (which is the basis of decision analytic model in the next chapter) are derived these catalogues (15,234).
2. The catalogues were derived using a patient-level national dataset, therefore methods applied here could be applied to analysis of predicted utilities in TILDA.

The characteristics of each catalogue are summarised in Table 6.2. All three are based on EQ-5D-3L data collected in the national US Medical Expenditure Panel Survey (MEPS). While, MEPS is a longitudinal survey, Sullivan et al only exploited the cross-sectional portion of the dataset (233–235).

The catalogues differ in the value sets applied, the disease classification system used, and coverage years applied. Three different disease classification systems are employed:

- Quality Priority Conditions (QPC)
- 9th Version International Classification of Diseases (ICD-9)
- Clinical classification categories (CCC)

QPC classifications collect data on conditions experienced by patients in the past. In contrast, the CCC and ICD-9 classifications systems collected data on conditions experienced in the survey year only (233,235). The discriminatory capacity of the CCC and ICD-9 across diseases is greater than the QPC (with each accounting for 100s of disease categories compared to only ten in the QPC). The third version of the catalogue is of most relevance to our research (234). It values health using the UK EQ-5D-3L value set (preferred by the NCPE, Ireland) and other international HTA agencies (189).

Table 6.2 Summary of utility catalogue characteristics estimated by Sullivan et al (233–235)

		N	Years Covered	US Population	EQ-5D-3L Value Set	Disease Classification.		
						QPC	CCC	ICD-9
1	Sullivan et al (2005) (233)	38,678	2000-2002	Yes	US	✓	✓	×
2	Sullivan and Ghushchyan (2006) (235)	38,678	2000-2002	Yes	US	×	×	✓
3	Sullivan et al (2011) (234)	79,522	2000-2003	Yes	UK	×	✓	✓

CCC, Clinical Classification Categories; ICD-9, International Classification of Disease – 9th Version. QPC, Quality Priority Conditions.

6.3.1.2 Presentation of results

In each of the three catalogues, the authors present two sets of analyses for each disease classification system – adjusted and unadjusted (233–235).

6.3.1.2.1 Unadjusted

The directly observed mean, median and interquartile utility values are presented for each disease category and for the total sample (after adjusting for survey methods) (233–235). This

is supported with information about the number of people, the mean and median age and the mean number of co-morbidities for each disease category. The unadjusted mean utility values are baseline utility values for each cohort.

6.3.1.2.2 Adjusted

Regression methods were employed and accounted for the sampling frame used in the MEPS (including clusters and sampling weights[10]) (233–235). CLAD regression models were used to calculate adjusted utility values by regressing predicted utility on a range of independent factors. Factors included age, sex, race, income, education, dummy variables for each of the disease classifications and a variable classifying the number of co-morbidities. The regression indicates that multiple conditions have a synergistic additive effect on utility when the number of co-morbid conditions is six or lower. When the number is seven or higher, the effect of multiple conditions may be synergistic or subtractive (depending on the size of disease coefficients).

6.3.1.2.3 How to use the catalogues to predict utility values

Sullivan et al provides an example describing the intended use of the catalogues (233–235). The baseline value should be estimated from the unadjusted results. This value incorporates the decrements associated with all co-morbidities and the effects of socio-demographic factors. When the starting population are modelled to experience additional health events, these utility decrements should be taken from the adjusted catalogue. For modelling the effects of diabetes and stroke, for example, the decrements for diabetes, stroke, and the decrement associated with having two co-morbid conditions should be subtracted from the unadjusted baseline value. An age decrement should be used to adjust the value for population aging.

6.3.1.3 Advantages and limitations associated with the catalogue

There are a number of advantages associated with the Sullivan et al 2011 catalogue (234):

1. It was based on a US national survey with a very large sample size which increases the precision of the estimates;
2. Trained coders coded patient reported health status into ICD-9 Codes increasing consistency in classifications;
3. It clearly differentiates between the baseline utilities of a population with a disease and decrements associated with the health events after controlling for confounders.

[10] A more detailed description for regression methods adjustments in survey data is described in Section 6.5.1.1.2 below.

However, it also has several limitations which limits its generalisability to our research question:

1. Measures of US health may not be transferable to Ireland.
2. Marginal decrements were estimated using CLAD regressions which capture median rather than mean decrements in health (226). Mean decrements are most relevant in an economic model.
3. In the catalogue that used the EQ-5D-3L UK value set, decrements are only estimated for patients who experienced an event in the year preceding the interview. Disutilities will not be representative of the chronic health state in the years following the CV event.
4. The age decrement was estimated after adjusting for all acute conditions. Therefore, the effect of increasing co-morbidity with age is not fully incorporated.

While the catalogue has several advantages, the extent of the limitations means that its use could introduce substantial bias into our economic evaluation. Other sources of data are required.

6.3.2 Health Survey for England data

The Health Survey for England is an annual national UK survey measuring health in the population of England. Unlike similar Irish studies, EQ-5D data was collected. Every year, data on core questions relating to general health was also collected. The 2003 and 2006 surveys included questions on CVD. A detailed analysis has been previously presented by Ara et al (253).

Baseline utilities were estimated with an OLS regression analysis, of predicted utility, on age and sex in the no CVD population only (253). But this approach means that factors associated with both CVD and age are not fully accounted for in the age decrements. Utility multipliers were calculated using the methods outlined in Section 6.2.3.2.2. The results are presented in Table 6.3 below.

Table 6.3 Utility multipliers calculated from Health Survey for England data reported by Ara et al (253).

	Utility	Male*	Age (years)	Baseline	Multiplier Event
Acute MI	0.721	0.57	65.4	0.8199	0.8794
Acute Stroke	0.626	0.57	67.9	0.8095	0.7733
No event <12 months, history of angina only	0.775	0.57	68.8	0.8057	0.9619
No event <12 months, history of heart attack only	0.742	0.57	65.1	0.8211	0.9037
No event <12 months, history of stroke only	0.668	0.57	66.8	0.8141	0.8205

*The proportion who are male is not reported by Ara et al(253). Therefore, we assumed it to be 0.57 in line with the proportion who are male in the base case population in TILDA.
MI, Myocardial Infarction; TILDA, The Irish Longitudinal Study on Aging.

Using the same data, Stafford et al examined socioeconomic differences in the HRQoL impact of CV conditions (272). In contrast to Ara et al (253), Stafford et al conducted multiple regressions of utility on CV variables and a number of confounders (272). After adjusting for age group, sex, ethnicity, education and socioeconomic position, the mean utility decrement for each condition was angina (-0.141), heart attack (0.139) and stroke (-0.16). After accounting for the impact of all conditions simultaneously (including obesity and hypertension and diabetes), mean decrements fell to -0.090, -0.060 and -0.101 for angina, heart attack and stroke respectively.

Strengths of the Health Survey for England data that it is based on nationally representative data (albeit in England) and also the breadth of analyses conducted. However as highlighted in Section 6.2.1, values from England may not be transferable to the Irish setting. Utilities are not controlled for potential confounders. No information is presented for patients with other forms of CVD.

6.3.3 Other studies

In a longitudinal study, Munyombwe et al administered the EQ-5D 3L questionnaire and EQ-VAS scores at hospitalisation and at 1, 6 and 12 months post-event, to an eligible cohort of 9,566 patients (273). Response rates were high initially but declined over the length of the study. The mean EQ-5D-3L value were 0.72 (Standard error [SE] 0.3) at hospitalisation and 0.78 (SE 0.3) at 12 months post event. However, while some analyses of EQ-VAS scores was adjusted for missingness, these results are not. Over 40% of data was missing at 12 months. Predictors of missingness for the EQ-VAS are likely to be similar to the EQ-5D including baseline value,

ethnicity, age, diabetes and previous MI. Therefore, these values are very uncertain. At all time-points, patients with a history of non-ST elevation myocardial infarction (NSTEMI) had lower HRQoL scores compared to patients with a history of ST elevation myocardial infarction (STEMI). A limitation of the study is that most of the analysis was conducted on the EQ-VAS score. While an important outcome measure in its own right, it is a not a preference-based measure.

Amgen (the Marketing Authorisation holders of evolocumab) sponsored a systematic review of CV event utilities for MI and stroke (274). The authors highlighted the large differences in results across studies. They noted that this was due to the differing definitions of CVD events, heterogeneity in sample characteristics and the different assessment tools used.

Baseline utility values from a pooled analysis of 4,203 patients registered in early alirocumab studies in the ODYSSEY program are available in abstract form (161). The mean age was 59 years and 63% were male. The results are presented in Table 6.4 below (as valued using the UK value set).

Table 6.4 Mean utility values in the ODYSSEY program

		Total Population		No other CV event condition		At least one other CV event condition	
Condition	**Age**	**N**	**Utility (SE)**	**N**	**Utility (SE)**	**N**	**Utility (SE)**
ACS in last 0-12 months	56.2	198	0.844 (0.197)	142	0.848 (0.201)	56	0.832 (0.189)
ACS in last 12-24 months	58.7	192	0.858 (0.187)	120	0.874 (0.185)	72	0.832 (0.190)
CHD	61.4	2731	0.851 (0.194)	813	0.860 (0.191)	1918	0.847 (0.195)
Ischemic Stroke	63.8	344	0.797 (0.228)	164	0.804 (0.212)	180	0.791 (0.242)
PAD	62.8	188	0.771 (0.233)	98	0.775 (0.253)	90	0.767 (0.211)

ACS, Acute Coronary Syndrome; CHD, Coronary Heart Disease; CV, Cardiovascular; PAD, Peripheral Arterial Disease; SE, standard error.

Having reviewed those CV utilities (estimated using the EQ-5D-3L) that have been published in the literature, approaches taken to estimate utility in economic models of CVD are now examined.

6.4 Approaches taken in other models

6.4.1 CALIBER model approach.

The CALIBER model adopted utilities developed from the Sullivan et al 2011 catalogue (see Section 6.3.1) (15,234). However, when estimating their utility values, Asaria et al did not use the approach described by Sullivan et al (see Section 6.3.1.2.3) (15,234).

Instead of using a starting utility value from the unadjusted disease specific values, Asaria et al used a general population value and applied acute utility decrements according to the history of CVD disease to predict the model starting value (15,234). Combining multiple utility values increases uncertainty. A co-efficient for male sex was included but differences between male and female are already incorporated into the baseline value.

There was minimal differentiation in utility between acute and chronic health states. Asaria et al considered two approaches (15):

a. assumed that the disutility derived from an acute event lasted a lifetime.
b. assumed the disutility derived from an acute event lasted for one year only. In this scenario, the starting utility value in the model was not adjusted for the proportions of stroke or previous MI. The utility decrements used are presented in Table 6.5.

Table 6.5 Utilities used in the CALIBER economic model by Asaria et al (based on Sullivan et al 2011 catalogue) (15,234)

General population	Age	Male	Acute MI	Previous MI	Angina	HF	Stroke
0.828	-0.0003	0.001	-0.0626	-0.0368	-0.0854	-0.1167	-0.1171

HF, Heart Failure; MI, Myocardial Infarction.

6.4.2 Ara et al assessment of cardiovascular models.

When undertaking an evaluation, it is useful to examine approaches adopted in other models in CVD. However, this process is difficult when the source of utility values or the estimates used in the model themselves are unclear. Ara et al conducted a review of utilities, employed in cost-effectiveness evaluations of lipid lowering therapies, published between Jan 2014 – July 2017 (269). Only four of the 24 studies identified reported all health state utility values accurately when compared to values in the original source studies. Only six studies referenced the original sources for all utilities. Further, half the studies did not specify the instrument or method used to obtain the utility values. In many cases, it is not possible to identify the source even when referring to multiple reference sources (269). While reporting quality does not equate to the quality of the study, it is not possible critically assess the quality of a study when poorly reported.

Ara et al summarised the health state utility values used in each model identified and the methods used to adjust for multiple events (269). Relevant results are reproduced in Table A.12 in the Supplementary Appendix. None of the secondary prevention studies included age adjusted utilities for baseline health. Only seven of the primary prevention included age adjusted estimates of baseline health. Many studies inappropriately assumed a constant baseline health of 1 in the primary prevention. Reporting methods for multiple event adjustments were poor (269).

6.4.3 Utilities adopted in NICE technology appraisal of PCSK9 inhibitors.

As only summaries of NCPE reports are published, it is not possible to describe utilities used in company submissions to the NCPE (9–11). NICE technology appraisals are publicly available.

In the UK technology appraisal for evolocumab, the company based the health state utility values on those used in a NICE's technology appraisal for lipid modification in CVD (36,275). The Evidence Review Group had concerns; many utility values did not match the states in the economic model. Most of the utilities were estimated from studies that used the time trade off and did not meet the NICE reference case for use of the EQ-5D-3L. Therefore, this data is unlikely to inform our economic model.

The utilities used in the UK alirocumab technology appraisal were based on the Health Survey for England data reported by Ara et al (33,253) (see Section 6.3.2).

6.5 Generation of cardiovascular utilities for the economic model

Given the multiple considerations required of deriving utilities, an analysis of a patient-level national longitudinal survey of utilities would be an extremely valuable resource to estimate utilities for a CVD economic model. This would allow the analyst to control for multiple CVD events, co-morbidities and confounding factors in a dataset measuring health in the population of Ireland.

While such a dataset is not available in Ireland, the mapping model generated in Chapter 5, allows a utility value to be predicted for each TILDA participant. While TILDA is representative only of the population over 50 years of age, this captures most of the CVD population in Ireland.

A strength of TILDA is its longitudinal design; this allows follow-up of patients for over 10 years and allows participants' utility to be observed pre and post event. The analysis of longitudinal data is complex. Therefore, analysis of the national TILDA population presented here represents a cross sectional analysis (given that analysis is limited to Wave 1 only).

Given the parameter and methodological uncertainty, four scenarios are considered in our economic model:

1. TILDA-1: analysis of the predicted utilities in TILDA using an approach (similar to that of Sullivan et al) (233–235) but with a number of modifications. TILDA-1a describes the base case scenario. Several variations of this scenario (TILDA-1 b-d) are also examined so that the effect of methodological assumptions can be ascertained.
2. TILDA-2: analysis of the mean predicted utility values of TILDA adjusted using a multiplicative approach (similar to that of Ara et al) (253).
3. Scenario 3 Uses published mean Health Survey for England data (253,272).
4. Scenario 4 uses the utility estimates applied by Asaria et al (15).

The methods used to derive the utilities under the TILDA-1 and TILDA-2 scenarios are described below.

6.5.1 Methods

6.5.1.1 Common methods for TILDA scenarios

6.5.1.1.1 Application of mapping model

Statistical Analysis was conducted in Stata® version 15.1 (StataCorp, College Station, Texas). The ALDV mapping model was applied to national TILDA population using the *'predict'* command in Stata® to generate a predicted utility value for each TILDA participant. As highlighted in Section 5.5.3.1, these values represent predicted mean utility value rather than data. It would be best practice to use the mapping model to generate simulated data using multiple imputation techniques and then conduct the second stage of analysis of TILDA using the simulated dataset(s). For simplicity this analysis was not conducted, and the predicted means were used directly instead. Consequently, the uncertainty associated with the utility values is underestimated.

Complete case analysis is used for all analyses. There was no imputation for missing values. Less than 1% of participants were missing a predicted utility value.

6.5.1.1.2 Accounting for the TILDA sampling frame

As the sampling frame used to recruit participants to TILDA has implications for the analysis of the data, the method is described here. Participants were recruited into TILDA using samples based on the RANSAM system which has been used widely for survey design in Ireland (195). All residential addresses in the country were aggregated into 3,155 clusters. 640 of these clusters were randomly selected with implicit proportionate stratification of clusters by socioeconomic

group and geography. The second stage involved the selection of 50 addresses within each cluster. Each of the fifty addresses were approached and every household member over 50 years of age was invited to take part in the study. Partners of eligible participants were also enrolled even if they were under 50 years of age (195). These participants were excluded from this analysis.

Stratification and clustering have different implications on standard errors. Proportional stratification reduces sampling error below those of a random sample. Clustering increases sampling error because members of clusters are more likely to resemble other members of the cluster compared to a random member of the sample. However, clustering is often used in survey design methods as it improves the efficiency of the data collection process. While, stratification and clustering do not affect the estimation of point estimates, they effect the calculation of the standard error which has implications for uncertainty and interpreting statistically significant differences (276).

The overall response rate to TILDA was 62% (195). Hypothetically, if the response rate were unrelated to sample characteristics, no adjustment would be needed. But differential non-response, where response rates vary across subgroups (such as sex or age) can lead to biased estimates. To account for this, sampling weights are calculated corresponding to the number of members of the total population of Ireland that are represented by this participant. This means that participants from subgroups who are less likely to respond are assigned higher weights than those from subgroups more likely to respond.

To estimate these weights, TILDA calculated weights based on age, sex and educational attainments compared to the Quarterly National Household survey in 2010 (195). Adjusting patient weights in this way accounts for non-response bias attributable to these characteristics only. If non-response is due to other characteristics (such as income or health) weighting will not eliminate the source of bias for these estimates. Given the different rates of non-response across the three different components of TILDA, (interview, self-completion questionnaire and health assessment) different weights are used depending on each analysis.

To incorporate the strata, clusters and weighting into the analysis, the survey data analyses features of Stata® were employed. At the start of the analysis, the '*svyset*' command was used to define the pre-defined factors in TILDA corresponding to the clusters, strata and sampling weight. The '*svy*' command was incorporated into all further analyses to incorporate these features.

6.5.1.1.3 Choice of regression model

The potential choices of regression models were described in Section 5.3.4.3 of Chapter 5. We want to regress utility on age, sex and CV characteristics. In this case, our interest lies in the estimates of the beta coefficients representing the mean effects of the co-variates (rather than on predicting the individual results accurately). In line with the results of Pullenayegum et al, OLS models will predict beta coefficients accurately provided the model is correctly specified (227). Therefore, we use OLS models to estimate the average associations between predicted utility and the included co-variates.

6.5.1.2 Initial utility value

Identical approaches to defining initial utility values are used in both the TILDA-1 and TILDA-2 scenarios. In line with the approach adopted by Sullivan et al, the initial utility is the mean predicted utility for the sample under consideration (233–235). Predicted utilities are adjusted for TILDA sampling frame described in Section 6.5.1.1.2 but no other adjustments are applied.

Given the heterogeneity of the population, 23 populations were defined which represent different manifestations and severity of CVD. Derivation of population characteristics is defined in greater detail in the Chapter 7.

6.5.1.3 Utility decrements for age

6.5.1.3.1 TILDA-1

Modelling the age decrement in the model is important because it captures the decline in utility associated with increased morbidities with aging. In the TILDA-1 scenario, these are estimated using regression. It is not advisable to categorise continuous factors (such as age). However, in our version of TILDA dataset, age was censored at 80 (for anonymity reasons); categorisation was necessary.

Two regressions were conducted. Regression 1 regressed predicted utility on age, sex, and a range of CV variables. Some analyses have shown that the expected utility decline with age differs by sex. This is rarely incorporated into economic models. Therefore, Regression 2 included the same co-variates as Regression 1, plus a variable modelling an Age-Sex interaction.

Like Sullivan et al, co-variates for CVD are included; the disutility associated with CVD is incorporated separately in the model through the initial starting utility and disutilities for modelled CVD events (233–235). However, unlike Sullivan et al, co-variates for other co-morbidities and socioeconomic factors are not included. This is because the aim of the analysis is to incorporate the effect of other co-morbidities on utility over time (233–235).

The three TILDA-1 scenarios (1a, 1c, 1d) account for the age-sex interaction by adopting the age decrements in Regression 2. In scenario TILDA – 1b, the interaction variable is not included so that its impact can be elucidated by comparing model results to TILDA-1a.

6.5.1.3.2 TILDA-2

Scenario TILDA-2 is based on the methods adopted by Ara et al so that a comparison between results estimated using data from Ireland and England can be conducted (253). The regression analysis for TILDA-1b (which does not account for the age-sex interaction) is used estimating the no-CVD decline in utility associated with age in this scenario consistent with the approach by Ara et al which did not include an interaction.

6.5.1.4 Acute utility decrements for MI and stroke

While information on time of MI and stroke was collected in TILDA, detailed data was not reported in the dataset available for this analysis. Therefore, it is not possible to calculate acute utility decrements from TILDA for this analysis. Given the strengths of the Health Survey for England data, utility decrements, from the analysis of acute CVD events by Ara et al were utilised for all analyses in TILDA-1 and TILDA-2 scenarios (253). Utility multipliers are estimated calculated from the paper using the methods described in Section 6.2.3.2.2.The multiplier is applied to the initial utility value after adjustments for model time.

6.5.1.5 Chronic utility decrement associated with MI and stroke.

6.5.1.5.1 TILDA-1

Utility decrements for chronic CVD events were estimated by regressing predicted utility on CV event variables. Given the complexity, five additional regression analyses were conducted (Regressions 3-8).

Regression 3 and 4 examine the association between predicted utility and MI and stroke respectively after controlling for age, and sex.

Regression 7 examines the association between predicted utility, MI and stroke after controlling for a range of CV and other co-morbidities in addition to education and sociodemographic factors. The inclusion of factors representing multiple events facilitates the measurement of disutility associated with multiple events. Variables for heart failure and for psychiatric problems are not included in the list of co-variates. This is because CV events can lead to depression and damage to the heart caused by an MI can lead to heart failure (277). Therefore, including these factors as co-variates could lead to an underestimation of the utility decrements associated with MI and stroke.

Regression 6 is the equivalent to Regression 5 with a variable representing an interaction between MI and stroke added.

Regression 7 is equivalent to Regression 5 except coefficients for education and socio-economic factors are not included (to account for the impact of these factors).

6.5.1.5.2 TILDA-2

Chronic utility decrements were investigated by applying the method used by Ara et al to the TILDA dataset (253). Utility multipliers were calculated using the following steps:

1. The mean predicted utility value was calculated for the just MI population (including other heart condition) and just stroke population (excluding other heart condition) as already described in Section 6.5.1.2.
2. The equivalent no-CV utility value was estimated using Regression 2 to the equivalent age and sex characteristics of the CV event population.
3. The utility multiplier for the event was calculated by dividing the mean predicted event utility multiplier by the equivalent no-CV utility value obtained from the Regression 2.

6.5.1.6 Treatment associated disutility

In the systematic review of PCSK9 inhibitors in Chapter 3, no increase in serious adverse events was observed for PCSK9 inhibitors versus placebo. Therefore, no utility decrement is applied.

6.5.1.7 Sensitivity and scenario analysis

It is important to note that predicted utility values generated from application of the mapping model to TILDA are predictions means rather than data. Any analysis of predicted utility in TILDA should ideally incorporate both the uncertainty of the TILDA analysis and that of the mapping model. It can be argued that directly observed utilities are themselves estimates or predictions (given that they are also estimated from regression equations derived from health state valuations on characteristics of the EQ-5D-3L descriptive system) (200,224).

As described in Chapter 3, to incorporate the mapping uncertainty, variance co-variance matrices representing the uncertainty should be used to generate multiple versions of the mapping model using multiple imputation methods. The datasets should be combined using Rubin's rules to generate a final version of predicted values for TILDA which incorporates the uncertainty of the mapping model. However, given the complexity, the uncertainty in the mapping algorithm is not incorporated into this analysis which is a substantial limitation of this analysis.

However, we do incorporate the remainder of the uncertainty associated with estimating utilities from the National TILDA sample. The uncertainty in the initial utility values is estimated by sampling from the utility values using the beta distribution.

Ideally, the correlation between predicted utility decrements in should be accounted for in the sampling frame by using the variance co-variance matrix to estimate the Cholesky Decomposition Matrix and jointly sampling from the normal distribution for the analysis in this way. However, for simplicity in this analysis each decrement was sampled independently.

6.5.2 Results

6.5.2.1 Initial Starting Values

Mean predicted utility values for the 23 defined populations are presented in Table A.13 in the Appendix. Sociodemographic characteristics for each group are also presented. The mean utility value for the base-case population (Population 1) for anyone with a history of angina MI, TIA or stroke is 0.7788 (95% CI 0.7631 to 0.7945). The other CVD population includes TILDA participants with other forms of heart disease including those with a history of revascularisation. But it also contains participants with non-atherosclerotic heart disease. When these participants are also included (Population 2), the mean predicted utility is similar at 0.7847 (95% CI 0.7705 – to 0.7990).

The lowest predicted mean utility value was observed in Population 15 which includes TILDA participants with a history of more than one stroke (0.5483; 95% CI 0.4235 to 0.6730). The highest mean predicted utility was observed in Population 12 which includes TILDA participants with only a history of TIA (0.8313; 95% CI 0.8050 to 0.8577).

6.5.2.2 Utility Decrements for Age

The results of Regressions 1 and 2 which regressed predicted utility on age and sex is shown in Table 6.6. After controlling for a range of CV factors, Regression 1 shows that being male is associated with a higher predicted utility than female, and that predicted utility declines with increasing age. An interaction between age and sex was added in Regression 2. The results show that the association between utility and age differs by sex and the decline is much greater for women than men. Age decrements are used to account for the decline in utility associated with age and the related co-morbidities in the economic model. Age coefficients in Regression 2 are used in the base case – Scenario TILDA-1a and its variants 1c-1e. Age coefficients from Regression 1 are used in TILDA-1b and TILDA-2 utility scenarios. Sex decrements are not included in the model as the effect of sex is already incorporated into the initial utility values in the model.

Table 6.6 Coefficients and standard errors for age and sex for predicted utility regressed on sociodemographic and cv variables in an ordinary least squares (OLS) model.

	Regression Model (1) **Coefficient (SE)**	**Regression Model (2)** **Coefficient (SE)**
Age 50-59 (Ref)		
Age 60-69	0.0005	-0.0023
	(0.0039)	(0.0057)
Age 70-79	-0.0129*	-0.0249**
	(0.0056)	(0.0078)
Age 80+	-0.0499***	-0.0728***
	(0.0108)	(0.0149)
Male	0.0337***	0.0221***
	(0.0036)	(0.0055)
Age 60-69 and Male		0.00564
		(0.00799)
Age 70-79 and Male		0.0256*
		(0.0109)
Age 80+ and Male		0.0576**
		(0.0179)
Constant	0.8690***	0.8749***
	(0.0034)	(0.0040)
N	8461	8461

Additional covariates included in both models Any MI, >1 MI, angina, diabetes, congestive heart failure, transient ischemic attack, any stroke, >1 stroke, other heart trouble (as defined by variable PH201_15 in TILDA).
N, Number; Ref, Reference; SE, Standard Error.

6.5.2.3 Acute Utility Decrements

Acute utility decrements from Ara et al were used for all scenarios (253). To calculate utility multipliers, the proportion who are male is required. But as this was not reported by Ara et al (253), it was assumed that 57% of the population were male in line with sex proportions reported in Population 1 of the analysis of the national TILDA population (Supplementary Appendix Table A.13)

The results are shown in Table 6.7. There were two options available in Ara et al for acute utility for both heart attack and stroke (253):

- history of event and other CV condition
- history of event only.

The utility multipliers for “just heart attack” and “just stroke” were chosen for use in the model so as not to double count the effect of co-morbid conditions. Utility values in Ara et al were consistent with those by Munyombwe et al (253,273).

Table 6.7 Acute event utility multipliers estimated from Ara et al (253)

	N	Observed Utility Ara et al.	Male proportion Assumption from TILDA	Age Ara et al,	Comparable No CV baseline*	Multiplier
Heart Attack <12 months, history of heart attack + another CV condition	36	0.431	0.57	66.7	0.81	0.53
Stroke <12 months, history of stroke + another CV condition	18	0.479	0.57	73.5	0.79	0.61
Heart Attack <12 months, history of just heart attack	31	0.721	0.57	65.4	0.82	0.88
Stroke <12 months, history of just stroke	76	0.626	0.57	67.9	0.81	0.77

*No CV baseline is calculated from the No-CV baseline reported by Ara et al: EQ-5D = 0.9454933 + 0.0256466*male - 0.0002213*age - 0.0000294*age^2.
CV, cardiovascular; TILDA, The Irish Longitudinal Study on Aging.

6.5.2.4 Chronic Utility Decrements

6.5.2.4.1 TILDA-1

The predicted utility decrements associated with CV events after adjusting for age and sex only are presented in Table 6.8.

The results of Regression 3a show that MI is associated with a large reduction in utility of -0.071 (95% CI -0.094 to -0.0482). After, including co-variates for angina and heart failure which represent the main symptomatic manifestations of coronary artery disease the decrement decreases to -0.0273 (95% CI-0.0517 to -0.0028). Regression 4 shows that stroke is associated with a substantial utility decrement, more than twice the decrement after MI.

Table 6.8 Coefficients and standard errors for age and sex for predicted utility regressed on sociodemographic and cardiovascular variables in an OLS model

	Regression Model (3a)		Regression Model (3b)		Regression Modell 4	
	Mean (SE)	**95% CI**	**Mean (SE)**	**95% CI**	**Mean (SE)**	**95% CI**
MI	-0.0711 (0.0116)	-0.094 to -0.0482	-0.0273 (0.0125)	-0.0517 to -0.0028		
Angina			-0.0968 (0.0247)	-0.1453 to -0.0482		
Heart failure			-0.0943 (0.0133)	-0.1203 to -0.0682		
Stroke					-0.1537 (0.0252)	-0.2031 to -0.1042

Additional coefficients include age, Sex and Age-Sex interaction.
MI, Myocardial Infarction; CI, Confidence Interval; OLS Ordinary Least Squares; SE, Standard Error

The decrements for the chronic utilities associated with CV events (after adjusting for confounders and co-morbidities) are presented in Table 6.9. As described in Section 6.5.1.5.1, heart failure is not included as a co-variate, as to do so may underestimate the utility decrement associated with MI.

Regression 5 calculates the chronic utility decrements used in the base-case utility scenario TILDA-1a. The size of the decrements decreases compared to those estimated in Table 6.8 which did not adjust for co-morbidities.

Regressions 6-9 examine various scenarios analysis. Regression 6 includes a variable representing a MI-Stroke interaction. The results of the analysis do not meet the expectation that utility should decline after each subsequent event. For example, the utility decrement associated with stroke is -0.1038. If a person goes onto experience an MI, utility is modelled to increase by 0.0149, (-0.0207 [Any MI] +0.0356 [MI and stroke]). This could be due to chance because of the wide confidence intervals around each estimate or because of a survivor bias.

Regression 7 excludes the age-sex interaction. As expected, there is little change in the utility decrements in this scenario (Scenario TILDA-1b).

Regression 8 excludes the education and socioeconomic factors. There is minimal change in the chronic utility decrements (Scenario TILDA-1c).

Regression 9 excludes the angina variable (Scenario TILDA-1d). Therefore, much of the disutility associated with angina is captured in the MI and >1 MI coefficients. It is excluded in this regression; as angina is one of the main symptomatic manifestations of CHD. Therefore, the

base-case analysis in Regression 5 may not reflecting the disutility associated with a reduction in chronic events sufficiently.

Table 6.9 Coefficients and standard errors for MI and stroke coefficients for predicted utility regressed on sociodemographic and cardiovascular variables in an OLS model

	Mean (SE)	95% CI
Regression 5 – Base case TILDA1a		
Any MI	-0.0191 (0.0146)	-0.0477 to 0.0095
>1 MI	-0.0291 (0.029)	-0.086 to 0.0279
Any stroke	-0.0994 (0.0253)	-0.1492 to -0.0497
> 1 stroke	-0.1677 (0.0667)	-0.2986 to -0.0368
Regression 6 – includes MI -Stroke interaction		
Any MI	-0.0207 (0.0145)	-0.0491 to 0.0078
>1 MI	-0.0311 (0.0295)	-0.0891 to 0.0268
Any stroke	-0.1038 (0.0255)	-0.154 to -0.0537
> 1 stroke	-0.1767 (0.0712)	-0.3166 to -0.0368
MI and stroke	0.0356 (0.0696)	-0.101 to 0.1722
Regression 7 – TILDA -1b excludes Age-Sex interaction		
Any MI	-0.019 (0.0147)	-0.048 to 0.0099
>1 MI	-0.0275 (0.0292)	-0.0847 to 0.0298
Any stroke	-0.0965 (0.0252)	-0.146 to -0.0471
> 1 stroke	-0.1676 (0.0672)	-0.2996 to -0.0357
Regression 8 –TILDA -1c excludes education and socioeconomic factors		
Any MI	-0.0142 (0.012)	-0.0377 to 0.0094
>1 MI	-0.0365 (0.0254)	-0.0864 to 0.0135
Any stroke	-0.1015 (0.0228)	-0.1463 to -0.0567
> 1 stroke	-0.1635 (0.0647)	-0.2906 to -0.0365
Regression 9 TILDA 1d excludes angina-variable)		
Any MI	-0.0391 (0.014)	-0.0666 to -0.0116
>1 MI	-0.0423 (0.0284)	-0.098 to 0.0134
Any stroke	-0.1016 (0.0257)	-0.1521 to -0.051
> 1 stroke	-0.1654 (0.068)	-0.299 to -0.0318

Regression 5 includes covariates for age, sex, age-sex interaction, angina, diabetes, TIA, other heart, long term lung condition, asthma, cancer liver, arthritis, stomach ulcer, varicose ulcer, osteoporosis, education and socioeconomic factors
Regression 6 = Regression 5 + MI-Stroke Interaction
Regression 7 = Regression 5 - Age-Sex Interaction
Regression 8 = Regression 5 - Education and Socioeconomic factors
Regression 9 = Regression 5 – Angina
MI, Myocardial Infarction; CI, Confidence Interval; OLS Ordinary Least Squares; SE, Standard Error, TIA, transient ischemic attack.

6.5.2.4.2 TILDA-2

The utility multipliers for chronic MI and stroke calculated from TILDA are presented in Table 6.10. The relative utility decrement is greater for stroke (0.8416) than for MI (0.9303).

Table 6.10 Calculation of utility multipliers from the national TILDA population

	Mean Predicted Utility	Male	Median Age (Years)	No CV Baseline	Multiplier
	A	B	C	D	E
Just Stroke (14)	0.7364	0.525	71	0.87507	0.8416
Just MI (9)	0.8313	0.757	69	0.89359	0.9303

A, B, C are from Table A.13 in Supplementary Appendix. D = From Regression 1 in Table 6.6; E= A/D.

CV, cardiovascular; MI, Myocardial Infarction, TILDA, The Irish Longitudinal Study on Aging.

6.5.3 Summary of utilities used in the economic model

A summary of utility values used in the model in each scenario and their source is presented in Table 6.11.

As described in the previous chapter, the use of predicted utilities generated through mapping over observed utility values should be justified. The multiple factors that must be considered before selecting populating an economic model with health state utilities are described above.

Although mapping utilities have more uncertainty that directly observed values, the selection of the predicted utility from TILDA is justified because:

- They measure population health in Ireland.
- They allow the estimation of utility values across subgroups of the population.
- They had good face validity with other estimates.
- Values were slightly higher than that observed in the Health Survey for England data. Higher utility values in the Irish population compared to the UK have previously been observed (258).
- They will allow for the control of co-morbidities.
- The allow regression techniques to be used.

We use utilities from the Health Survey for England data in scenario analysis given their national coverage of the English population and the depth of subgroup specific data reported (253).

Table 6.11 Utilities used in base case and scenario analysis for the economic model evaluating PCSK9 Inhibitors in Chapter 7

Utility Scenario	Purpose	Starting model value*	Age Decrement	MI Acute	Stroke Acute	MI Chronic		Stroke Chronic		Adjustments for multiple events
						Any MI	>1 MI	Any Stroke	>1 Stroke	
TILDA-1	a Base case – TILDA Reg	0.7778 Appendix Table A.13. Analysis of TILDA	Reg 2 Table 6.6	0.88 – m Ara et al (253).	0.77-m Ara et al (253).	-0.0191 Reg 5 Table 6.9	-0.0291 Reg 5 Table 6.9	-0.0994 Reg 5 Table 6.9	-0.01677 Reg 5 Table 6.9	Additively
	b No Age-Sex interaction		Reg 1 Table 6.6			-0.019 Reg 7 Table 6.9	-0.0275 Reg 7 Table 6.9	-0.0965 Reg 7 Table 6.9	-0.1676 Reg 7 Table 6.9	
	c No socio-economic for chronic disutility		Reg 2 Table 6.6			-0.0142 Reg 8 Table 6.9	-0.0365 Reg 8 Table 6.9	-0.1015 Reg 8 Table 6.9	-0.1635 Reg 8 Table 6.9	
	d No angina for chronic disutility					-0.0391 Reg 9 Table 6.9	-0.0423 Reg 9 Table 6.9	-0.1016 Reg 9 Table 6.9	-0.1654 Reg 9 Table 6.9	
TILDA 2	TILDA utilities applied through multipliers derived from mean values	0.7778 Appendix Table A.13. Analysis of TILDA	Reg 1 Table 6.6	0.88 – m Ara et al (253)..	0.77-m Ara et al (253)..	0.9303-m Table 6.10 and TILDA		0.8416 – m Table 6.10 and TILDA		Multiplicatively
3. Health Survey for England	Scenario analysis given TILDA utilities are derived from a mapping model	0.7052 Ara et al. (253). TIA multiplier was estimated from Luengo-Fernandez et al. (278)	No CVD equation Ara et al. (253)	0.88 – M Ara et al (253)..	0.77-M Ara et al (253)..	0.90365– m Multiplier calculated from Ara et al. (253).		0.8205m Multiplier calculated from Ara et al. (253).		Multiplicatively
4a. CALIBER – 1 year scenario	To compare with the original utility scenario in CALIBER	0.7289 (15,234)	-0.003	-0.062	-0.1171	No decrement		No decrement		Decrements applied for one year only
4b CALIBER - Constant scenario	To compare with the original utility scenario in CALIBER	0.7289 (15,234)	-0.003	-0.062 + -- .068	-0.1171	-0.0368		-0.1171		Additively

CVD, cardiovascular disease; m. multiplicatively, MI, myocardial infarction; Reg, Regression; TILDA, The Irish Longitudinal Study on Aging;

6.6 Discussion

6.6.1 Main findings

The difficulties of estimating utility values for use in economic models in CVD has been described. Standards of reporting of CV utilities in the literature has generally been poor and important considerations such as the application of age are not always implemented (269).

The application of the mapping model to national TILDA population has generated a patient dataset of predicted utility values for the community dwelling population in Ireland who are 50 years and older. The dataset was analysed to generate health state utility values to inform our economic model in CVD. The depth of the data available allows the heterogeneity in initial baseline utility values across 23 defined populations to be captured.

Unadjusted values had good face validity but were higher than that observed in a UK national study (253). Higher utility values in the Irish population compared to the UK have previously been observed (258,259). Accounting for this difference is important because results of economic evaluations can be sensitive to these parameters (171). However, the possibility that the difference is due to a bias in the mapping model used to derive them cannot be excluded. The baseline utility values were generally lower than those reported by a review of recent CV models (Table A.12) (269).

There was minimal difference in utility between the ages of 50 and 70. Ara et al previously observed a levelling off in the mean utility in the age 65 to 70 years. They hypothesised that it could be due to a relationship between utility and all-cause mortality (261). Another explanation could be due to the number of people retiring at these ages. A reduction in manual labour and an increase in time available for health promoting activities could theoretically stem some of the utility decline associated with age. The depth of data available in TILDA regarding employment history and retirement will allow the examination of this question in future research.

The results of Regression 2 are consistent with the results of Ara et al (253). On average, the utility of men is approximately 0.03 higher than women. The fact that men generally report higher utility than women is well established, but the cause is generally unknown. De Schemdt et al outlined some of the potential causes focusing on the area of coronary artery disease (279). Females may perceive symptoms differently or there may be physiological differences in the manifestation of disease. Higher depression rates in women is associated with poorer recovery rates (279).

The results of Regression 3 suggest that the decline associated with age differs between men and women, with a greater decline observed for women. An interaction between age and sex was previously observed in an study by Wu et al. examining the impact of 11 longstanding health conditions in a South Yorkshire cohort study (280). Such an interaction was not considered in any of the economic models in CVD reviewed by Ara et al (269). All else equal, Regression 3 predicts that a man aged 80 years and over would have a 0.0835 greater utility than a woman of the same age. This might reflect a survivor bias. Regardless of the cause, incorporating such a differential decline may have important implications for results of economic evaluations. This will be explored in the decision analytic model in the next chapter. Depending on its importance, this may lead to further debate regarding the ethics of subgroup analysis and decision-making based on sex.

The predicted utility decrement associated with multiple MIs is the MI decrement plus the '>1 MI' decrement. The decrement associated with more than one MI or stroke is greater than the decrement associated with a single event which indicates a synergistic effect. This could mean that there is less capacity for the body to compensate for the loss in health or it could indicate a more severe event. An alternative explanation is that the >1 MI decrement is capturing the effects of more than two events which makes the decrement greater.

There are few studies reporting population utility norms in Ireland. EQ-5D-5L population norms have been presented in the literature in the form of responses to the descriptive system (281). But the Irish EQ-5D-5L value set has not been applied to date. Information was also collected on many sociodemographic factors include religion, income, education in this data. Coverage is over the range of entire population and is not limited to those fifty years of age and over. However, no information is garnered regarding the respondent's own health apart from a 'yes – no' statement regarding self or family member experience of serious illness. Therefore, while further analysis of this dataset will inform baseline values for the Irish value EQ-5D-5L value set, it provides little information on providing utilities for CV events.

Detailed respondent-level datasets allow us to explore the complex relationship between variables and the EQ-5D-3L (282). Regression 3a shows that, after accounting for age and sex, MI is associated with a substantial utility decrement. We believe that that the loss of health after a MI is captured in the model through the coefficients for heart failure and angina (which account for the primary long-term sequelae after a MI). An understanding of the disease course is critical before deriving utilities for an economic evaluation (283). Controlling for heart failure in the regression would mean that some of the important consequences of MI would not be

captured in the coefficient. Respondent-level data allows us to generate utilities for subgroups of the modelled population and for respondents with joint health conditions.

There has been limited appreciation for the importance of controlling for confounders when estimating the disutility associated with CV events. This may be as due to the absence of the data required in many datasets. The depth of data in TILDA has allowed for their consideration here. It is important to note that baseline values for participants with a history of CV events and utility decrements associated with experiencing CV events are both required for economic evaluations. One of the most important implications of our mapping model is that, by applying it to further waves of TILDA, a longitudinal dataset of predicted utilities can be generated. This dataset would capture predicted utility before and after acute health events.

The MI and stroke decrements in each analysis represent average utility decrements relative to a no-CVD baseline. It is unclear if these decrements should be subtracted directly form the starting utility values in the economic model. There are too many variables to consider multiple interaction variables; it would lead to a combinatorial explosion. But as many co-morbidities are already controlled for (in the initial starting values) it is unclear if these decrements should be applied directly. Therefore, a conservative approach was taken here, and they were applied directly in our analysis.

6.6.2 Limitations and potential for future research

6.6.2.1 Accounting for changes over time

Since the Wave 1 data was collected in TILDA, there have improvements in the provision and organisation of care of CVD in Ireland. However, analysis of later waves of TILDA in future research will allow for an analysis of the changing in population health over time.

6.6.2.2 Systematic review

Ideally, a systematic review of the literature would be conducted to identify potential sources of utility values for an economic model. However, this was not conducted here. The depth of data in CVD and the number of health states required would make locating values from similar populations in comparable studies unrealistic. However, published systematic reviews have been examined in addition to more recent research. We are not aware of a detailed study of CV utility in Ireland using the EQ-5D-3L. While preferred, it is acknowledged that literature searches of utilities are not always necessary or feasible. But that the criteria used to choose studies should always be described (189).

6.6.2.3 Self-reported health conditions

The data captured in TILDA may affect the generalisability of the predicted utilities. Health conditions are self-reported. However, in a comparative study of information sources for healthcare utilisation in Ireland, the rate of healthcare utilisation from the primary care database was comparable to the self-reported group (284). This is explored more in the next chapter where the characteristics of the population with CVD in TILDA are compared to the CALIBER population used in the economic model and in other Irish research (15). No information is reported in TILDA regarding whether participants experienced haemorrhagic versus ischemic stroke. Therefore, differences in utility decrements between these could not be estimated here. In a UK study, fewer participants reported haemorrhagic stroke compared to ischemic stroke preluding meaningful comparisons between these events (278).

6.6.2.4 Recruitment and response bias

In Wave 1 of TILDA, participants were only sampled from the community dwelling population in Ireland. Further, participants who were not able to give informed consent were excluded. This means that participants with dementia or those in long-term care facilities are not accounted for in the sample. Their exclusion means that the utility decrements estimated here will be somewhat biased. However, participants recruited in Wave 1 who progress to these health states in future waves are not excluded. Therefore, this issue can be examined in future research.

Participant responses in TILDA were weighted for age, sex and education to account for differential rates of response, but we cannot exclude a health response bias (195).

6.6.2.5 Data availability

Only the publicly accessible version of the TILDA dataset was accessed for this study. For reasons of anonymity, many variables are modified or not included in this version. For example, age is top coded at 80 years. In addition, further information was collected in TILDA regarding the timing of CV events, and the forms of CVD that make up the "other heart disease" variable in TILDA. Access to this dataset would give more precise estimates than those presented here.

6.6.2.6 Apply to longitudinal data in TILDA

Only Wave 1 of TILDA was analysed here. Over time, the TILDA cohort would have experienced further CV events. Therefore, analysis of Waves 2 to 5 would increase the precision of the estimates as the number events increases. It would also provide further information on the trajectory of utility and HRQoL in an Irish context over time.

An important feature of TILDA is that from Wave 2 forward, TILDA has information both before and after participants have experienced CV events. This means that longitudinal analysis can attempt to quantify the causal predicted utility decrement associated with the event. This is an advantage compared to longitudinal studies which follow participants only after an initial event occurred.

6.7 Conclusion

Mapping TILDA to the EQ-5D-3L is a pragmatic method of generating further data to inform the quantification of the utility of the CVD population in Ireland. As highlighted in the conclusions of the previous chapter, the comparative strengths and limitations of data sources should be considered before using the mapping model to predict utilities for an economic model. The limitations of using predicted utility values was discussed extensively in the previous chapter. Despite these, the absence of alternative Irish data sources means that it is reasonable to include the TILDA estimate in the base-case economic evaluation given its comparative strengths versus other sources. Scenario analysis using Health Survey for England data is important given the uncertainty in the results.

In the next chapter, the health state utility values derived here will be used in a decision analytic model to quantify the cost-effectiveness of PCSK9 inhibitors in the Irish setting.

Chapter 7. Economic evaluation of PCSK9 inhibitors for the secondary prevention of CVD

7.1 Introduction

In order to conduct an economic evaluation of PCSK9 inhibitors in the Irish setting, an economic model is required to predict incremental changes in expected costs and outcomes associated with PCSK9 inhibitors compared with standard of care over a lifetime time horizon. To do this, robust estimates of the baseline CV risk are required. Given the disparate healthcare system in Ireland and limited use of electronic health records, there is limited availability of such data in the Irish setting.

In contrast, there are many sources of electronic data in the English setting. The CALIBER economic model is based on an analysis of patient-level data of over 94,966 patients in England between 2001 and 2010 (15). Given the lack of Irish data available to build a de novo economic model, this model represents the best source of evidence available to answer this research question.

The model code for the CALIBER model is freely available online at https://github.com/miqdadasaria/caliber-scad-model. Dr Miqdad Asaria, the corresponding author of the paper, gave permission to adapt the model for this thesis (15).

7.1.1 Aim

The aim of this chapter is to conduct an economic evaluation of PCSK9 inhibitors in Ireland for the secondary prevention population in Ireland. Given the uncertainty in the clinical evidence, two primary comparisons will be conducted:

- PCSK9 inhibitors versus standard of care.
- Alirocumab versus evolocumab versus standard of care.

Treatment effectiveness data for PCSK9 inhibitors is derived from the systematic review in Chapter 3. Irish utility data from Chapter 6 and other Irish evidence are used to populate the CALIBER model to meet the aims of this analysis.

7.1.2 Chapter outline

First, previous economic evaluations of PCSK9 inhibitors in the literature are described. The model developed by Asaria et al and related studies analysing the CALIBER population are presented (15). Notable data and methods used by Asaria et al are discussed (15). Next, the methods and data used to adapt the CALIBER model for the aims of this thesis are described. In the following section, the results of the economic evaluation are described. Differences in cost-effectiveness across subgroups of the secondary prevention population are analysed. Scenario

analyses are conducted to examine the robustness of the results to alternative assumptions. The chapter concludes with a discussion of findings.

7.2 Previous economic evaluations of PCSK9 inhibitors

A systematic review of cost-effectiveness evaluations was previously conducted by Korman et al (285). Ten studies were identified (286–294). Some of these were updates of other included cost-effectiveness analyses, to account for increased evidence after publication of the FOURIER economic outcomes trial. In the secondary prevention population, estimated ICERs ranged from €45,340 per QALY in an industry sponsored analysis to over US$1.3 million per QALY in an independent study. Analyses which applied direct treatment effects had lower ICERs than those who applied indirect estimates. None of the papers identified by Korman et al compared PCSK9 inhibitors to each other (285).

The papers identified by Korman et al were reviewed to examine how the ten studies accounted for subgroups in their analysis (285). Most studies did not examine cost-effectiveness across any subgroups across the secondary prevention population. A de novo evaluation by Korman et al looked at the impact of age, diabetes and statin intolerance. The impact of different LDL-C thresholds were not examined (289). Fonarow et al examined cost-effectiveness in US secondary prevention population but conducted a subgroup analysis using the FOURIER population (287). The Institute for Clinical and Economic Review examined the statin tolerant versus intolerant population through different LDL-C thresholds (294).

More recently, Wisloff et al modelled the cost-effectiveness of PCSK9 inhibitors across different subgroups in the secondary prevention HeFH population (164). However, the subgroups considered were simple including gender statin intolerance (modelled through LDL-C level) and age. They stated that PCSK9 inhibitors were not cost-effective in only one subgroup when treatment effects were estimated through the surrogate endpoint of LDL-C. Subgroups based on other clinical characteristics were not considered. They did not account for the fact that the percentage LDL-C reduction is dependent on the baseline LDL-C. Therefore, the cost-effectiveness is overestimated.

The studies to date have primarily examined PCSK9 inhibitors in the broad secondary prevention population. ICERs have generally been very high. No study identified simultaneously examined the impact of different forms of cardiovascular disease history such as MI, ischemic stroke or combinations of these. While PCSK9 inhibitors may not be cost-effective in the broad secondary prevention population, it may be possible to identify subgroups of the population in which they are cost-effective.

In the next section, the CALIBER economic model developed by Asaria et al is described. This model forms the basis of our own evaluation (15).

7.3 The CALIBER CV model

7.3.1 Population

The CALIBER population is a cohort of 94,966 patients with stable CAD in the UK. Patients were observed between January 2001 and March 2010 for a median follow up time of 4.2 years (inter quartile range: 1.9 to 6.9 years) (15). The population is a mixture of populations; a prevalent cohort who were eligible at the study start date in 2001 and an incident cohort who had a CV event during the study period. Patients were required to have no event in the 180 days post diagnosis of stable CAD. These patients entered the cohort at this point. The index date was the time of cohort entry.

A significant strength of the study is the number of data sources that were used to populate the dataset including primary care data from the Clinical Practice Research data link, hospital discharge records from the UK Hospital Episode Statistics, cause specific mortality from the UK Office of National Statistics and CV data from the Myocardial Ischaemia National Audit Project Registry (MINAP). Asaria et al highlighted the strength of this approach referencing a study by Herrett et al (15,295).

Asaria et al took the following approach to modelling the population (15). First, the baseline risk equations were applied to each of the 94,966 patients. Patients were then segregated into 10 groups according to their CV risk (15). Costs and outcomes were predicted at the mean baseline co-variate value across patients within each risk group. Costs and outcomes were also predicted for a representative patient from each group. Asaria et al justified the approach taken as there were non-linearities across patients characteristics and risk (15). For example, the five-year risk, when averaged across patents was 16.68%. But when the risk was estimated using the average co-variate values of the same population, the five-year risk was 11.64%.

7.3.2 Model structure

7.3.2.1 Multi-state model

Using the CALIBER data, Asaria et al developed a decision analytic model (which was based on previous models in CVD and expert clinical advice) (15). A simplified version of the model is presented in Figure 7.1. All patients entered the model in the stable CAD state. As a first event, patients can experience a non-fatal MI, non-fatal ischemic stroke, non-fatal haemorrhagic

stroke, a CVD death or a non-CVD death. Following a non-fatal event, only transitions to a fatal health state are permitted. The cycle length is 90 days. The time-horizon is lifetime.

There are multiple types of decision analytic models and there are often confusions in the terminology used to describe them. The CALIBER model is a multi-state model. This means that transitions between states are measured in continuous time. This differs from more familiar types of decision analytic model such as a state-transition model which model transitions between states in discrete time using transition cycle probabilities (296). Multi-state models differ from a partitioned survival model as all transitions are modelled explicitly. The CALIBER model uses parametric survival distributions to model transitions between states. As the median follow-up was 4.9 years, some extrapolation was required. There is growing interest in the use of multi-state models in the health economic literature (297). Competing risks models are a form of multi-state models where transitions are modelled from one state only.

Like traditional state-transition models, multi-state models may have the Markov property (297). This means that the probability of transferring to a future state depends only on the properties of the current state; previous history is irrelevant. In this case the model is semi-Markov. Following a non-fatal event state, patients transition through a series of tunnel states – one for each cycle in the model after the initial cycle. Therefore, while technically each state has the Markov property, the time since a non-fatal CV event state is recorded through the number of tunnel states. Because of this, the total number of health states in the model is proportional to the number of cycles. Asaria et al used a time-horizon of 70 years and there was 852 health states in the model (15). The use of the software environment R increases the feasibility of using complex model structures.

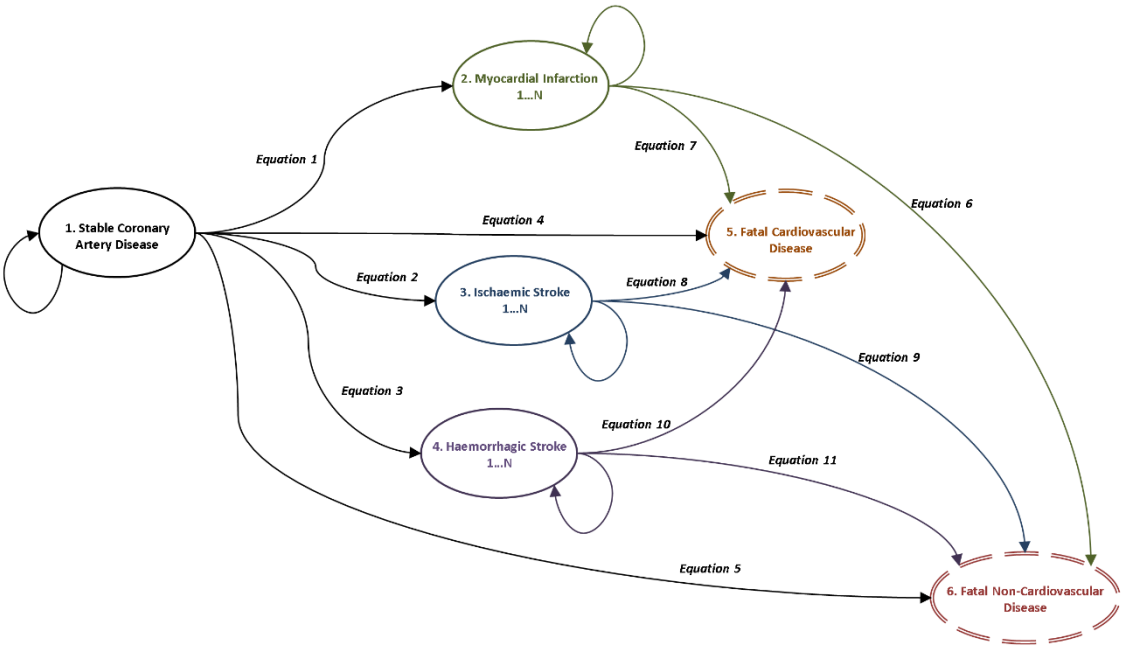


Figure 7.1 Economic model and risk equations adapted from Asaria et al. (15).

7.3.3 Baseline risk

7.3.3.1 Risk equations and parametric survival models

Rapsomaniki et al previously estimated CV risk in the CALBIER cohort (298). As Asaria et al built on these methods and used the same data, Rapsomaniki et al's work is described first (15,298).

A series of prognostic models for the CALIBER cohort were estimated for the endpoints of all-cause mortality and a composite endpoint of non-fatal MI or coronary death (298). Exponential proportional hazard models were used to measure the effect of co-variates. A large number of co-variates were incorporated, including sociodemographic characteristics (age, sex and social deprivation), CAD subtype, CV clinical history and risk factors (use of nitrates, history of revascularisation, previous MI , hypertension, lipids, heart failure, atrial fibrillation, smoking) and non-CVD co-morbidities (chronic renal disease, chronic obstructive pulmonary disease, cancer, chronic liver disease), mental health (depression, anxiety) as well as clinically assessed biomarkers (heart rate, white cell count, haemoglobin, creatinine). Rapsomaniki et al stated that the candidate predictors were drawn on recommendations in recent guidelines for the management of stable CAD (298). The models were internally and externally validated (298).

Rapsomaniki et al estimated that the Kaplan-Meier 5-year mortality rate was 20.6% (95% CI 20.3 to 20.9%) (298). Risk differed by CAD subtype with the highest risk in patients with NSTEMI and the lowest risk in patients with stable angina. The associations between risk and every

prognostic factor are reported by Rapsomaniki et al (298). It is important to note that the models represent prognostic rather than causal associations. For example, nitrates are a class of drug used to treat symptomatic angina. Rapsomaniki et al found that use of nitrates was associated with a 40.5% increase in risk of non-fatal MI or coronary death (298). However, this reflects the fact that the use of nitrates is associated with the severity of angina rather than a causal effect of the drug itself.

Building on this work, Asaria et al developed baseline risk equations for each model transition in the CALBIER economic model (15). Their approach differed from Rapsomaniki et al given the need to need to model specific transitions and the need to extrapolate to the model time horizon (15,298). The size of the population in the Asaria et al model was approximately 10% lower than that of the Rapsomaniki et al population (15,298). This is because Asaria et al did not use data from the undefined MI cohort in the derivation of their risk equations (15).

For each of the 11 possible model transitions presented in Figure 7.1, risk equations were estimated using parametric survival models. The co-variates chosen by Asaria et al, to adjust the baseline risk, mirrored those in the Rapsomaniki et al analysis (15,298). The parameters for each transition (as estimated by Asaria et al (15)) are reproduced in Supplementary Appendix 1.4.4. (15). Transitions to the non-fatal haemorrhagic stroke state and from any of the non-fatal event states only included age and sex as co-variates. Because of the fewer number of events observed, there was insufficient power to include more co-variates for these models. Four parametric survival distributions were assessed for each transition – generalised gamma, lognormal, Weibull and exponential. The parameters in each table represent acceleration factors. This means that if the median cause specific survival was 'y' years for someone without the risk factor, the median cause specific survival for someone with the risk factor is the coefficient multiplied by 'y' years. Therefore, coefficients less than one represent factors which increase the hazard for the event.

The strength of the risk equations comes from the large sample size and the long duration of follow-up. They incorporate data from 4,179 MI events, 3,222 ischemic strokes, 262 haemorrhagic strokes, 5,536 deaths attributable to CVD and 8,663 deaths attributable to non-CV causes (15,298).

Cohort mean predictions for cause specific survival model were plotted versus mean Kaplan-Meier survival estimates. This allowed the visual comparison of observed event rates to those predicted; it informed the assessment of the clinical plausibility of the estimates. Asaria et al used the AIC to assess the goodness of fit of the four different parametric distributions to the

data (15). The model with the lowest AIC was deemed to have the best fit relative to the other models examined and included in the economic model. However, this is not always appropriate as methods used to combine the hazards have implications for the assessment of model fit. In the next section competing risk methods are described. The implications of competing risk for the assessment of model fit is described Section 7.3.3.3.

7.3.3.2 Competing risk

7.3.3.2.1 Why accounting for competing risks is necessary

In the CALIBER model, transition hazards from each health state are appropriately combined in a competing risk framework as described by Putter et al (299). Williams highlights that there is often confusion in the literature about what is meant by competing risk (297). A summary of what is meant by competing risk and why it should be accounted for is presented below based on explanations provided by Williams and Putter et al (297,299).

The standard approach to survival analysis is the Kaplan-Meier approach where patients are censored if they do not have the event in the time observed. However, this method is only reliable where patients are at risk of one event only, for example all-cause death or a composite endpoint of non-fatal MI or death. Kaplan-Meier methods are biased in the presence of competing risks. Williams defines competing risks such that a competing risk prevents the event of interest occurring first (297). A common example used is the competing risks of multiple causes of death – when assessing the risk of CV death, non-CV death is a competing event. All patients are at risk of both events but only one will ever be observed. But death is not the only form of competing risk. In the CALIBER model, there are five possible transitions from the initial stable CAD health state. When quantifying the risks for a single transition, the remaining four transitions are competing risks for the transition of interest.

The method, where a competing risk is merely treated as censored at the time of the event, is called the naïve Kaplan-Meier analysis. This can lead to biased results (299). This is because it assumes that upon removal of one cause of failure, the risks of the remaining causes remain unchanged. Intuitively this is false. For example, if it were possible to completely remove the risk of CV death, more people would die from non-CV causes and the probability of non-CV death would increase. When using estimates derived from naïve Kaplan-Meier analysis, the estimated probabilities of CV and non-CV death could sum to more than one, which is clearly impossible. Even when such a critical failure does not occur, the risk of death is still overestimated. The risk of bias increases with increases in the hazard of the competing risk (299).

The focus in competing risks analysis is the derivation of the cumulative incidence function. This is defined as the cumulative probability of the event having occurred in the presence of competing events (299). The cumulative incidence function is preferred in a competing risks scenario as common assumptions may fail in a competing risk scenario.

For example, when conducting survival analysis, there is always a 1:1 correspondence between the hazard and the survival function. In other words, if the hazard decreases, the proportion surviving will increase. But this is not always the case in competing risk scenarios (297). For example, consider a scenario where a drug reduces the hazard of non-fatal MI but there is a competing risk of non-fatal stroke. Contrary to expectations, the cumulative incidence of non-fatal MI could theoretically still increase in certain scenarios. Theoretically, this can occur because the cumulative incidence of non-fatal MI depends not only on the hazard of non-fatal MI but also on the hazard of non-fatal stroke. If the drug also has an effect on the hazard of non-fatal stroke, the hazard of the competing risk of stroke could fall sufficiently such that more people are at risk of non-fatal MI (increasing the cumulative incidence of non-fatal MI). These examples show the importance of accounting for competing risks when modelling transitions from health states.

7.3.3.2.2 Statistical methods for competing risks

The statistical considerations, of how to account for competing risks, are outlined below. In a scenario where there are K competing risks, the cumulative incidence of one of these events k can be obtained by first calculating the cumulative cause specific hazard to time t and the overall survival function $S(t)$ for the risk of any of the K events. In the CALIBER model, there are five competing risks from stable CAD health state (K = 5): non-fatal MI, non-fatal ischemic stroke, non-fatal haemorrhagic stroke, CVD death and non-CVD death.

For each event k, the cumulative cause specific hazard is:

$$\Lambda_k(t) = \int_0^t \lambda_k(s)ds$$

Where $\lambda_k(t)$ represents the cause specific hazard at time t. When estimating the cause specific hazard, patients experiencing competing events are censored at the time of the competing event.

The overall survival function $S(t)$ for the risk of any of the K events is:

$$S(t) = \exp\left(-\sum_{k=1}^{K} \Lambda_k(t)\right)$$

In this equation, $\sum_{k=1}^{K} \Lambda_k(t)$ represents the sum of individual hazard functions k in K.

Using the results of these two equations, the cause specific cumulative incidence function $\mathrm{I}_k(t)$ for k can be determined by:

$$\mathrm{I}_k(t) = \int_0^t \lambda_k(s) S(s)\, ds$$

After defining the cumulative incidence function for each transition, Asaria et al (15) calculate the probability of transition from the stable CAD health state to each event state k in K for each cycle at time t to t +1 by

$$\frac{\mathrm{I}_k(t_{+1}) - \mathrm{I}_k(t)}{SCAD_t}$$

Where $SCAD_t$represents the probability of being in the stable CAD health state at time t and k represents the risk of any of the K events. Similar methods are used for transitions from non-fatal event states to death where the modelled population faces competing risks of CVD death and non-CV death.

7.3.3.3 Choice of survival distribution in the presence of competing risks

Combining hazards in a competing risk framework has implications for assessing model fit. The choice of distribution is more complex in a competing risk scenario. This is because the choice of distribution for the transition from the stable CAD health state, to any one of the five possible states, also affects the cumulative incidence of transitions from the stable CAD health state to the four states (each representing the remaining competing risks) (35).

Williams et al state that the term survival is usually reserved for survival - free of any competing risks (35). This is also described as overall survival. This is the survival function which is directly observed and is a function of all the competing risk hazard functions. The second form of survival function is the cumulative cause specific survival from a single event. This can be derived from the cause specific hazard using standard equations. However, this is only of theoretical relevance as it is based on an unrealistic scenario where competing risks do not occur.

In the Supplementary Appendix 1.4.4 the predicted cumulative cause specific survival functions are plotted versus naïve Kaplan-Meier plots. However as stated in 7.3.3.2 naïve Kaplan-Meier plots are biased in the case of competing risks. Because of this, they cannot be used to judge

the goodness of fit, as the Kaplan-Meier plots and predicted cumulative cause specific survival functions are not measuring the same thing.

Because the hazard is not affected by competing risks, the plots of the predicted hazard functions versus the observed piecewise exponential functions (also plotted by Asaria et al) are more relevant in competing risks scenarios (15). However, Williams et al state that, the choice of distribution for one event transition affects the observed probabilities for all the other transitions, because of competing risks (35). Thus, distributions should not be selected independently of the selection of the survival distribution for other transitions (35). Instead, selection for all transitions should occur at the same time. Candidate function for each transition should be plotted and combined in a competing risk framework to calculate the cumulative incidence of each event. Then the plotted cumulative incidence for each event should be compared to the observed data. Williams et al state that a visual assessment of relevant plots should be conducted, and distributions should be chosen based on a balance of good fit to the observed data and a reasonable extrapolation (35). However, this analysis is not reported by Asaria et al (15). Because the raw population data is not released for reasons of anonymity, it is not possible to conduct this analysis for this thesis.

These issues are important because Williams et al believe that choosing a distribution based on AIC is not appropriate in a competing risk scenario where the underlying hazards are modelled (35). Williams et al argue that comparing AICs of models for hazards of individual transitions does not correctly correspond to the state occupancy probabilities which are ultimately of interest (35,297).

These issues increase uncertainty in the baseline risk estimates applied in the model.

7.3.3.4 Non -CV death

As for CV risk, parametric survival models are estimated from the patient-level data in the CALIBER datasets and combined in a competing risks framework. After ten years, mortality is estimated using UK lifetables.

7.3.4 Costs

As a secondary prevention population, all patients in the model already incur significant ongoing costs associated with managing their condition. The costs applied by Asaria et al were based on previous analysis of the CALIBER population by Walker et al (15,300). Primary care, inpatient stays and diagnostic tests and prescriptions costs were included. Outpatient costs were not included because of a lack of data. Costs were presented as total costs, CVD costs or CAD costs.

A major strength of the Walker et al analysis is that the impact of baseline co-variates on costs were assessed by including them in a regression model (300). This means that the heterogeneity in costs across the modelled population can be accounted for. Co-variates were based on those used to derive the baseline risk equations in Section 7.4.6.2. Cost data is often right skewed. Therefore, a generalised linear model with a log link and gamma distribution were used to account for the skewness and the non-linear impact of co-variates. Panel data methods were used with time-invariant co-variates to estimate costs over each 90-day cycle. Co-variates were included to account for the impact of CV events on costs in the 90-day cycle in which they occurred and over in the long term following an event.

However, while, Asaria et al analysed the same data (15), the methods were modified slightly. Asaria et al used a linear regression model for his analysis as opposed to the generalised linear model with the log link used by Walker et al. (15,300). Further, Asaria et al included a smaller range of coefficients compared to Walker et al (15,300). Asaria et al state that aggregate costs were also estimated using a generalised linear model with a log link function but gave similar model results (giving reassurance regarding the appropriateness) (15). However, it is unclear what distribution was used to model the residuals in both analyses. The reason for the change in approach is not stated.

To account for CV event costs, Asaria et al included a dummy variable for the period in which the event occurred (cycle 1) and in cycles 2 to 4 following the event to capture marginal costs that occur later that year (15). A coefficient was also included to capture the ongoing costs associated with having a history of the event (15).

7.3.5 Utilities

The CALIBER model was populated with health state utility values from Sullivan et al (234). Two scenarios were considered:

1. Utility decrements after a CV event are constant
2. Utility decrements after a CV event applied for one year only.

The utility values and an appraisal of the methods used to derive them is presented in Section 6.4.1 of Chapter 6.

7.3.6 Treatment effect scenarios

The population in the decision analytic model are expected to undergo non-fatal MI, non-fatal ischemic stroke, non-fatal haemorrhagic stroke and CV death and non-CV death. The treatment benefits, of the interventions and comparators under consideration, are applied by reducing the

risk of these events. Asaria et al did not specifically examine the effect of PCSK9 inhibitors or any other licensed treatment for CVD in their model (15). Instead a series of hypothetical interventions were examined with different reductions in event rates. Four additional scenarios modelled indicative treatments that each reduced the risk of CV events by 10%, 20%, 30% and 40% (haemorrhagic stroke and non-CV death hazards were modelled to be unaffected by treatment). From this information, the maximum price that could be charged for new treatment in each risk group was calculated over a range of cost-effectiveness thresholds.

7.3.7 Results

An important point highlighted by Asaria et al is that, across all risk groups, the percentage of patients experiencing non-CV death was 94.46% in the lowest risk group but was still 54.05% in the highest risk group (15). Therefore, even if a drug substantially reduces the risk of CV death, patients may have limited capacity to benefit because of the competing risk of non-CV death (15). The results showed that even for the highest risk cohort who had a five year event risk of 44.2%, the maximum annual prices that could be achieved for a new intervention that reduced CV risk by 40% was £1,269 at a cost-effectiveness threshold of £20,000 per QALY. These prices are far lower than current annual list price of evolocumab of €6,331 excluding VAT (See Section 7.4.8.4)

Asaria et al acknowledge several limitations in their analysis (15). HRQoL was not recorded in the CALIBER dataset. Prognostic risk factors over time were not explicitly modelled. Further, only first events were explicitly modelled. Past event rates may not reflect future event rates.

7.3.8 Transferability of the economic model

The CALIBER economic model uses electronic health records to predict costs and outcomes in patients with stable CAD. The depth of data available means that the baseline risk and costs can be adjusted to account for the heterogeneity in the modelled population. Further, the use of survival distributions means that the complex changes in event rates over time can be incorporated into the model. The lack of accurate secondary-prevention event risk data in Ireland means that this model represents the best source of risk data available to model CV risk in Ireland. However, there are limitations. The utility data from Sullivan et al applied in the model may not reflect levels of health in Ireland (234). Further, patients with a history of TIA or stroke without a history of CAD are not included in the model. While a probabilistic sensitivity analysis was conducted, other forms of deterministic sensitivity are not reported. Asaria et al presents results by risk group (15), but such complex criteria are unsuitable to classify subgroups for reimbursement in the real world setting. There are also limitations which affect

transferability to the Irish setting including lack of Irish specific data. Further adaptations are required before the model can be used to estimate the cost-effectiveness of PCSK9 inhibitors in the secondary prevention population in Ireland.

7.4 Development of the CALIBER economic model for the economic evaluation of PCSK9 inhibitors in Ireland.

7.4.1 Irish reference case

The reference case described in national guidelines (for the economic evaluation of health technologies) informed the framework of the economic analysis (8). A cost utility analysis was conducted under the perspective of the healthcare payer in Ireland versus routine care. Incremental costs and utilities were estimated over 40 years to ensure a lifetime time horizon. Costs and utilities were discounted at a rate of 4%. Treatment effectiveness was informed by a systematic review of the clinical evidence.

7.4.2 Population

The licensed population for PCSK9 inhibitors is very broad. It includes both primary and secondary prevention of CVD in patents who are unable to reach LDL-C goals. LDL-C goals are undefined in the product license. The focus of this thesis is on the secondary prevention population only. No CV outcomes trials have been conducted in the primary prevention population.

The secondary prevention population is in itself very broad. One way of classifying CVD is by the vascular bed effected: coronary (heart), cerebrovascular (brain), peripheral (limbs and other arteries). Within each vascular system again, there are various manifestations and severity of disease including:

- CAD
    - stable angina
    - unstable angina
    - MI
        - NSTEMI
        - STEMI
    - history of revascularisation
        - percutaneous coronary intervention
        - coronary artery bypass graft (CABG)
    - other coronary artery disease

- cerebrovascular disease[11]
    - ischemic stroke
    - transient ischemic attack (TIA)
- PAD.

The secondary population can also be segregated based on:

- history of recurrent events
- disease in more than one vascular bed
- presence of co-morbidities such as diabetes
- risk factors including LDL-C.

The aim of this analysis is to quantify the cost-effectiveness of PCSK9 inhibitors in Ireland. Also, to identify subgroups of the population in which PCSK9 inhibitors may be cost-effective. This work will inform reimbursement recommendations and international treatment guidelines. Given the large number of classification factors, the potential number of subgroups is substantial. It is not feasible to model every possible subgroup. Therefore, an iterative approach was adopted. Based on the factors above and data availability, twenty-three populations were initially defined. These were previously defined in Table A.13. Where PCSK9 inhibitors were found to be cost-effective in a subgroup, it was planned that this subgroup would be re-investigated to examine whether the subgroup should be reclassified further. After further reclassification, subgroups, in which PCSK9 inhibitors were deemed to be cost-effective, would be removed from further analyses. This would ensure that any conclusions regarding differences in cost-effectiveness across the licensed population were based on mutually exclusive subgroups. This important adjustment is feasible because of the extent of the patient-level data in TILDA. Reclassification of subgroups by age or sex was not conducted given ethical concerns. However, analyses are presented to inform the audience regarding differences in cost-effectiveness across these subgroups.

Average population characteristics for each of these populations, including previous history of CVD and sociodemographic factors, were estimated directly from TILDA where data was available. Average CALIBER data was used where the required data was unavailable from TILDA.

[11]Haemorrhagic stroke is also a form of cerebrovascular disease. However, LDL-C lowering therapy is not indicated following haemorrhagic stroke

Baseline LDL-C has been used as reimbursement criteria in many countries including Ireland. When estimating cost-effectiveness at a given LDL-C threshold, previous UK analysis have modelled the average baseline LDL-C for patients above this threshold (33). However, this is inappropriate. The LDL-C threshold set should reflect the marginal LDL-C threshold at which PCSK9 inhibitors become cost-effective. Under this principle, the modelled baseline LDL-C in each analysis is the LDL-C treatment threshold considered.

PCSK9 inhibitors are recommended in European guidelines for patients with an LDL-C greater than 1.4mmol/L. However, achieving cost-effectiveness at such low thresholds is challenging. Therefore, LDL-C treatment threshold for the basecase evaluation here is 4mmol/L. This is the current treatment threshold for patients with a history of MI and/or CABG in Ireland. Scenario analysis are conducted at other baseline LDL-C treatment thresholds.

Pragmatically, it was assumed that that population characteristics did not differ across LDL-C treatment thresholds.

7.4.3 Intervention

The results of Chapter 3 show that PCSK9 inhibitors reduce the risk of non-fatal CV events across subgroups. However, the consistency in the magnitude of the clinical benefit across groups is unclear. Given the uncertainty in the clinical evidence base, two analyses are conducted.

In analysis one, the intervention modelled is the PCSK9 inhibitor class. The purpose of this scenario is to harness the whole of the evidence base by applying evidence derived from meta-analyses of trials pertaining to PCSK9 inhibitor class versus standard of care (SOC).

In the second scenario, evolocumab and alirocumab are modelled separately and compared to each other as well as SOC. This scenario is uncertain, given the number of uncertain assumptions underlying the adjusted indirect comparison.

For evolocumab, a dose of 140mg every two weeks subcutaneously is assumed. A regimen of 420mg subcutaneously once monthly is also available; this is considered therapeutically similar to former regimen. It is 50% more expensive and thus it is not considered further here.

For alirocumab, doses of 75mg or 150mg should be administered subcutaneously every two weeks. Patients may start at a dose of 75mg every week and titrate up as required. However, as our starting baseline LDL-C treatment threshold is high at least 4mmol/L, it is assumed that all patients start treatment at the higher dose of 150mg every two weeks.

7.4.4 Comparators

As described in Chapter 2 the appropriate place in therapy of PCSK9 inhibitors is in the third line setting in combination therapy with both maximum tolerated statins and ezetimibe. In this scenario, statins and ezetimibe represent the SOC. Therefore, the appropriate comparison for the economic evaluation is PCSK9 inhibitors + SOC versus SOC alone.

It is assumed that the effects of statins and ezetimibe are already incorporated into the baseline LDL-C estimates. Thus, no further adjustments are made to the baseline risk for these drugs.

7.4.5 Model structure

No adaptations were made to the model structure of the CALIBER model.

7.4.6 Baseline risk

The baseline risk is expected to be a primary driver of the model but there is a lot of uncertainty associated with it. The long-term follow up of CALIBER is a strength but it also means that the risk estimated may not reflect current CV risk. Changes in therapeutic options over time and changing demographics means that the baseline risks are a factor here. There is also uncertainty associated with the modelling approach. In the CALIBER population the average 5-year risk (when averaged across patients) was 43% higher than the average 5-year risk measured as the average of the population's covariate values (15).

7.4.6.1 Choice of survival distribution

The selection of the most appropriate survival distribution is more difficult in a competing risk scenario given the joint effect each distribution has on the cumulative incidence for each hazard. Given the uncertainty, scenario analysis was conducted using alternative distributions.

7.4.6.2 Populating the risk equations

As described in Section 7.4.2, the risk equations were populated with data from TILDA. Where data was unavailable on a co-variate, average values from CALIBER were used.

7.4.6.3 Stroke and TIA risk adjustments

The baseline risks in the model were derived from the CALIBER population which only included patients with stable CAD. Therefore, while patients with a history of both CAD and cerebrovascular disease were included in the CALIBER analysis, patients with only a history of cerebrovascular disease only are not accounted for. The coefficient for stroke in the baseline risk represents the additional risk associated with stroke on top of the CAD risk.

Therefore, several assumptions are required to incorporate the risk of patients with a history of cerebrovascular disease only into the model. It is assumed that:

- The CV risk after a TIA is equivalent to that after stroke.
- The effect of co-variates on baseline risk in cerebrovascular disease is the same as CAD.
- Survival distributions applicable to the CAD population are also transferable to the cerebrovascular disease only population.

The baseline risks can be adjusted downwards to account for the fact that a proportion of patients in the model will have a history of cerebrovascular disease only. Published hazard ratios from the Oxford Vascular Study (which quantify the increase in risk associated with disease in two vascular beds) are used to adjust the hazard for the proportion reporting a cerebrovascular event only weighted by the proportion of patients reporting any history of CVD (301). The hazard ratios used to perform the adjustment are presented in Table 7.1.

Table 7.1 Adjusted hazard ratios for 5-year risk of CV events for multiple versus single vascular beds adjusted for age and sex from the Oxford Vascular Study (301)

	Triple	**Triple/Double**	**Double #**
Number of Patients	104	712	608
	HR (95% CI)	**HR (95% CI)**	**HR**
Vascular Death	3.17 (1.97-5.10	1.52 (1.14 2.03)	1.24
Recurrent Ischemic Stroke	2.09 (1.19-3.66)	1.24 (0.91-1.68)	1.09
Recurrent non stroke acute vascular event *	5.28 (2.97-9.39)	3.34 (0.91-1.58)	3. 00

This was not reported in the study but was estimated using the data in the previous two columns.
*defined as myocardial infarction, acute peripheral vascular event or sudden cardiac death
CI, Confidence Interval; HR, Hazard Ratio.

7.4.6.4 Non-CVD mortality

The non-CV morality data for the post 10-years predictions were updated for Irish data. The annual probability of all cause-death by age and sex was estimated from the most recent version of the Irish life lifetables (302) and converted to a rate using the formula:

$$r = -\frac{1}{t}\ln(1-p)$$

-where r is the rate, t is time and p represents the probability.

The annual rate of CV death by age and sex for the corresponding years (2010 – 2012) was obtained as rates “Any circulatory system death” from Irish mortality data published by the Central Statistics Office (CSO) (303). The annual rate of non-CV death was estimated by subtracting the rate of CV death from the annual rate of all cause death from Irish life tables

estimated above. The daily non-CV mortality hazard was calculated by dividing the annual rate for each by 365 and multiplied by 90 to calculate the cumulative hazard for each 90-day cycle.

7.4.6.5 Scenario analyses

In scenario analyses, the hazards for each of the CV transitions were varied separately ±20%. A further scenario analysis was performed where all CV event transitions were varied jointly by ± 20% and ±50%. Scenario analysis was also conducted by substituting the parametric distributions for the non-fatal MI and first event CV death to the exponential distribution and for the first event non-CV death to the lognormal distribution. Further, changes in ICER, when estimating non-CV mortality using English lifetables, was explored.

Survival parameters were varied in the probabilistic sensitivity analysis (PSA) from the normal distribution.

7.4.7 Adverse events

Asaria et al assumed that the indicative treatments had no side effects (15). The results of our systematic review and meta-analysis in Chapter 3, indicated that there was no difference in serious adverse events observed between PCSK9 inhibitors and SOC. While, trials do indicate a slight increase in injection site reactions, it is assumed that there is no marginal increase in resource use and that such reactions are accounted for in routine follow-up.

7.4.8 Resource use and costs

7.4.8.1 Baseline costs

There is no Irish data source that is comparable to the depth and wealth of the CALIBER analysis for background costs. Therefore, data from Asaria et al's CALIBER analysis form the basis of cost parameters in the adapted model (15).

Asaria et al reported costs in 2011/2012 UK pound sterling (15). Therefore, costs were converted to 2019 euro using methods defined in national guidelines (8). Costs were inflated to 2019 values using the UK consumer price index for health and converted to euro using purchasing power parities. The base year was an average of 2011 and 2012 values. The adjusted parameters used in the model and an example of how they are computed to estimate costs is presented in the Supplementary Appendix 1.4.1 and Table A.14.

Outpatient costs were not included in the analysis by Asaria et al (15). Therefore, the cost of an annual outpatient consultant visit was applied pro-rata to each model cycle.

As described in Section 7.3.4, the costs were estimated from a population with CAD. While a proportion of this population have a history of cerebrovascular disease, the baseline costs of a population with cerebrovascular disease alone are not captured. However, the costs associated with the proportion of this population who subsequently have a stroke are accurately quantified. The costs from the population with CAD were considered transferable to the population with stroke alone. No further adjustment was made to the model. A coefficient for a baseline history of stroke was not included in the regression analysis by Asaria et al; this supports this assumption (15).

In the PSA, costs were sampled from the gamma distribution.

7.4.8.2 CV event costs

To incorporate Irish specific data, costs for acute MI and stroke from Healthcare Pricing Office Activity Based Price list were applied to the first cycle following the event (304). The Diagnosis Related Groups used to compute these costs are reported in Table A.15. As these Diagnosis Related Groups do not differentiate by type of stroke, the same costs were assumed for both haemorrhagic and ischemic stroke. No recent Irish data was located for event costs associated with cycles 2 to 4 or later. Therefore, the marginal costs associated with an event for cycles 2 to 4 were applied in the original CALIBER model after adjusting for conversion and inflation (15).

7.4.8.3 Non-drug costs associated with PCSK9 inhibitors use

To account for the extra monitoring at treatment initiation, the costs associated with additional out-patient consultant appointments were applied at treatment initiation (cycle 1) and at 3-month follow-up (cycle 2). This is assumed to cost €130 (from the HSE 2013 Ready Reckoner) which is equivalent to €136.50 in 2019 euro (305). Patients require training in order to self-inject. It is assumed that the cost of nurse training is equivalent to an additional outpatient cost of €136.50 given that this would also take place in the outpatient setting.

Going forward, it assumed that there are no additional resource implications beyond SOC.

7.4.8.4 Drug costs

Drug costs for PCSK9 inhibitors are presented in Table 7.1. Drug costs were calculated in reference to the NCPE Guidelines for the calculation of Drug Costs (306). In line with national guidelines for economic evaluation, 23% VAT was not included in the cost-effectiveness model (8). Costs including VAT are presented for information purposes only.

In cycles where a fatal event occurs, the event is assumed to occur mid-cycle and 50% of the estimated drug cost is applied.

Drug costs for statins, ezetimibe and other pharmacological therapies for CVD are already incorporated into the baseline costs in the model.

Table 7.2 Evolocumab and alirocumab drug costs

	Evolocumab		Alirocumab	
	Excl. VAT	Incl. VAT	Excl. VAT	Incl. VAT
90-day Cost	€1,633.82	€1,985.31	€1,666.93	€2,026.44
Annual Cost	€6,330.57	€8,057.04	€6,764.95	€8,223.98

Excl. Excluding; Incl. Including; VAT, Value Added Tax. VAT is not included in the cost-effectiveness model and costs including VAT are presented for comparative purposes only.

7.4.9 Utilities

The utilities applied by Asaria et al and the considerations and methods used to derive utilities representative of the Irish population were described in the previous chapter (15). A summary of the utility values applied in each scenario is presented in Table 6.11.

All scenarios use health state multipliers derived from Ara et al to adjust utilities for acute events (which included patients with a history of the event within the previous year) (253). The acute multipliers are applied for four cycles (equivalent to 360 days). Therefore, while utility will vary over the year following the event, it is assumed that by applying the utility decrement over one year, the total average utility decrement is accounted for.

Utilities were varied in the PSA. The distribution chosen depended on the type of utility value. Initial utility values and utility multipliers were sampled from the beta distribution. Utility decrements were independently sampled from the normal distribution.

7.4.10 Treatment effectiveness

7.4.10.1 Clinical evidence

In line with national guidelines, a systematic review was conducted in order to identify high quality randomised controlled trials examining the efficacy of PCSK9 inhibitors (8). The results of this review are described in Chapter 3. PCSK9 inhibitors were shown to reduce the risk of non-fatal MI and non-fatal ischemic stroke but there was no observed effect on CV mortality. No statistically significant difference in treatment effect was observed between evolocumab and alirocumab.

7.4.10.2 Application of the treatment effectiveness in the economic model

There are two possible methods of applying the treatment effect – direct and indirect. The direct method applies hazard ratios derived from the meta-analysis directly to the corresponding event hazards in the model.

Under the indirect method, hazard ratios of PCSK9 inhibitors versus SOC are derived indirectly through their ability to reduce the surrogate endpoint of LDL-C. There is no direct clinical evidence for the subgroup of patients with an LDL-C of 4mmol/L. Therefore, application of the indirect method is required to investigate cost-effectiveness under the assumption that the relative treatment effect of PCSK9 inhibitors is greater at higher levels of baseline LDL-C.

To translate the effect of PCSK9 inhibitors on the surrogate endpoint of LDL-C to hazard ratios for the risk of CV events, the following method is used:

1. The percentage LDL-C reduction associated with PCSK9 inhibitors is applied to the baseline LDL-C of the modelled population to calculate the absolute LDL-C reduction

$$Absolute\ LDL-C\ reduction\ =\ Baseline\ LDL-C\ \times Percentage\ Reduction$$

2. The absolute LDL-C reduction is used to adjust the event specific rate ratios per 1 mmol/L (published in the literature) for the modelled absolute reduction in LDL-C using the following formula:

$$Adjusted\ Hazard\ Ratios\ =\ Event\ Specific\ Rate\ Ratio^{Absolute\ LDL-C\ Reduction}$$

The event specific rate ratios for each event type are presented in Table 7.3. They are obtained from analyses published by the Cholesterol Treatment Trialists Collaboration (CTTC). The CTTC was established in 1994 because of the recognition that no single lipid-intervention trial would have sufficient statistical power to reliably assess mortality outcomes or perform multiple subgroup analysis (307). Several meta-analyses, using individual patient-level data from cholesterol lowering trials, have been published (41,162,163,308,309). Albeit, only trials examining statins have been included in analyses published to date.

Several limitations to the CTTC rate ratios should be noted:

1. While both statins and PCSK9 inhibitors reduce LDL-C, they belong to different drug classes and it may not be appropriate to assume the same treatment effect.
2. The CTTC rate ratios are derived from statin trials conducted over 20 years ago in heterogenous populations. Given the improvements in CV treatments in the interim, the rate ratios derived may not be relevant today.
3. In the trials, from which the CTTC treatment effects were derived, statins reduce CV mortality in two ways. The first is through a reduced risk of CV events (which avoids the

increased risk of CVD death after an event). The second is through preventing CVD mortality directly. As the prevention of non-fatal CV events is already incorporated into the model, applying the CTTC rate ratio may overestimate the applied treatment effect.

The application of the indirect treatment effects in the model should be considered as an optimistic but plausible assumption regarding treatment effectiveness. Assumptions are required in the absence of long-term clinical evidence.

7.4.10.2.1 Hazard ratios

Given the fewer assumptions required, direct application of hazard ratios is preferred. In the base-case, hazard ratios derived from the meta-analysis presented in Table 3.4 of Chapter 3 are applied directly to the event hazards for the comparison of PCSK9 inhibitors versus SOC. The results of the adjusted indirect comparison in Table 3.7 are applied for the comparison of evolocumab versus alirocumab.

Two scenario analysis are conducted, for the comparison of PCSK9 inhibitor versus SOC, using treatment effects derived using the indirect method.

- Scenario 1 uses LDL-C reductions attained in the subgroup in the ODYSSEY LONG TERM trial who had a baseline LDL-C level greater than 4.14 mmol/L. No other published economic evaluation of PCSK9 inhibitors has accounted for the difference in the LDL-C percentage lowering ability at this LDL-C threshold.
- Scenario 2 uses LDL-C reductions from the total population of the ODYSSEY LONG TERM trials rather than subgroups specific estimate.

LDL-C reductions from the evolocumab trials are not included here, given that no relevant trial published results by baseline LDL-C subgroup. Identification of the difference in treatment effect by subgroup was previously described in Chapter 3.

Table 7.3 Hazard ratios applied in base-case and scenario analysis for estimating the effect of PCSK9 inhibitors versus placebo

	Base-Case – Direct	Scenario - Indirect		
		CTTC Event Rate Ratio per 1mmol/L (95% CI)	HR Scenario 1	HR Scenario 2
Non-fatal MI	0.79	0.74 (0.71-0.77)	0.61	0.47
Non-fatal ischemic stroke	0.74	0.80 (0.75-0.85)	0.69	0.58
Non-fatal haemorrhagic Stroke	1.02	1.10 (0.927 – 1.37)	1.17	1.27
CVD Death	0.96	0.88 (0.84-0.92)	0.81	0.73

CI, Confidence Interval; CTTC, Cholesterol Treatment Trialists Collaboration; CVD, cardiovascular disease; HR, Hazard ratio; MI, myocardial infarction; mmol/L, millimoles per litre; PCSK9, proprotein convertase subtilisin/kexin type 9.

Scenario analyses are conducted at different LDL-C thresholds. When treatment effects are applied indirectly in this scenario, the percentage LDL-C reduction from the ODYSSEY LONG TERM trial corresponding to the baseline LDL-C value is used to estimate the treatment effect. The treatment effect is greater at higher baseline LDL-C levels under the indirect approach given the greater absolute LDL-C reduction.

7.4.11 Discontinuation rates

The treatment effect estimates derived from the meta-analysis were derived from a population where approximately 5.7% of the population discontinued therapy annually (6,7). The median follow-up was approximately 2.4 years (6,7). To account for this, this discontinuation rate was applied pro-rata for the first 10 cycles of the model and drug costs were adjusted accordingly. For simplicity, it is assumed that that there is no further drug discontinuation after cycle 10.

In scenarios where treatment effects are applied indirectly, it is assumed there are no discontinuations and treatment costs and effects are not adjusted.

7.4.12 Probabilistic analysis

Survival, cost, utility and treatment effectiveness parameters from the appropriate distribution as described in their relevant section above. Probabilistic results were estimated from the model using Monte Carlo simulation by running the model for 1,000 iterations.

7.5 Results

7.5.1 Population characteristics

The source and values of the modelled population characteristics for the TILDA Population 1 (defined as having experienced a MI, angina, TIA or stroke) is presented in Table 7.4. Patient characteristics for all 23 populations examined are presented in the Table A.16 in the Supplementary Appendix.

Table 7.4 Patient characteristics for TILDA population 1 (defined as having experienced MI, angina, TIA or stroke)

Covariate	Value	Source	Covariate	Value	Source
Sociodemographic	1		**Other characteristics**		
Female	0.43	TILDA	History of TIA or Stroke without CAD	0.22	TILDA. Proportion of population with only cerebrovascular disease. Used to adjust the CV hazards for those without CAD.
Age if Male	70	TILDA	History of Stroke	0.15	Not used in baseline Risk equation. Used to adjust utilities for patients who have a history of only-stroke.
Age if Female	72	TILDA	Baseline Utility	0.78	TILDA – as estimated in Chapter 6.
Most Deprived Quintile	0.2	CALIBER	**Biomarkers**		
Stable CAD diagnosis and severity			Pulse (bpm)	72	CALIBER
NSTEMI	0.24	The proportion of patients with an MI was estimated from TILDA. The relative proportion of NSTEMI to STEMI was taken from CALIBER	Creatinine (mmol/L)	100	CALIBER
STEMI	0.17		White Cell count (10^9 /L)	7.46	CALIBER
Unstable Angina	0.06	The proportion of patients with a history of angina was estimated from TILDA and distributed to the different forms of CHD using ratios from CALIBER	Haemoglobin (g/100ml)	1.36	CALIBER
Stable Angina	0.44		**Psychosocial**		
Nonspecific CHD	0.09		History Depression	0.06	History of Psychiatric problems reported in TILDA. Distribution between anxiety and depression assumed equivalent to CALIBER
PCI in last 6 months	0.09	CALIBER	History Anxiety	0.03	
CABG in last 6 months	0.04	CALIBER	**CVD and co-morbidities**		
Previous/Recurrent MI	0.09	TILDA	Heart Failure	0.06	TILDA
Use of Nitrates	0.28	CALIBER	PAD	0.08	CALIBER
Risk Factors			Atrial Fibrillation	0.2	TILDA - Assumed equal to the proportion reporting arrhythmia
Current smoker	0.18	TILDA	History of TIA or Stroke	0.32	TILDA
Ex-smoker	0.48	TILDA	Chronic Kidney Disease	0.07	CALIBER
Hypertension	0.86	TILDA	COPD	0.07	TILDA
diabetes	0.14	TILDA	Cancer	0.08	TILDA
HDL-C	1.56	Assumed equal to the HDL-C of patients with an LDL-C > 4 in TILDA	Chronic Liver Disease	0.01	TILDA
Total Cholesterol	6.17	Assumed LDL-C =4. Estimated equivalent total cholesterol using regression of Total Cholesterol by LDL-C in TILDA	CABG, coronary artery bypass graft; CAD, coronary artery disease; CHD, coronary heart disease; CVD, cardiovascular disease; HDL, high-density lipoprotein; MI, myocardial infarction; NSTEMI, non-ST segment elevation myocardial infarction; PCI, percutaneous coronary intervention; stable-CAD, stable coronary artery disease; STEMI, ST segment elevation myocardial infarction		

7.5.2 PCSK9 inhibitors versus standard of care

7.5.2.1 Markov trace

The basecase scenario examines the cost-effectiveness of PCSK9 inhibitors versus SOC at an LDL-C treatment threshold of >4mmol/L in TILDA population 1 (defined as having a history of TIA, angina, ischemic stroke or MI). The treatment effect is applied directly from the meta-analysis.

The Markov trace from the deterministic analysis is presented in Figure 7.2.This shows the proportion of patients in each of the six model health states over time (as predicted by the multi-state model). The results show a modest reduction in time spent in CV event states for PCSK9 inhibitors compared to SOC. The kink in the graph at ten years reflects the modelled increase in risk of non-CVD death associated with the switch in data source for non-CV mortality risk (from the non-parametric survival curves to Irish lifetables). This choice was made because after ten years non-CV mortality is extrapolated in CALIBER rather than observed and because it is implausible that patients with cardiovascular disease would have a lower rate of non-cardiovascular death than the general population. Such a dramatic increase in the rate of non-CV death at ten years is unrealistic and it is likely that the rate of non-CV death is already underestimated prior to this point. The discrepancy is a limitation of the analysis arising from the limited number of observations contributing to the end of the survival distribution increasing uncertainty. Asaria et al adopted the same approach and a kink was also observed by Asaria et al when using UK life tables (15).

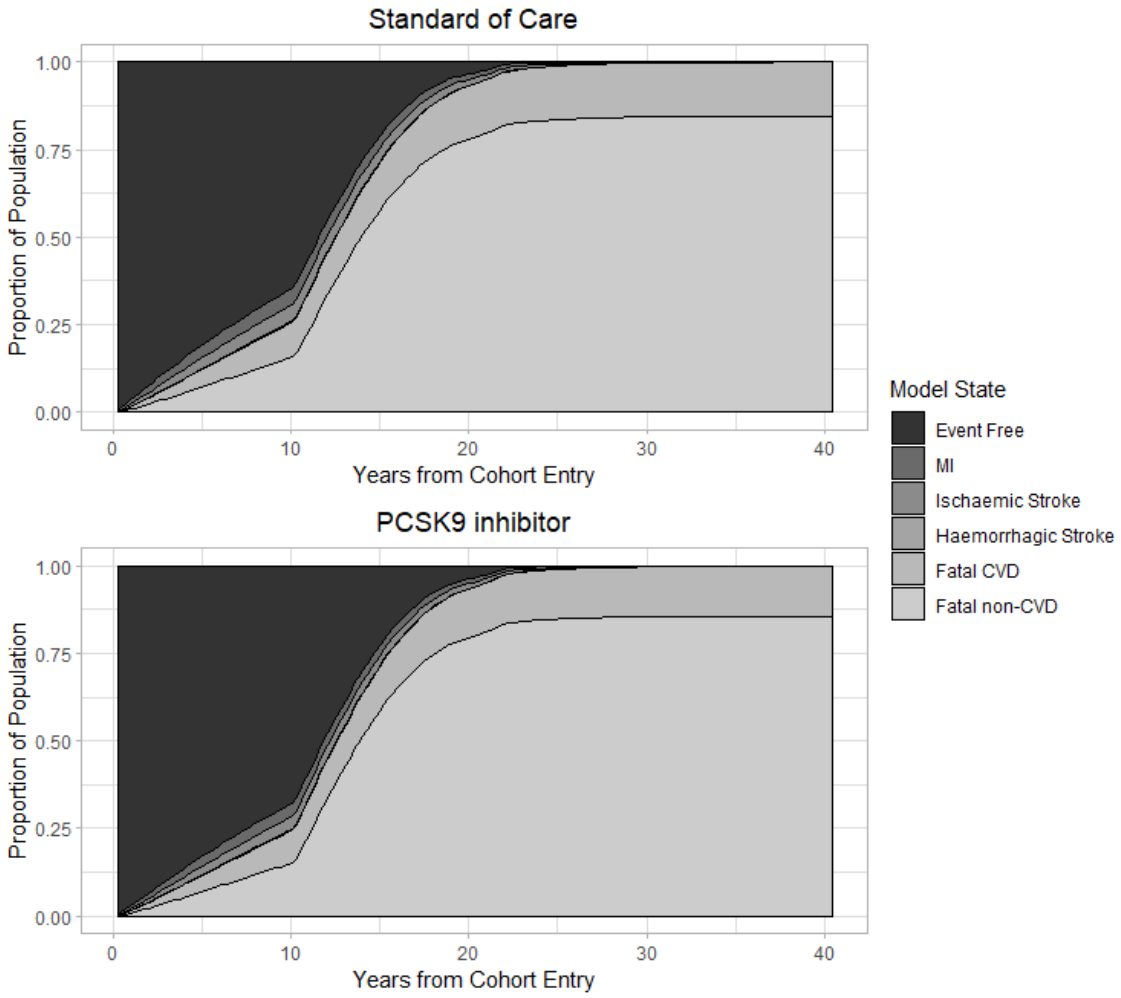


Figure 7.2 Proportion of patients in each of the six primary health states over time as predicted by the multi-state model in the base-case analysis of PCSK9 inhibitors versus standard of care.
Results are estimated from the deterministic analysis. CVD, Cardiovascular Disease; MI, Myocardial infarction.

7.5.2.2 Probabilistic analysis

The expected incremental costs and QALYS in the basecase scenario are plotted in a scatter plot in Figure 7.3. Mean expected costs and QALYS are presented in Table 7.5. The probabilistic ICER for PCSK9 inhibitors versus SOC is €2.3 million per QALY. The scatter plot shows that a substantial proportion of iterations are in the north-west quadrant (21.3%). Given the uncertainty in the CV mortality treatment effect hazard ratio, which includes a hazard ratio greater than one, PCSK9 inhibitors are associated with negative outcomes (versus SOC) in some iterations.

Table 7.5 Probabilistic results of the basecase analysis for the comparison of PCSK9 inhibitors versus standard of care

	Costs			**QALYS**		**ICER (€/QALY)**
PCSK9i	SoC	Incremental	PCSK9i	SoC	Incremental	
€85,052	€7,947	€77,105	7.26	7.22	0.03	2,266,095

ICER, Incremental Cost-Effectiveness Ratio; PCSK9i, proprotein convertase subtilisin/kexin type 9 inhibitor; SoC, Standard of Care.

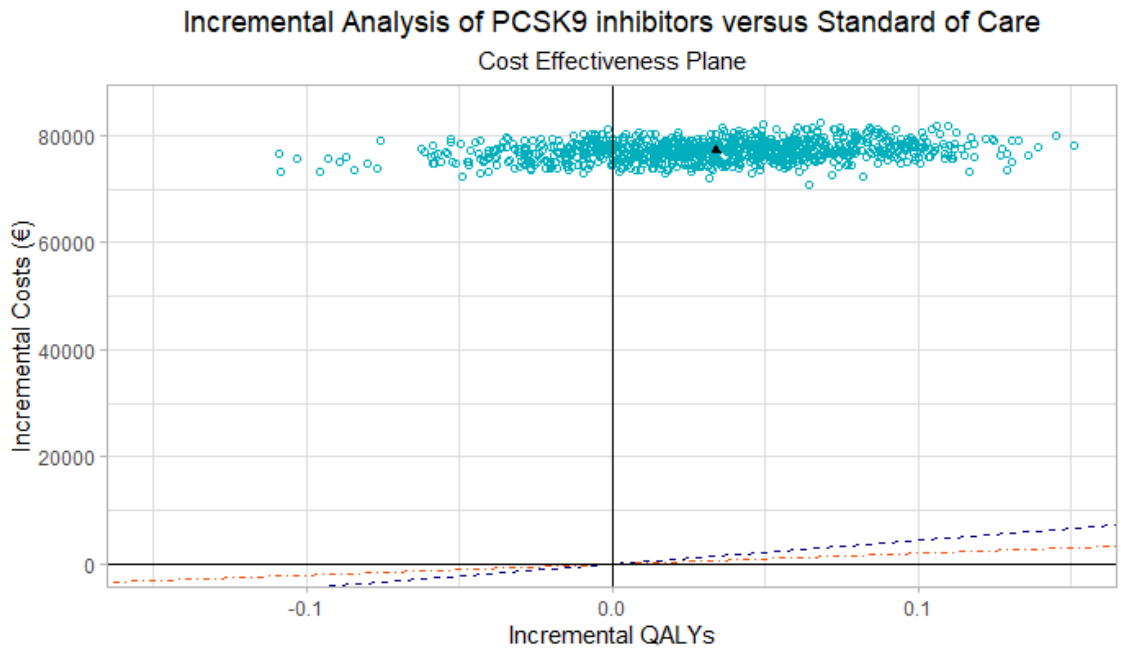


Figure 7.3 Scatter plot of incremental comparison of expected costs and QALYs for PCSK9 inhibitors versus standard of care at an LDL-C threshold of 4mmol/L. (Direct Treatment Effect)
Every blue circle represents the corresponding incremental costs and QALYS of 1,000 iterations of the decision analytic model. The black triangle represents the mean cost per QALY. The navy dashed line represents a cost-effectiveness threshold of €45,000 per QALY. The orange dashed line represents a threshold of €20,000 per QALY.

The cost-effectiveness acceptability curve of PCSK9 inhibitors versus SOC in the basecase scenario is shown in Figure 7.4. At the cost-effectiveness threshold of €45,000 per QALY, there is a 0% probability that PCSK9 inhibitors are cost-effective versus SOC. The curve appears to level off before the maximum cost-effectiveness threshold examined of €5 million per QALY. Because there are iterations in the north west quadrant, there is no cost-effectiveness threshold that would lead to a probability of cost-effectiveness of greater than 75%.

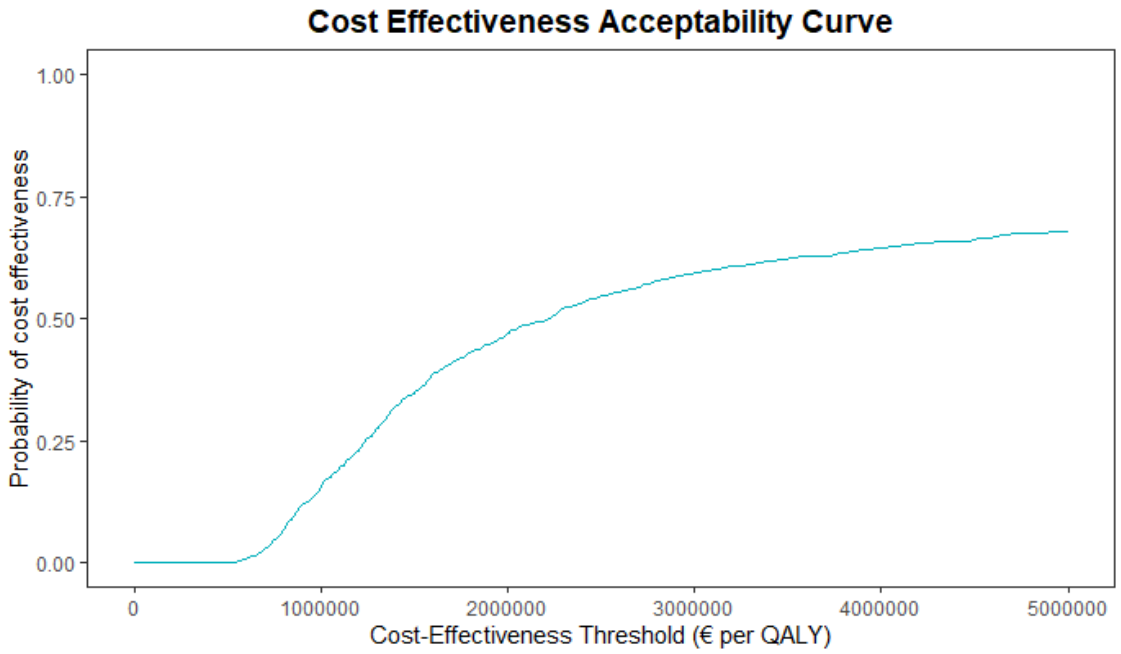


Figure 7.4 Cost effectiveness acceptability curve for PCSK9 inhibitors versus standard of care under the basecase analysis
PCKS9, QALY, Quality Adjusted Life Year.

7.5.2.3 Heterogeneity

To examine cost-effectiveness across heterogenous subgroups of the population, cost-effectiveness was modelled in 23 subgroups from TILDA (defined by clinical characteristics). Population characteristics were informed from TILDA. The model was cumbersome and thus, for pragmatic reasons, these analyses were calculated deterministically only. Therefore, these results should be considered indicative only (given differences between deterministic and probabilistic ICERs). This issue is described in greater detail in Section 7.6.1.

The total and incremental costs and QALYS are presented in Table 7.6. At an LDL-C threshold of 4mmol/L, PCSK9 inhibitors were not cost-effective in any of the subgroups examined. Of the

TILDA populations examined, the lowest ICER was observed in Population 20 (defined as a history of disease in coronary and cerebrovascular vascular beds).

In Ireland, PCSK9 inhibitors are reimbursed for patients with a history of MI at an LDL-C threshold of 4 mmol/L. The deterministic ICER estimated for this subgroup is €1.2 million per QALY.

As highlighted in Section 7.4.10, limitations in the clinical evidence base mean there is uncertainty surrounding the efficacy of PCSK9 inhibitors and the relationship between treatment benefit and baseline LDL-C. In the Supplementary Appendix Table A.17, results are estimated using the indirect approach for all 23 TILDA populations This assumes a treatment-effect equivalent to statins after accounting for the extent of LDL-C reduction. PCSK9 inhibitors were not cost-effective in any subgroup examined. However, ICERs were substantially lower to that estimated under the direct treatment effect approach. ICERs ranged from (€352,000 to €1 million per QALY).

The model was re-estimated using different baseline LDL-C values under the direct and indirect treatment effect assumptions for TILDA Population 1. Results are presented in Table 7.7. As expected, there is a negative correlation between baseline LDL-C level and the ICER. The range of ICERs across different baseline LDL-C values was much wider for ICERs estimated under the indirect approach. Even when assuming optimistic estimates of treatment effectiveness under the indirect treatment effect assumptions, the lowest ICER was €422,139 per QALY.

Subgroup analyses by age and sex are presented in the supplementary Appendix Table A.18. ICERs are lower in men compared to women. ICERs decrease with increasing age.

As the analysis indicated that PCSK9 inhibitors would not be cost-effective in any subgroup, no further exploration of these subgroups was conducted.

Table 7.6 Total and incremental costs and outcomes for PCSK9 inhibitors versus standard of care for 23 subgroups TILDA with a modelled of 4mmol/L , under the direct treatment effect approach (deterministic analysis)

		Costs			QALYs			Deterministic
		PCSK9i	SOC	Inc	PCSK9i	SOC	Inc	ICER (€ per QALY)
1	Any Angina, MI, TIA or Stroke (Base-case)	€94,417	€10,167	€84,250	7.21	7.15	0.05	€1,553,786
2	Any Angina MI TIA, Stroke or other CVD	€96,573	€9,952	€86,621	7.50	7.45	0.05	€1,721,293
3	Any Angina, MI or other CVD	€97,251	€10,388	€86,863	7.46	7.41	0.05	€1,670,174
4	Any Angina or MI	€95,283	€10,860	€84,423	7.13	7.08	0.06	€1,478,818
5	Any TIA or Stroke	€94,722	€12,506	€82,215	6.97	6.90	0.07	€1,150,256
6	Any Angina	€95,556	€9,942	€85,614	7.01	6.96	0.05	€1,684,591
7	Only Angina (No MI, TIA or stroke)	€98,393	€8,683	€89,711	7.62	7.58	0.05	€1,977,208
8	Any MI	€98,288	€14,405	€83,883	7.27	7.20	0.07	€1,195,151
9	Only MI (No Angina, TIA, or stroke)	€99,735	€13,764	€85,971	7.82	7.76	0.07	€1,284,711
10	>1 MI	€109,349	€20,622	€88,728	7.05	6.98	0.07	€1,217,024
11	Any TIA	€92,511	€12,247	€80,264	7.05	6.98	0.07	€1,133,942
12	Only TIA (other heart excluded)	€91,107	€10,733	€80,374	7.74	7.68	0.06	€1,312,937
13	Any Stroke	€93,868	€12,278	€81,590	6.46	6.39	0.07	€1,215,244
14	Only Stroke (No angina, MI, TIA, or other CVD)	€92,371	€10,786	€81,585	6.85	6.79	0.06	€1,279,424
15	>1 Stroke	€93,569	€14,164	€79,405	4.72	4.65	0.07	€1,093,668
16	Other CVD	€108,670	€8,648	€100,023	9.13	9.09	0.04	€2,640,474
17	Only other CVD	€103,448	€9,449	€93,998	8.25	8.20	0.04	€2,213,790
18	MI or Stroke	€97,404	€12,443	€84,961	7.24	7.18	0.06	€1,360,315
19	>1 MI or > 1 Stroke	€109,563	€20,343	€89,219	7.32	7.25	0.07	€1,209,163
20	Disease in two vascular beds i.e. (MI or Angina) and (TIA or stroke)	€91,878	€16,858	€75,020	5.19	5.10	0.09	€862,712
21	As population 1 but excluding Population 20	€96,655	€9,985	€86,670	7.55	7.50	0.05	€1,717,333
22	Diabetes and CVD (Angina, MI, TIA or Stroke)	€102,158	€13,563	€88,596	5.52	5.47	0.05	€1,739,324
23	Diabetes and MI	€102,402	€21,094	€81,309	6.55	6.47	0.08	€978,610
	Subgroups by Modified Baseline Risk Parameters in Population 1							
	TILDA population 1 with 26% Heart Failure instead of 7%.	€96,059	€10,561	€85,498	7.13	7.07	0.05	€1,561,546

CAD, coronary artery disease, CVD, CV disease, ICER, incremental cost-effectiveness ratio; Inc, Incremental; MI, myocardial infarction, PAD, peripheral artery disease; PCSK9, proprotein convertase subtilisin/kexin type 9; QALYs, Quality Adjusted Life Year; SoC, standard of care, TIA, Transient ischemic attack, TILDA, The Irish Longitudinal Study on Aging.

Table 7.7 Total and incremental costs and outcomes for PCSK9 inhibitors versus standard of care for TILDA population 1 by LDL-C threshold.

	Direct Treatment Effect							Indirect Treatment Effect with baseline LDL-C specific percentage LDL-C reduction						
Baseline LDL-C (mmol/L)	**Costs**			**QALYs**			**ICER (€ per QALY)**	**Costs**			**QALYs**			**ICER (€ per QALY)**
	PCSK9i	**SOC**	**Incremental**	**PCSK9i**	**SOC**	**Incremental**		**PCSK9i**	**SOC**	**Incremental**	**PCSK 9i**	**SOC**	**Incremental**	
6	€94,494	€11,013	€83,481	7.18	7.12	0.06	€1,406,701	€102,760	€11,013	€91,746	7.34	7.12	0.22	€422,139
5	€94,455	€10,581	€83,874	7.19	7.14	0.06	€1,478,358	€102,607	€10,581	€92,026	7.32	7.14	0.18	€508,180
4	€94,417	€10,167	€84,250	7.21	7.15	0.05	€1,553,786	€102,447	€10,167	€92,280	7.30	7.15	0.14	€636,931
3.5	€94,398	€9,968	€84,431	7.21	7.16	0.05	€1,592,972	€102,364	€9,968	€92,396	7.29	7.16	0.13	€728,869
3	€94,380	€9,773	€84,607	7.22	7.17	0.05	€1,633,172	€102,278	€9,773	€92,506	7.28	7.17	0.11	€851,570
2.5	€94,362	€9,582	€84,779	7.23	7.18	0.05	€1,674,410	€102,191	€9,582	€92,608	7.27	7.18	0.09	€1,023,741
1.4	€94,322	€9,179	€85,144	7.24	7.19	0.05	€1,768,915	€101,989	€9,179	€92,810	7.24	7.19	0.05	€1,844,335

ICER, incremental cost-effectiveness ratio; LDL-C, Low-density lipoprotein cholesterol; mmol/L, millimoles per litre; PCSK9i proprotein convertase subtilisin/kexin type 9 inhibitor, QALY, Quality Adjusted Life Year; SOC, standard of care.

7.5.2.4 Scenario analysis

7.5.2.4.1 Treatment effect

Scenario analysis, where treatment effectiveness is estimated using the surrogate endpoint of LDL-C, is presented in Table 7.8. The ICERs for PCSK9 inhibitors versus SOC fall substantially when an indirect treatment effect is assumed. But estimates remain substantially greater than a cost-effectiveness threshold of €45,000 per QALY.

Previous economic evaluations of PCSK9 inhibitors have applied mean trial estimates of the percentage lowering ability of PCSK9 inhibitors when estimating the cost-effectiveness of PCSK9 inhibitors with high baseline LDL-C. However, in Chapter 3, it was found that the relative percentage reduction in LDL-C depends on the baseline LDL-C. Therefore, while patients with higher baseline LDL-C have a greater absolute reduction in LDL-C than those with lower baseline values, applying mean trial percentage lowering LDL-C reductions, overestimates the absolute LDL -C reduction. This overestimates the predicted treatment benefit and biases results in favour of PCSK9 inhibitors. As shown in Table 7.8, applying the mean percentage LDL-C reduction from the total ODYSSEY LONG TERM trial population underestimates the ICER by 28% compared to applying the subgroup specific estimate.

Table 7.8 Deterministic costs and outcomes for TILDA population at an LDL-C treatment threshold of 4mmol/L under alternative treatment effect assumptions.

	Costs			QALYs			ICER
	PCSK9i	SOC	Inc.	PCSK9i	SOC	Inc	(€/QALY)
Basecase	€94,417	€10,167	€84,250	7.21	7.15	0.05	€1,553,786
Mean percentage reduction in LDL, estimated from subgroup, with a baseline LDL-C >4.14mmol/L, in ODYSSEY LONG TERM	€102,447	€10,167	€92,280	7.3	7.15	0.14	€636,931
Mean percentage reduction in LDL estimated from the full population in ODYSSEY LONG-TERM	€102,655	€10,167	€92,488	7.35	7.15	0.20	€461,207

ICER, incremental cost-effectiveness ratio; Inc, incremental; LDL-C, Low-density lipoprotein cholesterol; mmol/L, millimoles per litre; QALY, Quality Adjusted Life Year; SOC, Standard of Care.

7.5.2.4.2 Utility

The difference in total and incremental QALYs and the corresponding effect on the ICERs under different utility assumptions is presented in Table 7.9.

There is a minimal QALY gain between PCSK9 inhibitors and SOC under the direct treatment effect assumptions. Therefore, scenario analysis is also conducted, under the indirect treatment effect assumption, to examine the impact of different utility sources (where QALY gains are larger). The assumptions underlying each utility scenario are presented in Table 6.11.

There is minimal difference in utilities and QALY gain under any of the Utility Scenario 1 assumptions. This is expected given that these four scenarios represent minor variations of the same assumption.

Utility Scenario 2 analysis uses predicted utilities from TILDA, analysed using the multiplier approach (using the same methods as Ara et al in their analysis of Health Survey for England data) (253). This facilitates comparison with Ara et al's data which are used in Scenario 3 (253). Total QALYs using the TILDA multiplier (Utility Scenario 2) are substantially higher than those estimated using data from Ara et al (Utility Scenario 3) (253). This because of the higher initial values used in this population; these are then propagated throughout the model leading to higher QALY gain. Higher initial values also mean that chronic decrements post events are higher in absolute terms.

The Ara et al approach fails to control the disutility associated with common co-morbidities, when estimating disutility associated with CV events (253). This may lead to an over-estimation of the QALY gain attributable to the prevention of CV events. This may be the cause of the difference in utilities gained under the Utility Scenario1 and Utility Scenario 2 assumptions.

Asaria et al included two utility scenarios in his economic model (15). Utility Scenario 4b applies utility decrements after an event for one year only. MI and stroke have long term consequences for health which this scenario fails to capture. This means that QALY gains associated with PCSK9 inhibitors are underestimated. Under Utility Scenario 4a, the incremental QALY gain is similar to the QALY gain under TILDA base-case approach. However, the total QALYs gained in both the PCSK9 inhibitor and SOC arm are substantially lower than Utility scenario 1a indicating substantial differences in the underlying utilities. The initial Asaria et al utility (Scenario 4) is lower than the corresponding TILDA scenario, therefore the QALY gain associated with a reduction in mortality is underestimated compared to Utility Scenario 1a (15). However, this

underestimation is balanced by an overestimation in HRQoL lost by assuming substantial constant utility decrements post event.

Table 7.9 Scenario analysis for basecase analysis PCSK9 inhibitors under different utility source under direct and indirect treatment effect assumptions.

	Direct Treatment Effect					Indirect Treatment Effect				
Utility Scenario	**Total QALY PCSK9i**	**Total QALY SOC**	**Inc QALY**	**ICER**	**% Change**	**Total QALY PCSK9i**	**Total QALY SOC**	**Inc QALY**	**ICER**	**% Chang-e**
1a: TILDA Basecase	7.21	7.15	0.05	€1,553,786	N/A	7.30	7.15	0.14	€636,931	N/A
1b: TILDA (no age sex interaction)	7.21	7.15	0.05	€1,561,867	0.52%	7.30	7.15	0.14	€638,330	0.22%
1c: TILDA (no socio-economic control for chronic event disutility)	7.21	7.15	0.05	€1,550,449	-0.21%	7.30	7.15	0.15	€636,393	-0.08%
1d: TILDA (no angina control when estimating chronic event disutility)	7.2	7.15	0.06	€1,514,535	-2.53%	7.29	7.15	0.15	€626,393	-1.65%
2: TILDA using utility multipliers	7.18	7.13	0.06	€1,460,319	-6.02%	7.28	7.13	0.15	€614,914	-3.46%
3: Health Survey for England using multipliers	6.26	6.2	0.05	€1,592,355	2.48%	6.34	6.20	0.13	€693,502	8.88%
4a: Asaria constant event decrement	6.66	6.60	0.06	€1,527,983	-1.66%	6.74	6.60	0.14	€653,071	2.53%
4b: Asaria 1 year only event decrement	6.72	6.68	0.04	€2,195,542	41.30%	6.80	6.68	0.12	€765,315	20.16 %

ICER, incremental cost-effectiveness ratio; Inc, Incremental; QALY, Quality Adjusted Life Year; SOC, Standard of Care; TILDA, The Irish Longitudinal Study on Ageing.

7.5.2.4.3 Price analysis

Evolocumab was reimbursed in Ireland following confidential price negotiations. Therefore, the price modelled in this analysis may not reflect the actual price paid by the HSE. Alirocumab is not reimbursed in Ireland to date. Scenario analysis was conducted where the modelled cost of PCSK9 inhibitors (including wholesale mark-up and patient care fee) was reduced. The results are presented in Table 7.10. PCSK9 inhibitors were not cost-effective at any of the prices examined.

Table 7.10 Scenario Analysis examining cost-effectiveness under different price assumptions

	Costs			QALYs			ICER
	PCSK9 inhibitor	**SOC**	**Inc**	**PCSK9 inhibitor**	**SOC**	**Inc**	
Full Price	€94,417	€10,167	€84,250	7.21	7.15	0.05	€1,553,786
-20%	€83,256	€10,167	€73,089	7.21	7.15	0.05	€1,347,956
-40%	€72,096	€10,167	€61,928	7.21	7.15	0.05	€1,142,125
-60%	€60,935	€10,167	€50,768	7.21	7.15	0.05	€936,295

ICER, Incremental Cost-Effectiveness Ratio; Inc, Incremental; QALY, Quality Adjusted Life Year; SOC, Standard of Care.

7.5.2.4.4 Other scenarios

The results of further scenario analyses are presented in the tornado diagram in Figure 7.5. The CV hazard has a large non-linear effect on the ICER. Of note, discount rates of 0% and 8% both led to increases in the ICER (albeit marginal). The relationship is complex due to the long-term costs and benefits of PCSK9 inhibitors. ICERs are very sensitive to the baseline CV hazards. The result of an analysis where three survival distributions were changed had minimal effect on the ICER (<1% change). In contrast applying the English non-CV mortality life-table rates had a large impact on the ICER (-23%). As observed in Figure 7.2, most of the modelled population are expected to transition to the non-CV fatal health state. This means that changes in risk here can have a large impact on cost-effectiveness.

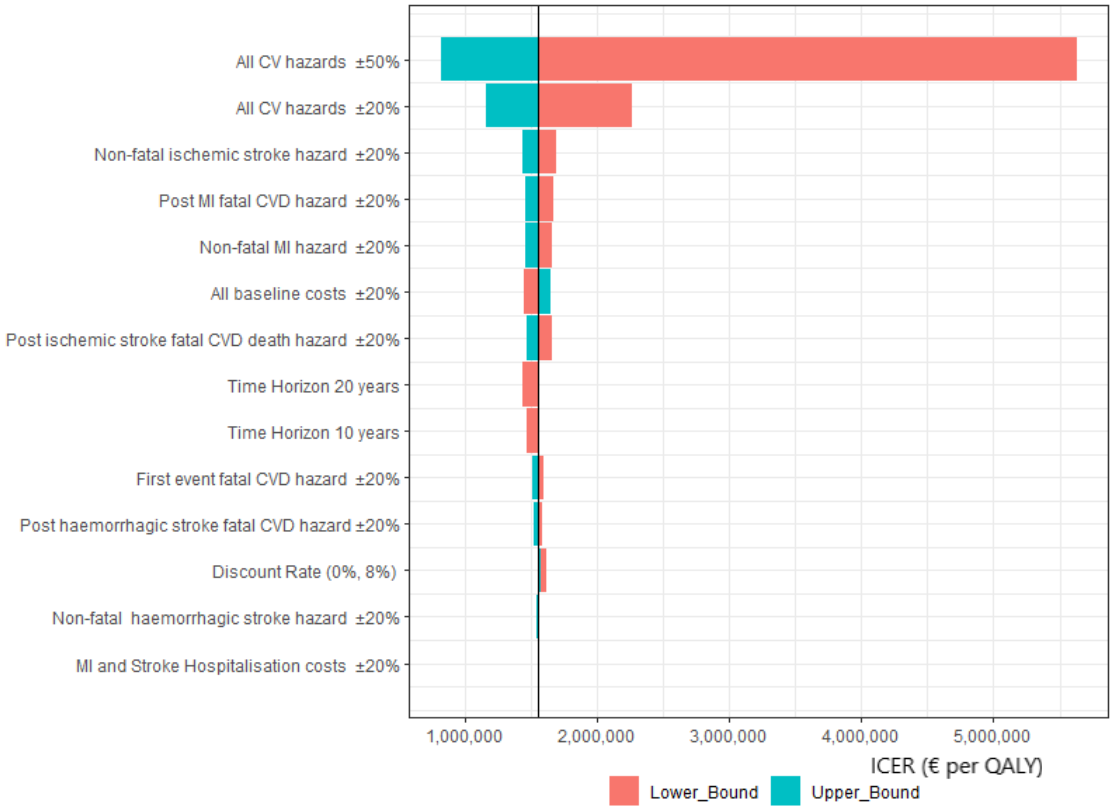


Figure 7.5 Tornado diagram of one way sensitivity analysis for PCSK9 inhibitors versus standard of care under the base-case analysis.
CV, cardiovascular; CVD, cardiovascular disease; ICER, incremental cost-effectiveness ratio; MI, myocardial infarction; QALY, quality adjusted life year

7.5.3 Evolocumab and alirocumab versus standard of care

7.5.3.1 Markov trace

The cost-effectiveness of the specific PCSK9 inhibitors evolocumab and alirocumab are also examined individually versus SOC. As for the comparison between PCSK9 inhibitors and SOC, the population examined in the base-case is TILDA population 1 (defined as a history of TIA, angina, ischemic stroke or MI). A LDL-C threshold of 4mmol/L is assumed.

The Markov trace showing the probability of health state measurement across the model time horizon is shown in Figure 7.6. The traces for all three treatments are similar with only modest differences in fatal and non-fatal CV events between each comparison.

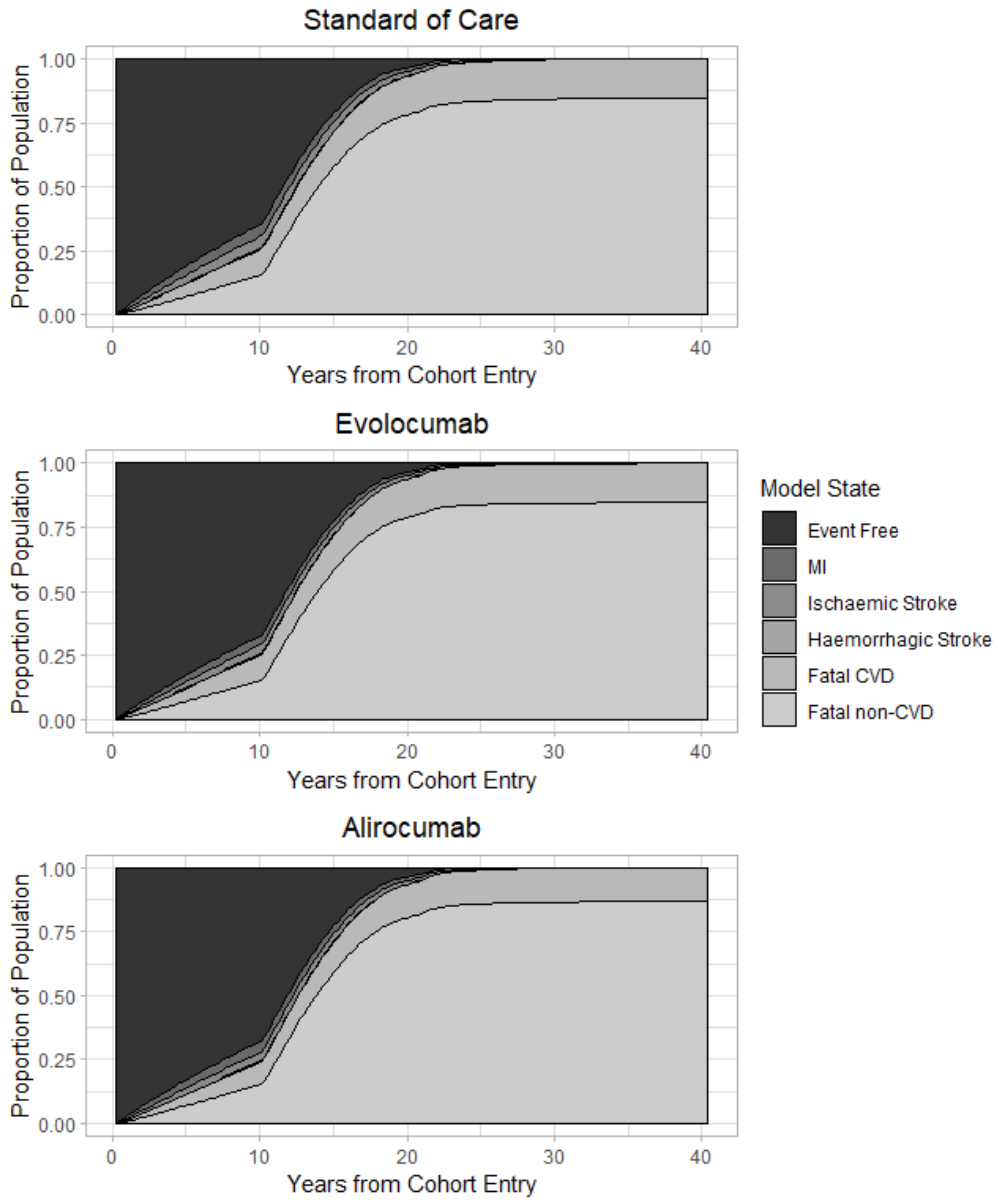


Figure 7.6 Proportion of patients in each of the six primary health states over time as predicted by the multi-state model in the base-case analysis of alirocumab, evolocumab and standard of care.
Results are estimated from the deterministic analysis. CVD, CV Disease; MI, Myocardial infarction;

7.5.3.2 Probabilistic analysis

Mean expected costs and QALYS are presented in Table 7.11. Evolocumab is strongly dominated by SOC; it is associated with higher expected costs and less expected QALYS. Therefore, evolocumab is excluded from the remaining comparison. The probabilistic ICER of alirocumab versus SOC is €1.3 million per QALY.

Table 7.11 Mean expected costed and outcomes for the comparison of alirocumab, evolocumab and SoC

	Mean Costs	Mean QALYs	All treatments versus SOC			Comparison of alirocumab versus evolocumab		
			Inc. Costs	Inc QALYS	ICER	Inc Costs	Inc. QALYS	ICER
SOC	€7,947	7.22						
Evolocumab	€83,778	7.21	Dominated					
Alirocumab	€85,915	7.28	€77,969	0.06	€1,278,817	€2,137	0.07	€30,768

ICER, incremental cost-effectiveness ratio; QALYs Quality Adjusted Life years; SOC, Standard of Care; Inc, Incremental.

The expected incremental costs and QALYS are plotted in a scatter plot in Figure 7.7. The scatter plot shows that there is a lot of uncertainty surrounding the expected costs and outcomes. Both evolocumab and alirocumab have a substantial number of iterations in the north-west quadrant which represents incrementally negative health outcomes. The scatter plot distribution for both PCSK9 inhibitors are largely overlapping. However, given the marginally higher list price, alirocumab iterations have slightly higher costs.

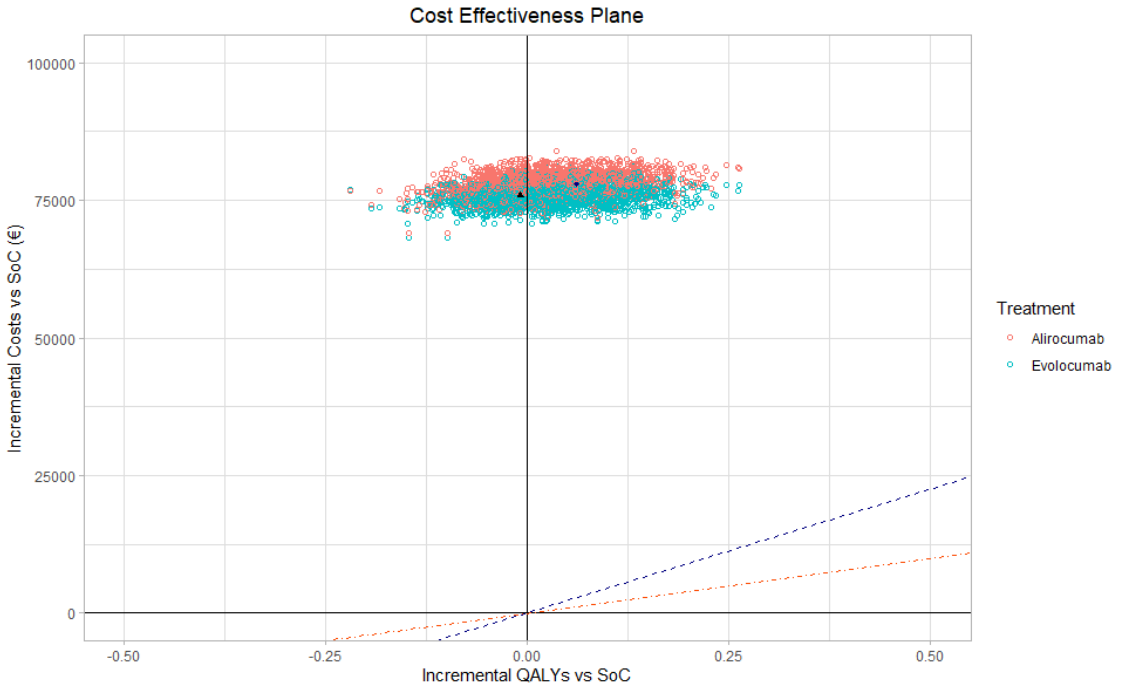


Figure 7.7 Scatter plot of incremental costs and QALYS for alirocumab and evolocumab versus standard of care
The origin represents the mean costs and QALYs with standard of care. Other points are plotted relative to this point. The black triangle and navy diamond represent the mean costs per QALY for evolocumab and alirocumab respectively. The navy and orange dashed lines represent cost-effectiveness thresholds of €45,000 and €20,000 per QALY respectively.
QALYs, Quality Adjusted Life year; SoC, Standard of Care.

At a cost-effectiveness threshold of €45,000 per QALY, neither evolocumab nor alirocumab are cost-effective under the base-case assumptions. The cost-effectiveness acceptability curves for all three treatment options are presented in Figure 7.8.

It has previously been stated that the basis of decision making should be the expected net benefit rather than the option with the highest probability of being cost-effective. When more than two options are being considered, the option with the highest probability of being cost-effective may not be the optimal option. This point is not illustrated clearly in the cost-effectiveness acceptability curve. Therefore, the cost-effectiveness acceptability frontier has been proposed as an alternative (310). This frontier only plots the option which has the highest expected net benefit at every given cost-effectiveness threshold. In Figure 7.8 below, the cost-effectiveness acceptability frontier is represented by the SOC curve until the alirocumab ICER is equal to the cost-effectiveness threshold at an threshold of €1.3 million per QALY, then the frontier is represented by the alirocumab curve.

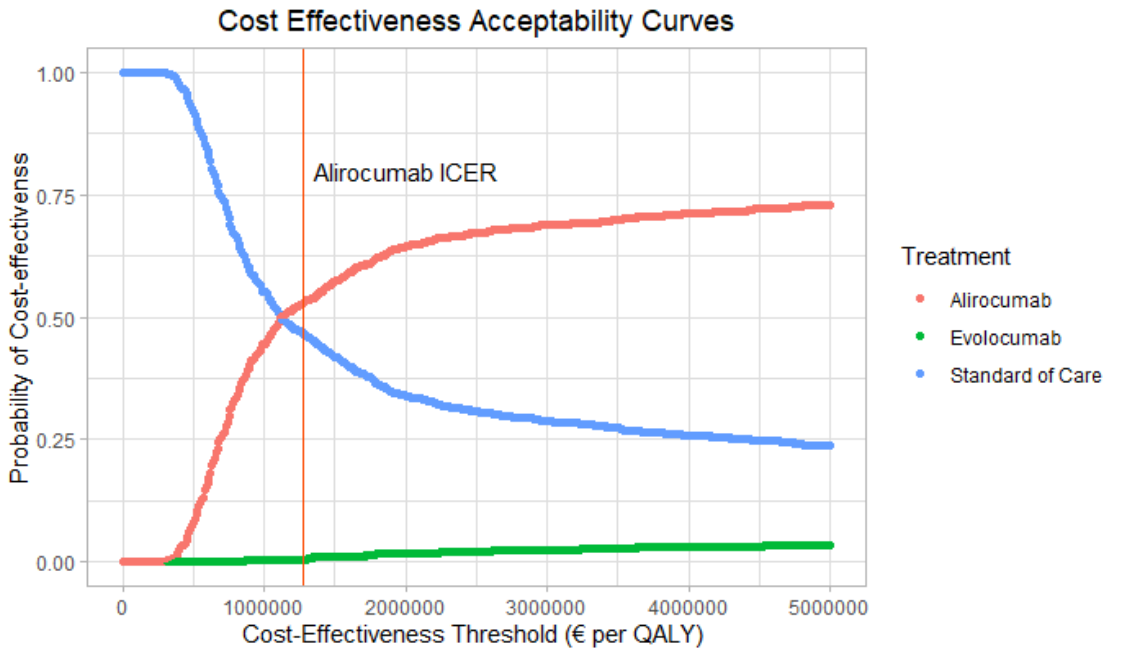


Figure 7.8 Cost-Effectiveness acceptability curve for the comparison of evolocumab , alirocumab and standard of care.

Population is TILDA population 1 at an LDL-C threshold of 4mmol/L under direct treatment effect assumptions. The cost-effectiveness acceptability frontier is represented by the standard of care probability until the alirocumab ICER at €1.3 million. After this point, the cost-effectiveness acceptability frontier is represented by the alirocumab probability.
ICER, Incremental Cost-Effectiveness Ratio; QALY, Quality Adjusted Life year.

7.5.4 Alirocumab versus evolocumab

As evolocumab is already reimbursed, an incremental comparison of alirocumab versus evolocumab is also of interest. The probabilistic ICER of alirocumab versus evolocumab in TILDA Population 1 is an ICER of €30,768 per QALY. However, evolocumab is not universally reimbursed in this population and is restricted to those with a history of MI or CABG only. The cost-effectiveness acceptability curve is presented in Figure 7.9. At a threshold of €45,000 per QALY, alirocumab has a 71.9% of being cost-effective relative to evolocumab.

It is important to note that the multi-technology appraisal of alirocumab versus evolocumab relies on the validity of the adjusted indirect comparison of alirocumab and evolocumab. As highlighted in Chapter 3, the comparison is very uncertain given the differences in populations between the two trials used to populate the comparison. Therefore, both the scatter plot on the cost-effectiveness plane and the cost-effectiveness acceptability curve fail to quantify all the uncertainty associated with the comparison of alirocumab versus evolocumab.

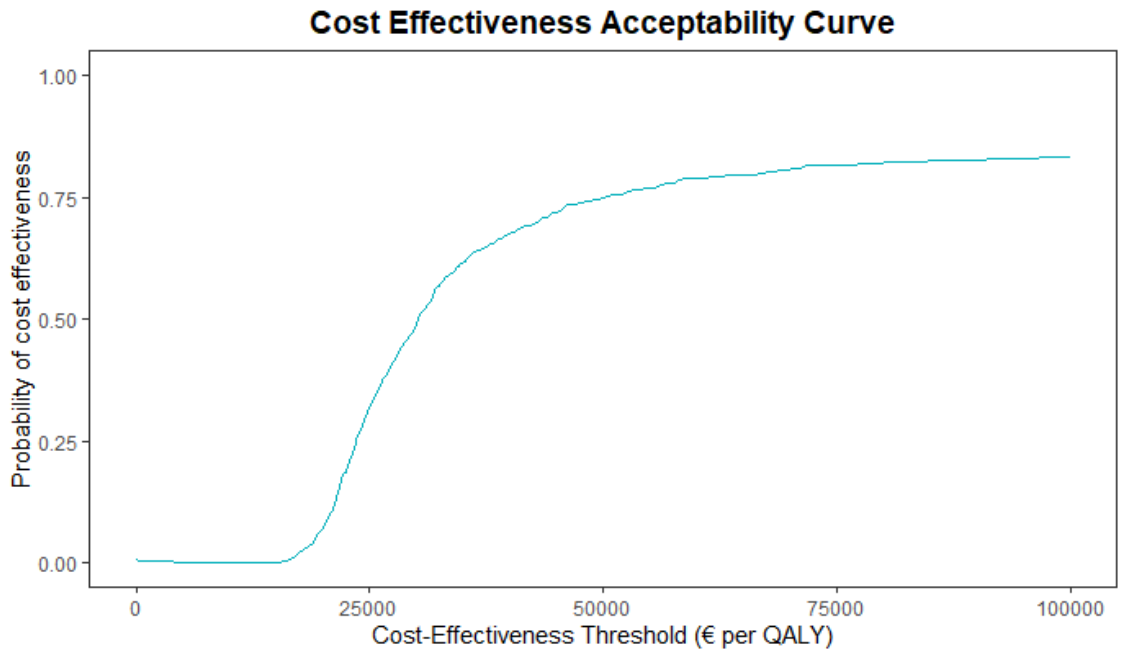


Figure 7.9 Cost Effectiveness acceptability curve for alirocumab versus evolocumab in TILDA population 1 at an LDL-C threshold of 4mmol/L assuming direct treatment effect.

7.6 Discussion

The results show that at current list prices, there is a low probability that PCSK9 inhibitors are cost-effective compared to SOC. This remains the case even if reimbursement is restricted to those with very high CV risk and with high baseline LDL-C levels. There is a lot of uncertainty in the clinical evidence base. When applying optimistic treatment effect assumptions estimated using the indirect approach, the ICER falls substantially. However, ICERs are still much higher than a cost-effectiveness threshold of €45,000 per QALY. Given the modest expected QALY gain, very large price reductions would be required for PCSK9 inhibitors to be cost-effective.

7.6.1 Deterministic versus probabilistic results

Probabilistic ICERs are preferred to deterministic ICERS as the latter do not account for non-linearities between the inputs and the results of the decision analytic model. In this analysis, there are large differences between deterministic and probabilistic ICERS with the deterministic ICER being approximately 32% less than its probabilistic equivalent. There are several reasons for this. First, the estimated QALY gain is relatively small. Therefore, even a small change in the estimated QALY gain can have a large effect on the ICER given its ratio characteristics. Further, the ICER is very sensitive to the CV risk. The tornado diagram in Figure 7.5 shows substantial non-linearity when the CV risk parameters are varied together. It is likely that the non-linearity derives from competing risks of non-CV death. At lower CV hazards, patients experience a fatal non-CVD event before having the opportunity to avoid a cardiovascular event.

For every analysis conducted, PCSK9 inhibitors were not cost-effective versus SOC. Therefore, the differences between probabilistic and deterministic ICERs is a theoretical rather than a practical consideration. The deterministic ICERs presented should be considered a lower bound to the expected ICER. If more precise cost-effectiveness estimates are required for decision making, ICERs should be re-estimated using the probabilistic method.

However, calculating probabilistic ICERs come at the expense of limitation of greater computational effort. On a standard computer one iteration of the multi-state model took less than a minute. However, it took over 16 hours per scenario to run the 1,000 iterations required for the Monte-Carlo simulation. The comparison of evolocumab, alirocumab and SOC took even longer given the need to model three different treatments in each iteration. Asaria et al conducted their analysis using a super-computer because of the computational burden (311). Future analysis should consider using similar methods.

7.6.2 Alirocumab versus evolocumab

While a previous comparison of evolocumab versus alirocumab was submitted to the Scottish Medicines Consortium (35), no published economic evaluation of alirocumab versus evolocumab has been identified that uses evidence from the relevant CV outcomes trials. This analysis was completed before the CV outcomes trials were published. This analysis shows that alirocumab is expected to be cost-effective relative to evolocumab. The probabilistic comparison should be repeated for the only-MI population to confirm the expected result.

7.6.3 Benefits of real-world data

One of the key strengths of this analysis was the richness of UK electronic health record data used to derive the risk equations and costs. Following publication of the economic model, Asaria et al published an article reflecting on the key challenges and opportunities of using such data (311). The computational demands of analysing very large datasets and the care required when handling sensitive data was noted. Further, it can be challenging to extract event data form datasets which were not designed to collect such data or where there are inconsistencies across datasets. The importance of having clear rules when defining clinical events across multiple datasets was highlighted. Changing event coding practices over time add to the complexity but is inevitable when analysing data in a long-term study. The ability of such data to capture key aspects of identifiable heterogeneity was highlighted (311).

7.6.4 Estimation of baseline risk

In some cases, none of the parametric models chosen offer a great fit to the data. If access to the raw data was possible, more flexible distributions could be considered (such as piece wise models) which may reflect the observed data more accurately. Access to the raw data would also allow the joint consideration of alternative survival distributions across competing risks as well as a comparison to the directly observed data.

7.6.5 Treatment effect

To our knowledge, this is the first academic study which has identified that the percentage lowering ability of PCSK9 inhibitors on LDL-C depends on the baseline LDL-C level. Previous economic evaluations which have conducted analyses using the indirect approach have overestimated the cost-effectiveness of PCSK9 inhibitors in patients with high baseline LDL-C values (312).

7.6.6 Utilities

Utilities can be a strong driver of economic models. In the absence of an ideal data source for Irish utilities, it is not possible to state whether the utility estimates applied in this model are truly representative of the population with CVD in Ireland. As outlined in Chapter 6, combining utility data from different data sources to account for patients with a history of multiple events is complex. In this case, the baseline predicted utility is estimated directly form patients regardless of their CV history. This decreases one form of uncertainty but comes at the expense of using predicted utility values. The limitations associated with this were outlined in the previous chapter.

7.6.7 Heterogeneity

A broad range of subgroups have been examined in an attempt to identify very high-risk subgroups that PCSK9 inhibitors may be cost -effective in.

The probabilistic analysis showed that non-linearity between the parameters and the expected outcomes mean that deterministic outcomes can be biased. In a similar way, modelling average population characters can lead to biased results if there is also non-linearity (313). In order to jointly account for both parameter uncertainty and heterogeneity, a nested Monte Carlo simulation is required with an outer loop *j* in which parameters are sampled and an inner loop *i* from which the population characteristics are sampled from the frequency distribution is recommended (313). Such an analysis was not computationally feasible in this case and is a limitation of our analysis.

However, extensive efforts were made to examine heterogeneity in the model. Over 23 subgroups were defined based on potential clinical reimbursement criteria. Population characteristics were derived from TILDA, a patient-level nationally representative database. Heterogeneity was captured in subgroup specific baseline risk equations, cost -parameters and utility estimates.

The importance of considering mutually exclusive subgroups was considered. In the UK technology appraisal of alirocumab, it appears that patients with disease in two CV beds were double counted in the derivation of the risk equations. This will bias the analysis in favour of alirocumab (10). The results of Table 7.6 show that the ICER from the remaining secondary prevention group is underestimated by about 12% (Population 1 versus population 20) in this scenario.

ESC guidelines report the results of analysis that state that PCSK9 inhibitors were not cost-effective at mid-2018 prices (32). They stated that cost-effectiveness is improved in higher risk patients but do not define such groups. They cite work of Kazi et al who reported that PCSK9 inhibitors would become cost-effective at lower prices (291). This thesis informs potential guidelines regarding what subgroups the drugs might be relatively most cost-effective in (including those with disease in more than one CV bed and patients with a history of MI and diabetes). However, PCSK9 inhibitors were not cost-effective in any subgroups of the population.

In this case, no formal evaluation of value of information was conducted. But given the that all ICERs are far from a cost-effectiveness threshold, any value associated with future research is likely to be low given the minimal uncertainty with the conclusion that PCSK9 inhibitors are not cost-effective.

A limitation of the analysis presented here is that due to an absence of data in TILDA, there was little analysis of the population with PAD. The avoidance of major adverse limb events in this population (such as amputation observed with trials of evolocumab) is expected to be a driver of cost-effectiveness (98). Given the specific issues in this population, further work is needed to drive robust estimates of cost-effectiveness in this population.

7.6.8 Modelling in R

There has been growing interest in the use of R in health economic and HTA. Jalal et al found that studies using R increased by nearly 50% over 5 years (314). The numerous advantages of R including its flexibility, integration with other packages and ease of documentation and transparency were also highlighted. Academic groups have actively promoted the use of R in HTA and decision analytic modelling including the R for HTA group and the DARTH working groups (315,316).

Hollman et al previously compared four software programs for implementing decision analytic cost-effectiveness models (317). They found that, while Excel and TreeAge Pro are sufficient for simpler models, the efficiency and transparency of using programming language such as R become more valuable when more complex analyses are required. New methodological techniques better represent real life complexity but require more computationally intensive methods (318). There is a steeper learning curve with R compared with Excel, but the script-based approach of R facilitates reproducibility and accuracy assessment. Given the need for more complex methods, tutorials have been written in the literature to facilitate adoption of these methods (319,320).

One of the advantages of R is its extensive use of vectorisation which has increased efficiency compared to other methods (314). However, as much of the CALIBER model and uses "for" loops in its implementation, the increased efficiency is not harnessed optimally. Future work could examine ways of optimising the code to make it more efficient.

An advantage of R is the number of user-built packages. This substantially increases the functionality of R. However, this comes at a price. For example, when preparing this thesis, an error was noted in the 'bcea' package in a function designed to automatically plot the cost-effectiveness acceptability frontier. User defined functions mean that errors in coding can be propagated across multiple analyses. Lists of validated packages may increase acceptability of R health economic methods to the HTA assessment body community. However, this is challenging given that packages are constantly updated and improved.

7.6.9 Limitations

There are several limitations to this analysis. These are in addition to those already outlined in the discussion thus far and those highlighted by Asaria et al (15). The non-linear effects of heterogeneity are not incorporated into the model. Baseline risks are based on English data which may not be transferable to Ireland. Limited data was available on the population with PAD. There is limited clinical evidence for PCSK9 inhibitors which mean that several strong assumptions regarding the clinical evidence base are required. All analyses are based on list prices which may not reflect the actual cost to the HSE. Pragmatically, it is assumed that there is no treatment discontinuation after 10 cycles. This does not reflect reality where patients will discontinue therapy over time. However, given the approximate correspondence between the accrual of drug costs incremental benefits, omission is unlikely to affect the conclusions of our analysis.

7.7 Conclusion

PCSK9 inhibitors have been shown to reduce the risk of CV events. However, regardless of whether optimistic or observed treatment effects are applied, the results show that PCSK9 inhibitors are not cost-effective in the secondary prevention CV population in Ireland. The implications of committing healthcare resources to interventions which are not cost-effective should be considered before extending reimbursement of PCSK9 inhibitors to a greater proportion of the population. Cost-effectiveness may vary by subgroup. An iterative approach to defining subgroups may reduce complexity associated with defining subgroups. It is important that where subgroup analysis is used for defining reimbursement criteria, and cost-effectiveness conclusions differ across the licensed population, recommendations should be

based on mutually exclusive subgroups. The results here may inform guidelines on the selection of patients for treatment with PCSK9 inhibitors (by identifying those who have the potential to benefit most).

The results show that relative to evolocumab, alirocumab is likely to represent a cost-effective alternative. Although further analysis at confidential prices paid would be required to confirm this. The availability of an alternative treatment may facilitate price negotiations to further lower prices.

Chapter 8. Conclusion

8.1 Introduction

The aim of this thesis was to evaluate the cost-effectiveness of PCSK9 inhibitors for the secondary prevention of CVD in Ireland. Also, to identify the subgroups of this population in which PCSK9 inhibitors might be deemed cost-effective. The outputs of this thesis also have a broader impact for economic evaluation in Ireland.

8.2 Main findings

The results of a meta-analysis conducted following a systematic review of the literature are presented in Chapter 3. It was found that PCSK9 inhibitors reduce the time to non-fatal MI and non-fatal stroke. No treatment effect on CVD death was observed over the time period of the clinical trials. But the quality of evidence was very low. To our knowledge, this is the first meta-analysis to measure treatment effectiveness in a meta-analysis using the hazard ratio. This chapter also presents the first full report of an indirect treatment comparison between alirocumab and evolocumab. Given the between-trial population heterogeneity, comparisons between PCSK9 inhibitors are uncertain. The similarity assumption was considered to hold sufficiently to justify the conclusion that there is no evidence of a difference in treatment effect between alirocumab and evolocumab.

Chapter 4 identified for the first time, the interaction that strategic behaviour such as price negotiations can introduce to economic evaluations. If a pharmaceutical company offers a conditional discount, which is dependent on obtaining reimbursement in two subgroups, cost savings generated from the discount in one subgroup may be used to offset the incremental cost of extending reimbursement in the other. A framework was presented to guide the economic evaluation process in the presence of an interaction. It was shown that failure to account for the interaction can lead to incorrect conclusions regarding the cost-effectiveness of interventions. Adoption of the framework is expected to increase population health through the increased recognition of cost-effective interventions. However, this conclusion does not consider the counterfactual that providers may present higher prices for drugs which would result in increased opportunity cost and a reduction in overall population health. Therefore, the net impact on population health is uncertain.

Population heterogeneity between jurisdictions means that utility values derived from international populations may not be transferable to Ireland. In Chapter 5, the links between the EQ-5D-3L and the TILDA dataset was identified. Mapping the TILDA dataset to the EQ-5D-3L was identified as a pragmatic method of generating utilities in the Irish setting in the absence of directly observed evidence (the data gap). A mapping model was derived in a population with

CVD. To our knowledge, this is the first time that a mapping study of this kind has been performed to address such a data gap.

In Chapter 6, the mapping model derived in Chapter 5 was applied to the national TILDA population. Utility values for 23 subgroups of the secondary prevention population were derived. The use of patient-level data means that heterogeneity in the HRQoL of the secondary prevention CVD population can be captured. Regression methods were used to estimate utility decrements for age, and to estimate chronic utility decrements for MI, stroke and multiple CV events. The importance of using adjusted or unadjusted utility values depending on the stage of the analysis was highlighted.

The results of Chapter 3, 5 and 6 were used to populate an economic model in Chapter 7. An incremental economic evaluation of PCSK9 inhibitors versus SOC was conducted in the identified 23 subgroups. The economic model allowed the capture of population heterogeneity on baseline risk, treatment effect, costs and utility values. Regardless of whether direct or conservative indirect treatment effects are applied, the results show that PCSK9 inhibitors are not cost-effective in the secondary prevention CV population in Ireland. To our knowledge, this is the most extensive assessment of the cost-effectiveness of PCSK9 inhibitors across subgroups in the academic literature to date. This is the first evaluation, that has applied subgroup specific LDL-C reduction estimates when predicting the expected costs and outcomes associated with PCSK9 inhibitors (through the surrogate endpoint of LDL-C). Previous estimates of cost-effectiveness in populations with high baseline LDL-C (which used assumed an indirect treatment effect) have been biased in favour of PCSK9 inhibitors. An incremental evaluation of alirocumab and evolocumab to SOC was also conducted. Neither alirocumab nor evolocumab are cost-effective relative to SOC. However, when compared to evolocumab, alirocumab is likely to represent a cost-effective alternative.

8.3 Implications for policy and practice

It is concerning that pre-specified outcomes from pivotal clinical trials are not made publicly available. The conclusions of this thesis were made with incomplete evidence given that to our knowledge, assessments of the effect of alirocumab on HRQoL from the ODYSSEY OUTCOMES trial, have not been made publicly available. It is recommended that international efforts to support publication of trial clinical study reports are supported.

The identification of the interaction between price negotiations and heterogeneity has many implications. Pharmaceutical companies now have additional considerations when devising reimbursement strategies. Strategic behaviour can potentially increase the reimbursement

price of new drugs. Economic evaluation guidelines should be updated to account for the use of this framework and the broader consideration of treatment alternatives that this requires.

Derivation of the mapping model to TILDA opens the door to the link between the EQ-5D-3L and the wealth of data in TILDA. The generation of a mapping model to TILDA means that more evidence is available for researchers who wish to use utilities which measure Irish health and harness the benefits associated with analysing patient-level data. Albeit this comes at the expense of using mapped values (which has its own uncertainty). The strengths and weakness of all utility value sources should be assessed before using the utilities derived from the mapping model in an economic evaluation.

Evolocumab is currently reimbursed in Ireland in a very restricted population under the management of the MMP. While the confidential price paid by the HSE is unknown, the results of the economic evaluation show that reimbursement of evolocumab is unlikely to represent a cost-effective use of scarce healthcare resources in the Irish setting. Consideration should be given to further restricting reimbursement criteria for future patients. At current list prices, alirocumab may represent a cost-effective alternative to evolocumab. If this conclusion holds at confidential net prices for both drugs, consideration should be given to reimbursing alirocumab. However, this is based on the provision that the reimbursement criteria for alirocumab are not broader than the comparable criteria for evolocumab.

International clinical guidelines recognise that PCSK9 inhibitors will not be cost-effective across the total licensed population but suggest cost-effectiveness will be improved in subgroups of the population (32). It is recommended that the results of this analysis are used to inform future subgroup specific recommendations in clinical guidelines. PCSK9 inhibitors are unlikely to represent cost-effective use of healthcare resources. However, where resources are available, access should be prioritised for those who have high baseline LDL-C values and who either have a history of MI and diabetes or those with a history of events in two vascular beds.

8.4 Future research

There is still much uncertainty regarding the relationship between baseline LDL-C and the magnitude of treatment benefit attributable to PCSK9 inhibitors. Further evidence to answer this uncertainty is unlikely to come from randomised controlled trials. It is recommended that the observational data is analysed to resolve this uncertainty. Reimbursement criteria (which restrict reimbursement to subgroups of the population) can be used to increase the power of observational data analysis methods.

The study of strategic behaviour and economic evaluation is just emerging as a research area. The role of game theory in economic evaluation should be further explored. Further research, to examine the prevalence of scenarios where the hybrid approach can be applied in practice and to examine the impact of this method on reimbursement decisions would be beneficial.

The mapping model derived in Chapter 5 is uncertain given the modest size of the population used to derive it. In the absence of directly observed EQ-5D-3L data in TILDA, it is recommended that a further mapping study is undertaken in a larger broader patient population to reduce associated uncertainty. Efforts should be made to ensure sufficient representation of patients in poorer health states. Lessons from this thesis can be used to inform the future study. Future research should harness the longitudinal nature of TILDA and attempt to measure causal utility decrements associated with CV events.

The cost-effectiveness of PCSK9 inhibitors in the primary prevention population was outside the scope of this thesis. Future research should be conducted to estimate the cost-effectiveness of PCSK9 inhibitors in the HeFH population in Ireland.

There is growing interest in the use of R as a tool to build economic models. However, there is a steep learning curve with its use. Future research should be conducted on how to optimise the presentation of economic models built in R to international HTA agencies in order to facilitate robustness, ease of use and transparency.

8.5 Conclusion

In this study, the cost-effectiveness of PCSK9 inhibitors (alirocumab and evolocumab) compared to SOC has been assessed. A primary focus was accounting for the implications of population heterogeneity on the results of the analysis. Clinical, economic and statistical expertise were required to apply a variety of research methods. These included the building of a theoretical framework, systematic review and evidence synthesis, primary data collection, as well as regression and other statistical methods. PCSK9 inhibitors were not found to be cost-effective across any subgroup examined. Several important implications for policy and practice have been outlined.

REFERENCES

1. Ference BA, Ginsberg HN, Graham I, Ray KK, Packard CJ, Bruckert E, et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. Eur Heart J. 2017;385(32):2264–71.

2. Sabatine MS, Giugliano RP, Wiviott SD, Raal FJ, Blom DJ, Robinson J, et al. Efficacy and safety of evolocumab in reducing lipids and cardiovascular events. N Engl J Med. 2015;372(16):1500–9.

3. Colhoun HM, Robinson JG, Farnier M, Cariou B, Blom D, Kereiakes DJ, et al. Efficacy and safety of alirocumab, a fully human PCSK9 monoclonal antibody, in high cardiovascular risk patients with poorly controlled hypercholesterolemia on maximally tolerated doses of statins: rationale and design of the ODYSSEY COMBO I and II trials. BMC Cardiovasc Disord. 2014;14(1):121.

4. European Medicines Agency (EMA). Praluent: Summary of Product Characteristics. Amsterdam: EMA; 2020. Available from https://www.ema.europa.eu/en/documents/product-information/praluent-epar-product-information_en.pdf. Date accessed 31 March 2020.

5. European Medicines Agency (EMA). Repatha: Summary of Product Characteristics. Amsterdam: EMA; 2020. Available from https://www.ema.europa.eu/en/documents/product-information/repatha-epar-product-information_en.pdf. Date accessed 31 March 2020.

6. Schwartz GG, Steg PG, Szarek M, Bhatt DL, Bittner VA, Diaz R, et al. Alirocumab and Cardiovascular Outcomes after Acute Coronary Syndrome. N Engl J Med. 2018;379(22):2097-2107.

7. Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, et al. Evolocumab and clinical outcomes in patients with cardiovascular disease. N Engl J Med. 2017;376(18):1713–22.

8. Health Information Quality Authority (HIQA). Guidelines for the Economic Evaluation of Health Technologies in Ireland. Dublin: HIQA; 2019. Available from https://www.hiqa.ie/sites/default/files/2019-07/HTA-Economic-Guidelines-2019.pdf Date accessed 31 March 2020.

9. National Centre for Pharmacoeconomics (NCPE) Cost-effectiveness of evolocumab (Repatha®) for hypercholesterolemia. Dublin: NCPE; 2018. Available from http://www.ncpe.ie/wp-content/uploads/2016/10/Evolocumab-Summary-Final-20180516.pdf Date accessed 31 March 2020.

10. National Centre for Pharmacoeconomics (NCPE) Cost-effectiveness of alirocumab (Praluent®) for hypercholesterolemia. Dublin: NCPE; 2017. Available from http://www.ncpe.ie/wp-content/uploads/2015/11/Alirocumab-Summary-Final.pdf Date accessed 31 March 2020.

11. National Centre for Pharmacoeconomics (NCPE) Cost-effectiveness of evolocumab (Repatha®) for hypercholesterolemia. Dublin: NCPE; 2016. Available from http://www.ncpe.ie/wp-content/uploads/2016/10/Evolocumab-Website-Summary.pdf Date accessed 31 March 2020.

12. Medicines Management Programme (MMP). Managed Access Protocol – Repatha® (Evolocumab) v1.2. Dublin: HSE; 2020. Available from:

hse.ie/eng/about/who/cspd/ncps/medicines-management/pcsk9-inhibitors/hse-managed-access-protocol-repatha1.pdf Date accessed 31 March 2020.

13. Committee for Medicinal Products for Human Use (CHMP). Summary of Opinion Bempedoic Acid (Nilemdo®). Amsterdam; EMA, 2020. Available from: https://www.ema.europa.eu/en/documents/smop-initial/chmp-summary-positive-opinion-nilemdo_en.pdf Date accessed 31 March 2020.

14. Specialist Pharmacy Service. Inclisiran [Internet]. 2020. Available from: https://www.sps.nhs.uk/medicines/inclisiran/ Date accessed 31 March 2020.

15. Asaria M, Walker S, Palmer S, Gale CP, Shah AD, Abrams KR, et al. Using electronic health records to predict costs and outcomes in stable coronary artery disease. Heart. 2016;102(10):755–62.

16. WHO. HTA Definitions [Internet]. 2020. Available from: https://www.who.int/health-technology-assessment/about/Defining/en/ Date accessed 31 March 2020.

17. National Centre for Pharmacoeconomics [Internet]. 2020. Available from: www.ncpe.ie Date accessed 31 March 2020.

18. Government of Ireland (Oireachtas). Health (Pricing and Supply of Medical Goods) Act 2013, Number 14. Irish Statue Book. Dublin: Office of the Attorney General; 2013. Available from: http://www.irishstatutebook.ie/eli/2013/act/14/enacted/en/html. Date accessed 31 March 2020.

19. Framework agreement on the supply and pricing of medicines. Dublin: HSE; 2016. Available from https://www.hse.ie/eng/about/who/cpu/iphaagreement2016.pdf Date accessed 31 March 2020.

20. O'Donnell H, Lamrock F, Tilson L, Barry M. To HTA or Not to HTA: Identifying the Factors Influencing the Rapid Review Outcome in Ireland. Value Health. 2020;23(2):274-275

21. Murphy A, Redmond S. To HTA or Not to HTA: Identifying the Factors Influencing the Rapid Review Outcome in Ireland—Reply to Letter to the Editor by Helen O'Donnell et al. Value Health. 2020;23(2):275

22. Johannesen KM, Claxton K, Sculpher MJ, Wailoo AJ. How to design the cost-effectiveness appraisal process of new healthcare technologies to maximise population health: A conceptual framework. Health Econ. 2018;27(2):e41–54.

23. O'Mahony JF, Coughlan D. The Irish Cost-Effectiveness Threshold: Does it Support Rational Rationing or Might it Lead to Unintended Harm to Ireland's Health System? Pharmacoeconomics. 2016;34(1):5–11.

24. Chen TC, Wanniarachige D, Murphy S, Lockhart K, O'Mahony J. Surveying the Cost-Effectiveness of the 20 Procedures with the Largest Public Health Services Waiting Lists in Ireland: Implications for Ireland's Cost-Effectiveness Threshold. Value Health. 2018;21(8):897–904.

25. Phelps CE. Good Technologies Gone Bad. Med Decis Mak. 1997;17(1):107–17.

26. Coyle D, Buxton MJ, O'Brien BJ. Stratified cost-effectiveness analysis: a framework for establishing efficient limited use criteria. Health Econ. 2003;12(5):421–7.

27. Basu A, Meltzer D. Value of Information on Preference Heterogeneity and Individualized Care. Med Decis Mak. 2007;27(2):112–27.

28. Kim DD, Basu A. New Metrics for Economic Evaluation in the Presence of Heterogeneity: Focusing on Evaluating Policy Alternatives Rather than Treatment Alternatives. Med Decis Mak. 2017;37(8):930–41.

29. Espinoza MA, Manca A, Claxton K, Sculpher MJ. The Value of Heterogeneity for Cost-Effectiveness Subgroup Analysis: Conceptual Framework and Application. 2014;34(8):951-64

30. Page MM, Watts GF. PCSK9 inhibitors - mechanisms of action. Aust Prescr. 2016;39(5):164–7.

31. Ridker PM, Revkin J, Amarenco P, Brunell R, Curto M, Civeira F, et al. Cardiovascular Efficacy and Safety of Bococizumab in High-Risk Patients. N Engl J Med. 2017;376(16):1527–39.

32. Mach F, Baigent C, Catapano AL, Koskinas KC, Casula M, Badimon L, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. Eur Heart J. 2020;41(1):111-88.

33. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal Alirocumab for treating primary hypercholesterolaemia and mixed dyslipidaemia Committee Papers. London: NICE; 2016. Available from https://www.nice.org.uk/guidance/ta393/documents/committee-papers Date accessed 31 March 2020.

34. Scottish Medicines Consortium (SMC). Alirocumab for patients with primary hypercholesterolaemia (heterozygous familial and non-familial) or mixed dyslipidaemia ID 1147/16. Glasgow: NHS Scotland; 2016. Available from: https://www.scottishmedicines.org.uk/media/1228/alirocumab_praluent_final_july_2016_amended_040816_for_website.pdf Date accessed 31 March 2020.

35. Scottish Medicines Consortium (SMC). Evolocumab for patients with primary hypercholesterolaemia (heterozygous familial and non-familial) or mixed dyslipidaemia ID 1148/16. Glasgow: NHS Scotland; 2017. Available from https://www.scottishmedicines.org.uk/media/1701/evolocumab__repatha__resubmission_final_jan_2017_for_website.pdf Date accessed 31 March 2020.

36. National Institute for Health and Care Excellence (NICE). Final appraisal determination Evolocumab for treating primary hypercholesterolaemia and mixed dyslipidaemia. London: NICE; 2016. Available from https://www.nice.org.uk/guidance/ta394/resources/evolocumab-for-treating-primary-hypercholesterolaemia-and-mixed-dyslipidaemia-pdf-82602910172869 Date accessed 31 March 2020.

37. Health Information Quality Authority (HIQA). Guidelines for Evaluating the Clinical Effectiveness of Health Technologies in Ireland. Dublin: HIQA; 2018. Available from https://www.hiqa.ie/sites/default/files/2019-01/Clinical-Effectiveness-Guidelines.pdf Date accessed 31 March 2020.

38. Robinson JG, Farnier M, Krempf M, Bergeron J, Luc G, Averna M, et al. Efficacy and safety of alirocumab in reducing lipids and cardiovascular events. N Engl J Med. 2015;372(16):1489–99.

39. Blom DJ, Hala T, Bolognese M, Lillestol MJ, Toth PD, Burgess L, et al. A 52-week placebo-controlled trial of evolocumab in hyperlipidemia. N Engl J Med. 2014;370:1809–19.

40. Navarese EP, Kolodziejczak M, Schulze V, Gurbel PA, Tantry U, Lin Y, et al. Effects of Proprotein Convertase Subtilisin/Kexin Type 9 Antibodies in Adults With Hypercholesterolemia. Ann Intern Med. 2015;163(1):40-51

41. Mihaylova B, Emberson J, Blackwell L, Keech A, Simes J, Barnes EH, et al. The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. Lancet. 2012;380(9841):581–90.

42. O'Donnell H, McCullagh L, Walsh C, Barry M. Coronary Heart Disease Mortality vs Cardiovascular Mortality Comments on Effects of Proprotein Convertase Subtilisin/Kexin Type 9 Antibodies in Adults With Hypercholesterolemia A Systematic Review and Meta Analysis. Ann Intern Med. 2017. Available from https://www.acpjournals.org/doi/10.7326/M14-2957?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed Date accessed 31 March 2020.

43. Kereiakes DJ, Robinson JG, Cannon CP, Lorenzato C, Pordy R, Chaudhari U. Efficacy and safety of the proprotein convertase subtilisin/kexin type 9 inhibitor alirocumab among high cardiovascular risk patients on maximally tolerated statin therapy: the ODYSSEY COMBO I study. Am Heart J. 2015;169(6):906-15.e13

44. Schmidt AF, Pearce LS, Wilkins JT, Overington JP, Hingorani AD, Casas JP. PCSK9 monoclonal antibodies for the primary and secondary prevention of cardiovascular disease. 2017;4(4):CD011748.

45. Guedeney P, Giustino G, Sorrentino S, Claessen BE, Camaj A, Kalkman DN, et al. Efficacy and safety of alirocumab and evolocumab: a systematic review and meta-analysis of randomized controlled trials. Eur Heart J. 2019;ehz430.

46. Lipinski MJ, Benedetto U, Escarcega RO, Biondi-Zoccai G, Lhermusier T, Baker NC, et al. The impact of proprotein convertase subtilisin-kexin type 9 serine protease inhibitors on lipid levels and outcomes in patients with primary hypercholesterolaemia: a network meta-analysis. Eur Heart J. 2016;37(6):536–45.

47. Giugliano RP, Mach F, Zavitz K, Kurtz C, Im K, Kanevsky E, et al. Cognitive Function in a Randomized Trial of Evolocumab. N Engl J Med. 2017;377(7):633–43.

48. ClinicalTrials.gov. Evaluating Effect of the Study Drug Praluent (Alirocumab) on Neurocognitive Function When Compared to Placebo NCT02957682 [Internet]. 2016. Available from https://clinicaltrials.gov/ct2/show/NCT02957682 Date accessed 31 March 2020.

49. Toth PP, Worthy G, Gandra SR, Sattar N, Bray S, Cheng LI, et al. Systematic Review and Network Meta-Analysis on the Efficacy of Evolocumab and Other Therapies for the Management of Lipid Levels in Hyperlipidemia. J Am Heart Assoc. 2017;6(10):e005367.

50. Du H, Li X, Su N, Li L, Hao X, Gao H, et al. Proprotein convertase subtilisin/kexin 9 inhibitors in reducing cardiovascular outcomes: a systematic review and meta-analysis. Heart. 2019;105:1149–59.

51. Khan SU, Riaz H, Rahman H, Khan MU, Khan MS, Alkhouli M, et al. Association of baseline LDL-C with total and cardiovascular mortality in patients using proprotein convertase subtilisin-kexin type 9 inhibitors: A systematic review and meta-analysis. J Clin Lipidol. 2019;13(4):538–49.

52. Khan SU, Talluri S, Riaz H, Rahman H, Nasir F, Bin Riaz I, et al. A Bayesian network meta-analysis of PCSK9 inhibitors, statins and ezetimibe with or without statins for cardiovascular outcomes. Eur J Prev Cardiol. 2018;25(8):844–53.

53. Zhao Z, Du S, Shen S, Luo P, Ding S, Wang G, et al. Comparative efficacy and safety of lipid-lowering agents in patients with hypercholesterolemia: A frequentist network meta-analysis. Medicine. 2019;98(6):e14400

54. Guedeney P, Sorrentino S, Giustino G, Chapelle C, Claessen B, Ollier E, et al. P5367 Indirect comparison of the safety and efficacy of alirocumab and evolocumab: from a comprehensive meta-analysis of 30 randomized controlled trials. Eur Heart J. 2019;40(Suppl_1):ehz746.0332.

55. Cao YX, Liu HH, Dong QT, Li S, Li JJ. Effect of proprotein convertase subtilisin/kexin type 9 (PCSK9) monoclonal antibodies on new-onset diabetes mellitus and glucose metabolism: A systematic review and meta-analysis. Diabetes Obes Metab. 2018;20(6):1391–8.

56. de Carvalho LSF, Campos AM, Sposito AC. Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9) Inhibitors and Incident Type 2 Diabetes: A Systematic Review and Meta-analysis With Over 96,000 Patient-Years. Diabetes Care. 2018;41(2):364–7.

57. Cao YX, Li S, Liu HH, Li JJ. Impact of PCSK9 monoclonal antibodies on circulating hs-CRP levels: a systematic review and meta-analysis of randomised controlled trials. BMJ Open. 2018;8:e022348.

58. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane; 2020. Available from www.training.cochrane.org/handbook. Date accessed 31 March 2020.

59. Moher D, Liberati A, Tetzlaff J, Altman DG, Altman D, Antes G, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. Ann Intern Med. 2009;151(4):264-69.

60. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.

61. Hartling L, Hamm MP, Milne A, Vandermeer B, Santaguida PL, Ansari M, et al. Testing the Risk of Bias tool showed low reliability between individual reviewers and across consensus assessments of reviewer pairs. J Clin Epidemiol. 2013;66(9):973–81.

62. Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health. 2019;22(4):153–60.

63. R Core Team. R: A language and environment for statistical computing [Software]. Vienna, Austria: R Foundation for Statistical Computing; 2020.

64. Microsoft Corporation. Microsoft Excel 2016 [Software]. United States:Microsoft; 2020.

65. Veroniki AA, Jackson D, Bender R, Kuss O, Langan D, Higgins JPT, et al. Methods to calculate uncertainty in the estimated overall effect size from a random-effects meta-analysis. Res Synth Methods. 2019;10(1):23–43.

66. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177–88.

67. Gonnermann A, Framke T, Großhennig A, Koch A. No solution yet for combining two independent studies in the presence of heterogeneity. Stat Med. 2015;34(16):2476–80.

68. Hartung J, Knapp G. A refined method for the meta-analysis of controlled clinical trials with binary outcome. Stat Med. 2001;20(24):3875–89.

69. Friede T, Röver C, Wandel S, Neuenschwander B. Meta-analysis of two studies in the presence of heterogeneity with applications in rare diseases. Biometrical J. 2017;59(4):658–71.

70. Langan D, Higgins JPT, Simmonds M. Comparative performance of heterogeneity variance estimators in meta-analysis: a review of simulation studies. Res Synth Methods. 2017;8(2):181–98.

71. Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc. 2009;172(1):137-159.

72. Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. BMJ. 2011;342(7804):964–7.

73. Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Itzler R, Cappelleri JC, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 2. Value Health. 2011;14(4):429–37.

74. Dias S, Caldwell DM. Network meta-analysis explained. Arch Dis Child Fetal Neonatal Ed. 2019;104(1):F8-F12.

75. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol. 1997;50(6):683–91.

76. Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. Value Health. 2011;14(4):417–28.

77. Schünemann H, Brożek J, Guyatt G, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. The GRADE Working Group, 2013. Available from guidelinedevelopment.org/handbook. Date accessed 31 March 2020.

78. GRADEpro GDT. GRADEpro Guideline Development Tool [Software]. McMaster University, 2020 (developed by Evidence Prime, Inc.). Available from gradepro.org. Date accessed 14 Jan 2020.

79. Murphy SA, Pedersen TR, Gaciong ZA, Ceska R, Ezhov M V, Connolly DL, et al. Effect of the PCSK9 Inhibitor Evolocumab on Total Cardiovascular Events in Patients with Cardiovascular Disease: a Prespecified Analysis from the FOURIER Trial. JAMA Cardiol. 2019;4(7):613-619.

80. Charytan DM, Sabatine MS, Pedersen TR, Im K, Park JG, Pineda AL, et al. Efficacy and Safety of Evolocumab in Chronic Kidney Disease in the FOURIER Trial. J Am Coll Cardiol. 2019;73(23):2961–70.

81. Martin SS, Giugliano RP, Murphy SA, Wasserman SM, Stein EA, Ceska R, et al. Comparison of low-density lipoprotein cholesterol assessment by Martin/Hopkins estimation, friedewald estimation, and preparative ultracentrifugation insights from the FOURIER trial. JAMA Cardiol. 2018;3(8):749-753.

82. Ramirez L, Saver JL. Tetrad therapy with antiplatelets, antihypertensives, high intensity statins and a PCSK9 inhibitor reduces vascular events in patients with a history of stroke. Stroke. 2018;49(Supplement. 1):ATP390.

83. Sabatine MS, Leiter LA, Wiviott SD, Giugliano RP, Deedwania P, De Ferrari GM, et al. Cardiovascular safety and efficacy of the PCSK9 inhibitor evolocumab in patients with and without diabetes and the effect of evolocumab on glycaemia and risk of new-onset diabetes: a prespecified analysis of the FOURIER randomised controlled trial. Lancet Diabetes Endocrinol. 2017;5(12):941–50.

84. Sabatine MS, De Ferrari GM, Giugliano RP, Huber K, Lewis BS, Ferreira J, et al. Clinical Benefit of Evolocumab by Severity and Extent of Coronary Artery Disease: Analysis From FOURIER . Circulation. 2018;138(8):756-66.

85. Mach F, Giugliano R, Zavitz K, Keech A, Pedersen T, Sabatine M, et al. Primary results of EBBINGHAUS, a cognitive study of patients enrolled in the Fourier trial. Eur stroke J. 2017;1(Supplement 1):16.

86. Sabatine MS, De Ferrari GM, Giugliano RP, Huber K, Lewis BS, Ferreira J, et al. Clinical Benefit of Evolocumab by Severity and Extent of Coronary Artery Disease. Circulation. 2018;138(8):756-766.

87. Murphy SA, Pedersen TR, Gaciong ZA, Ceska R, Ezhov M V, Connolly D, et al. Reduction in Total Cardiovascular Events With the PCSK9 Inhibitor Evolocumab in Patients With Cardiovascular Disease in the FOURIER Trial. Circulation. 2017;136(Suppl_1):A15100.

88. Wiviott SD, Giugliano RP, Morrow DA, De Ferrari GM, Lewis BS, Huber K, et al. Characterization of types and sizes of myocardial infarction reduced with evolocumab in FOURIER. Circulation. 2017;136(Supplement 1):A16714.

89. Bohula EA, Giugliano R, Leiter L, Verma S, Park JG, Sever P, et al. Residual inflammatory and cholesterol risk in the FOURIER trial. J Am Coll Cardiol. 2018;71(11 Supplement):A2.

90. Qamar A, Giugliano RP, Keech AC, Kuder JF, Murphy SA, Kurtz CE, et al. Interindividual Variation in Low-Density Lipoprotein Cholesterol Level Reduction With Evolocumab An Analysis of FOURIER Trial Data. JAMA Cardiol. 2019;4(1):59–63.

91. Mach F, Sabatine M, Giugliano R, Keech A, Honarpour N, Wasserman SM, et al. Further cardiovascular outcomes research with PCSK9 inhibition in subjects with eelevated risk, the FOURIER trial. Kardiovaskulare medizin. 2017;20(5 Supplementum 27):S115

92. Marston NA, Kamanu FK, Nordio F, Gurmu Y, Roselli C, Sever PS, et al. Predicting Benefit From Evolocumab Therapy in Patients With Atherosclerotic Disease Using a Genetic Risk Score: Results From the FOURIER Trial. Circulation. 2019;141(8):616-23

93. Giugliano RP, Keech A, Murphy SA, Huber K, Tokgozoglu SL, Lewis BS, et al. Clinical Efficacy and Safety of Evolocumab in High-Risk Patients Receiving a Statin: secondary Analysis of Patients With Low LDL Cholesterol Levels and in Those Already Receiving a Maximal-Potency Statin in a Randomized Clinical Trial. JAMA Cardiol. 2017;2(12):1385-1391.

94. Pedersen TR, Giullano RP, Sever P, Keech A, Wasserman SM, Honarpour N. The further cardiovascular outcomes research with PCSK9 inhibitor in subjects with elevated risk (FOURIER) trial: effect of evolocumab on cerebrovascular disease. Eur stroke J. 2017;2(Suppl 1):485.

95. International Clinical Trials Registry. Study assessing impact of additional reduction of bad cholesterol on heart related diseases with AMG 145 when used in combination with cholesterol lowering medications CTRI/2014/01/004324 [Intenet]. 2014. Available from

http://www.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2014/01/004324 Date accessed 31 March 2020.

96. Sabatine MS, Giugliano RP, Keech A, Honarpour N, Wang H, Liu T, et al. Rationale and design of the Further cardiovascular OUtcomes Research with PCSK9 Inhibition in subjects with Elevated Risk trial. Am Hear J. 2016;173:94–101.

97. Sabatine MS, Leiter LA, Wiviott SD. Questioning the safety and benefits of evolocumab – Authors' reply. Lancet Diabetes Endocrinol. 2018;6(1):11-12.

98. Bonaca MP, Nault P, Giugliano RP, Keech AC, Pineda A L, Kanevsky E, et al. Low-Density Lipoprotein Cholesterol Lowering With Evolocumab and Outcomes in Patients With Peripheral Artery Disease: insights From the FOURIER Trial (Further Cardiovascular Outcomes Research With PCSK9 Inhibition in Subjects With Elevated Risk). Circulation. 2018;137(4):338-350.

99. Sever PS, Gouni-Berthold I, Keech A, Giugliano R, Pedersen T, Wasserman S, et al. Benefit of LDL-C lowering with evolocumab on cardiovascular outcomes by age & sex: An analysis of the FOURIER trial. Eur Heart J. 2018;39:1041–2.

100. Nault P, Bonaca M, Giugliano RP, Honarpour N, Keech AC, Sever PS, et al. Risk of Major Adverse Limb Events and Benefits of Evolocumab in Patients With Peripheral Artery Disease by History of Prior Peripheral Revascularization. J Vasc Surg. 2019;69(6):e195.

101. Giugliano RP, Mach F, Zavitz K, Kurtz C, Schneider J, Wang H, et al. Design and rationale of the EBBINGHAUS trial: a phase 3, double-blind, placebo-controlled, multicenter study to assess the effect of evolocumab on cognitive function in patients with clinically evident cardiovascular disease and receiving statin background lipid-lowering therapy—A cognitive study of patients enrolled in the FOURIER trial. Clin Cardiol. 2017;40(2):59–65.

102. Giugliano RP, Pedersen TR, Park J-GG, De Ferrari GM, Gaciong ZA, Ceska R, et al. Clinical efficacy and safety of achieving very low LDL-cholesterol concentrations with the PCSK9 inhibitor evolocumab: a prespecified secondary analysis of the FOURIER trial. Lancet. 2017;390(10106):1962–71.

103. Giugliano RP, Sabatine MS, Ott BR. Cognitive Function in a Randomized Trial of Evolocumab REPLY. N Engl J Med. 2017;377(20):1997.

104. Sabatine M, Giugliano Robert P, Pedersen TR. Evolocumab in Patients with Cardiovascular Disease- Author's Response. N Engl J Med. 2017;377(8):785–8.

105. Bohula EA, Giugliano RP, Leiter LA, Verma S, Park JG, Sever PS, et al. Inflammatory and Cholesterol Risk in the FOURIER Trial. Circulation. 2018;138(2):131–40.

106. Bonaca MP, Sabatine MS. Response by Bonaca and Sabatine to Letters Regarding Article, "Low-Density Lipoprotein Cholesterol Lowering With Evolocumab and Outcomes in Patients With Peripheral Artery Disease: Insights From the FOURIER Trial (Further Cardiovascular Outcomes Research. Circulation. 2018;138(2):222–3.

107. Committee for Medicinal Products for Human Use (CHMP). Assessment Report Repatha. London: EMA;2018. Available from: https://www.ema.europa.eu/en/documents/variation-report/repatha-h-c-3766-ii-0017-g-epar-assessment-report-variation_en.pdf Date accessed14 January 2020.

108. O'Donoghue M, Giugliano R, Keech A, Kanevsky E, Im K, Pineda AL, et al. Lipoprotein(a), PCSK9 Inhibition and cardiovascular risk: Insights from the FOURIER trial. Atherosclerosis. 2018;275:e9–10.

109. Gencer B, Mach F, Guo J, Im K, Ruzza A, Wang H, et al. Abstract 13833: Impact of Lowering LDL-C With Evolocumab on Everyday Cognition in Participants From the FOURIER Trial. Circulation. 2019;140(Suppl_1):A13833

110. Gencer B, Mach F, Murphy SA, Ferrari GM De, Huber K, Lewis BS, et al. Evolocumab and Cardiovascular Outcomes in Patients With Recent Myocardial Infarction: Analysis From FOURIER. Circulation. 2019;140(Suppl_1):A1387.

111. Keech AC, Sever P, Jiang LX, Hirayama A, Lu C, Tay L, et al. Effiacy and Safety of Long-term Evolocumab Use in Asian versus Other Subjects: The FOURIER Trial. J Am Coll Cardiol. 2019;73(9):A189.

112. ClinicalTrials.gov. Further Cardiovascular Outcomes Research With PCSK9 Inhibition in Subjects With Elevated Risk (FOURIER) NCT01764633 [Internet]. 2019. Available from https://clinicaltrials.gov/ct2/show/NCT01764633 Date accessed 14 January 2020.

113. International Clinical Trial Registry Platform. A Study Assessing the Impact of Additional LDL-Cholesterol Reduction on Major Cardiovascular Events when Evolocumab (AMG 145) is used in combination With Statin Therapy In Patients with Clinically Evident Cardiovascular Disease EUCTR2012-001398-97-BE [Internet]. 2012. Available from http://www.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-001398-97-BE Date accessed 31 March 2020.

114. Qamar A, Giugliano RP, Keech AC, Kuder JF, Murphy SA, Liu T, et al. Consistency of LDL-C reduction with evolocumab: An analysis from FOURIER. J Am Coll Cardiol. 2018;71(11):1727.

115. Bonaca MP, Nault P, Giugliano RP, Keech AC, Pineda AL, Kanevsky E, et al. Evolocumab and outcomes in patients with peripheral artery disease. Circulation. 2017;136:e451-e452.

116. Bohula EA, Morrow DA, Pedersen TR, Kanevsky E, Murphy SA, Giugliano RP, et al. Atherothrombotic risk stratification and magnitude of benefit of evolocumab in FOURIER. Circulation. 2017;136(Supplement 1):A20183.

117. Bonaca MP, Giugliano RP, Nault P, Scirica BM, Pedersen TR, Bhatt DL, et al. Patient selection for therapies to prevent major adverse limb events. Circulation. 2018;138(Supplement 1):A16771.

118. Correction to: Effect of Alirocumab on Mortality After Acute Coronary Syndromes: An Analysis of the ODYSSEY OUTCOMES Randomized Clinical Trial. Circulation. 2019;140(4):e171.

119. Goodman SG, Aylward PE, Szarek M, Chumburidze V, Bhatt DL, Bittner VA, et al. Effects of Alirocumab on Cardiovascular Events After Coronary Bypass Surgery. J Am Coll Cardiol. 2019;74(9):1177–86.

120. Goodman SG, Aylward P, Szarek M, Chumburidze V, Bhatt DL, Bittner V, et al. Patients With Acute Coronary Syndrome, Elevated Atherogenic Lipoproteins, And Prior Coronary Artery Bypass Grafting Derive Large Absolute Benefit From Alirocumab: Insights From The ODYSSEY OUTCOMES Trial. J Am Coll Cardiol. 2019;73(9 Suppl_1):30.

121. Goodman SG, Schwartz G, Bhatt DL, Bittner V, Diaz R, Hanotin C, et al. Use of high-intensity statin therapy post-acute coronary syndrome in the ongoing odyssey outcomes trial of alirocumab, a proprotein convertase subtilisin/kexin type 9 monoclonal antibody, versus placebo: Interim baseline data. J Am Coll Cardiol. 2017;69(11):153.

122. Jukema JW, Szarek M, Roe M, Bhatt DL, Bittner V, de Silva HA, et al. Post-acute coronary syndrome patients with polyvascular disease derive large absolute benefit from alirocumab: ODYSSEY OUTCOMES. J Am Coll Cardiol. 2019;73(9 Supple_1):2034.

123. Jukema JW, Szarek M, Zijlstra LE, de Silva HA, Bhatt DL, Bittner VA, et al. Alirocumab in Patients With Polyvascular Disease and Recent Acute Coronary Syndrome: ODYSSEY OUTCOMES Trial. J Am Coll Cardiol. 2019;74(9):1167–76.

124. Jukema JW, Szarek M, Zijlstra L, Bhatt DL, Bittner VA, Diaz R, et al. Effect of Low-density Lipoprotein Cholesterol Lowering With Alirocumab on Stroke in ODYSSEY OUTCOMES. Circulation. 2019;140(Suppl_1):A11149.

125. Jukema JW, Zijlstra LE, Bhatt DL, Bittner VA, Diaz R, Drexel H, et al. Effect of Alirocumab on Stroke in ODYSSEY OUTCOMES. Circulation. 2019;140(25):2054–62.

126. Labos C, Brophy JM, Sniderman A, Thanassoulis G. Mortality Benefit of Alirocumab: A Bayesian Perspective. J Am Heart Assoc. 2019;8(20): e013170.

127. Moriarty PM, Steg PG, Gorby LK, Zeiher AM, White HD, Sourdille T, et al. Relation of Red Blood Cell Distribution Width to Risk of Major Adverse Cardiovascular Events, Death, and Effect of Alirocumab After Acute Coronary Syndrome. Circulation. 2019;140(Suppl_1):A12751.

128. O'Donoghue ML, Fazio S, Giugliano RP, Stroes ESG, Kanevsky E, Gouni-Berthold I, et al. Lipoprotein(a), PCSK9 Inhibition, and Cardiovascular Risk. Circulation. 2019;139(12):1483-1492.

129. Correction to Lancet Diabetes Endocrinol 2019; 7: 618–28. Lancet Diabetes Endocrinol. 2019;7(9):e21.

130. Pordy R, White H, Schwartz G, Alings M, Bhatt D, Bittner V, et al. Alirocumab Reduces Total Nonfatal Cardiovascular and Fatal Events in the ODYSSEY OUTCOMES Trial. J Clin Lipidol. 2019;13(3):e54–5.

131. Ray KK, Colhoun HM, Szarek M. Effects of alirocumab on cardiovascular and metabolic outcomes after acute coronary syndrome in patients with or without diabetes: a prespecified analysis of the ODYSSEY OUTCOMES randomised controlled trial. Lancet Diabetes Endocrinol. 2019 Aug;7(8):618-628.

132. Ray KK, Colhoun H, Szarek M, Baccara-Dinet M, Bhatt DL, Bittner V, et al. Alirocumab and cardiovascular outcomes in patients with acute coronary syndrome (ACS) and diabetes-prespecified analyses of odyssey outcomes. Diabetes. 2018;67 (Suppl_1) :LB2-LB3.

133. Roe MT, Szarek M, Li QH, Bhatt DL, Bittner V, Goodman SG, et al. Efficacy of alirocumab treatment after acute coronary syndrome according to new ACC/AHA guidelines for lipid-lowering therapy. Eur Heart J. 2019;40(Supplement_1):2474.

134. Schwartz GG, Bessac L, Berdan LG, Bhatt DL, Bittner V, Diaz R, et al. Effect of alirocumab, a monoclonal antibody to PCSK9, on long-term cardiovascular outcomes following acute coronary syndromes: rationale and design of the ODYSSEY outcomes trial. Am Heart J. 2014;168(5):682-689.

135. Committee for Medicinal Products for Human Use (CHMP). Assessment Report Praluent. Amsterdam: EMA;2019. Available from: https://www.ema.europa.eu/en/documents/variation-report/praluent-h-c-003882-ii-0042-epar-assessment-report-variation_en.pdf Date accessed14 January 2020.

136. Bhatt DL, Briggs A, Reed SD, Annemans L, Szarek M, Bittner VA, et al. Cost-effectiveness of alirocumab based on evidence from a large multinational outcome trial: the ODYSSEY OUTCOMES economics study. Circulation. 2018;138(25):E753-754.

137. Bittner V, Szarek M, Aylward PE, Bhatt DL, Diaz R, Fras Z, et al. Lp(A) and cardiovascular outcomes: An analysis from the odyssey outcomes trial. Atheroscler Suppl. 2018;32:24–5.

138. ClinicalTrials.gov. ODYSSEY Outcomes: Evaluation of Cardiovascular Outcomes After an Acute Coronary Syndrome During Treatment With Alirocumab NCT01663402 [Internet]. 2019. Available from: https://clinicaltrials.gov/ct2/show/NCT01663402 Date accessed 31 March 2020.

139. Damask A, Steg PG, Schwartz GG, Szarek M, Hagström E, Badimon L, et al. Patients with High Genome-Wide Polygenic Risk Scores for Coronary Artery Disease May Receive Greater Clinical Benefit from Alirocumab Treatment in the Odyssey Outcomes Trial. Circulation. 2020;141(8):624-636.

140. Damask A, Steg PG, Schwartz GG, Szarek M, Hagstrom E, Badimon L, et al. Abstract 13322: Patients With a High Genetic Risk Score for Coronary Artery Disease May Receive Greater Benefit From Alirocumab Treatment in the ODYSSEY OUTCOMES Trial. Circulation. 2019;140(Suppl_1):A13322.

141. Diaz R, Li QH, Bhatt DL, Bittner VA, Baccara-Dinet MT, Goodman SG, et al. Effect of alirocumab on recurrent cardiovascular events after acute coronary syndrome, according to the intensity of background statin treatment. Eur Heart J. 2019;40(Supp_1):2475.

142. Kastelein JJP, Kereiakes DJ, Cannon CP, Bays HE, Minini P, Lee L V, et al. Effect of alirocumab dose increase on LDL lowering and lipid goal attainment in patients with dyslipidemia. Coron Artery Dis. 2017;28(3):190-197.

143. Leiter LA, Zamorano JL, Bujas-Bobanovic M, Louie MJ, Lecorps G, Cannon CP, et al. Lipid-lowering efficacy and safety of alirocumab in patients with or without diabetes: a sub-analysis of ODYSSEY COMBO II. Diabetes Obes Metab. 2017;19(7):989-996.

144. Toth PP, Dwyer JP, Cannon CP, Colhoun HM, Rader DJ, Upadhyay A, et al. Efficacy and safety of lipid lowering by alirocumab in chronic kidney disease. Kidney Int. 2018;93(6):1397–408.

145. Krempf M, Bergeron J, Elassal J, Minini P, Miller K, Kastelein JJP. Efficacy of alirocumab according to background statin intensity and other lipid-lowering therapy in heterozygous familial hypercholesterolemia or high CV risk populations: phase 3 sub-group analyses. Atherosclerosis. 2015;241(Suppl_1):e21.

146. Catapano AL, Lee L V, Louie MJ, Thompson D, Bergeron J, Krempf M. Efficacy of alirocumab according to background statin type and dose: pooled analysis of 8 ODYSSEY Phase 3 clinical trials. Sci Rep. 2017;7:45788.

147. Ferdinand KC, Jacobson TA, Williams KA, A TK, Elassal J, Thompson DE, et al. Alirocumab efficacy and safety by race and ethnicity: Analysis from three odyssey phase 3 trials. Circulation. 2017;136(Suppl_1):A15748.

148. Ferdinand KC, Jacobson TA, Koren A, Elassal J, Thompson D, Deedwania P. Alirocumab efficacy and safety by race and ethnicity: Analysis from 3 ODYSSEY phase 3 trials. J Clin Lipidol. 2019;13(4):586–93

149. Harvey PD, Sabbagh MN, Harrison JE, Ginsberg HN, Chapman MJ, Manvelian G, et al. No evidence of neurocognitive adverse events associated with alirocumab treatment in

3340 patients from 14 randomized Phase 2 and 3 controlled trials: A meta-analysis of individual patient data. Eur Heart J. 2018;39(5):374–81.

150. Cannon CP, Cariou B, Blom D, McKenney JM, Lorenzato C, Pordy R, et al. Efficacy and safety of alirocumab in high cardiovascular risk patients with inadequately controlled hypercholesterolaemia on maximally tolerated doses of statins: the ODYSSEY COMBO II randomized controlled trial. Eur Heart J. 2015;36(19):1186–94.

151. El Shahawy M, Cannon C, Blom D, McKenney J, Cariou B, Lecorps G, et al. Alirocumab versus ezetimibe over 104 weeks in individuals with hypercholesterolemia and high cardiovascular risk: final results from odyssey combo II. J Clin Lipidol. 2016;10(3):717-718.

152. ClinicalTrials.gov. Efficacy and Safety of Alirocumab (SAR236553/REGN727) Versus Ezetimibe on Top of Statin in High Cardiovascular Risk Patients With Hypercholesterolemia (ODYSSEY COMBO II) NCT01644188. [Internet]. 2016. Available from https://clinicaltrials.gov/ct2/show/NCT01644188 Date accessed 31 March 2020

153. Vallejo-Vaz AJ, Roth EM, Hovingh GK, Louie MJ, Valcheva V, Letierce A, et al. Predictive factors for alirocumab dose increase in patients with hypercholesterolaemia and high cardiovascular risk: from the ODYSSEY COMBO I and II trials. Eur Heart J. 2017;38(Suppl_1):P6238.

154. Farnier M, Colhoun HM, Sasiela WJ, Edelberg J, Asset G, Robinson JG. Treatment adherence to the proprotein convertase subtilisin/kexin type 9 inhibitor alirocumab in six ODYSSEY phase 3 clinical studies with treatment duration of 52 - 104 weeks. Atherosclerosis. 2016;252:e254.

155. Robinson JG, Farnier M, Kastelein JJP, Roth EM, Taskinen MR, Colhoun HM, et al. Relationship between alirocumab, PCSK9, and LDL-C levels in four phase 3 ODYSSEY trials using 75 and 150 mg doses. J Clin Lipidol. 2019;13:979–88.

156. Kastelein JJ, Kereiakes DJ, Cannon CP, Bays HE, Minini P, Lee L V, et al. Additional LDL-C reduction achieved with alirocumab dose increase on background statin. Circulation. 2015;132(Supp_1):A17099.

157. Jones PH, Bays HE, Chaudhari U, Pordy R, Lorenzato C, Miller K, et al. Safety of Alirocumab (A PCSK9 Monoclonal Antibody) from 14 Randomized Trials. Am J Cardiol. 2016;118(12):1805-1811.

158. ClinicalTrials.gov. Effect of Evolocumab in Patients at High Cardiovascular Risk Without Prior Myocardial Infarction or Stroke (VESALIUS-CV) NCT03872401 [Internet]. 2019. Available from https://www.clinicaltrials.gov/ct2/show/NCT03872401 Date accessed 14 January 2020.

159. ClinicalTrials.gov. Effect of Evolocumab on Saphenous Vein Graft Patency Following Coronary Artery Bypass Surgery (NEWTON-CABG) NCT03900026 [Internet]. 2019. Available from https://clinicaltrials.gov/ct2/show/NCT03900026 Date accessed 14 January 2020.

160. Cannon CP, Cariou B, Blom D, McKenney JM, Lorenzato C, Pordy R, et al. Efficacy and safety of alirocumab in high cardiovascular risk patients with inadequately controlled hypercholesterolaemia on maximally tolerated doses of statins: The ODYSSEY COMBO II randomized controlled trial. Eur Heart J. 2015;36(19):1186–94.

161. Joulain F, Bechet M, Valcheva V, Gorcyca K, Gooch K. Patient Utility Estimates In Patients At High Cardiovascular Risk. Value Health. 2015;18(7):A398.

162. Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, Bhala N, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet. 2010;376(9753):1670–81.

163. Cholesterol Treatment Trialists' (CTT) Collaboration, Fulcher J, O'Connell R, Voysey M, et al. Efficacy and safety of LDL-lowering therapy among men and women: meta-analysis of individual data from 174 000 participants in 27 randomised trials. Lancet. 2015;385(9976):1397–405.

164. Wisløff T, Mundal LJ, Retterstøl K, Igland J, Kristiansen IS. Economic evaluation of lipid lowering with PCSK9 inhibitors in patients with familial hypercholesterolemia: Methodological aspects. Atherosclerosis. 2019;287:140–6.

165. Ginsberg HN, Rader DJ, Raal FJ, Guyton JR, Baccara-Dinet MT, Lorenzato C, et al. Efficacy and Safety of Alirocumab in Patients with Heterozygous Familial Hypercholesterolemia and LDL-C of 160 mg/dl or Higher. Cardiovasc Drugs Ther. 2016;30(5):473-483.

166. Dixon DL, Virani SS. Mortality reduction with PCSK9 inhibition: A case of cautious optimism. J Clin Lipidol. 2019;13(5):712–3.

167. Steg PG, Szarek M, Bhatt DL, Bittner VA, Bregeault M-F, Dalby AJ, et al. Effect of Alirocumab on Mortality After Acute Coronary Syndromes. Circulation. 2019;140(2):103–12.

168. Guedeney P, Giustino G, Sorrentino S, Claessen BE, Camaj A, Kalkman DN, et al. Efficacy and safety of alirocumab and evolocumab: a systematic review and meta-analysis of randomized controlled trials. 2019. Jul 3;ehz430. doi: 10.1093/eurheartj/ehz430. Online ahead of print.

169. Myers KD, Farboodi N, Mwamburi M, Howard W, Staszak D, Gidding S, et al. Effect of Access to Prescribed PCSK9 Inhibitors on Cardiovascular Outcomes. Circ Cardiovasc Qual Outcomes. 2019;12(8):e005404.

170. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008.

171. McCullagh L, Barry M. The Pharmacoeconomic Evaluation Process in Ireland. Pharmacoeconomics. 2016 Dec;34(12):1267–76.

172. National Institute for Health and Care Excellence (NICE). Budget Impact Test [Internet]. 2018. Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance/budget-impact-test Date accessed 31 March 2020.

173. National Institute for Health and Care Excellence (NICE). TA358 Tolvaptan for treating autosomal dominant polycystic kidney disease: London: NICE; 2015. Available from https://www.nice.org.uk/guidance/ta358/resources/tolvaptan-for-treating-autosomal-dominant-polycystic-kidney-disease-pdf-82602675026629 Date accessed 31 March 2020.

174. Sculpher M. Subgroups and heterogeneity in cost-effectiveness analysis. Pharmacoeconomics. 2008;26(9):799–806.

175. Chandra A, Garthwaite C. The Economics of Indication-Based Drug Pricing. N Engl J Med. 2017;377(2):103–6.

176. Dakin H, Gray A. Decision Making for Healthcare Resource Allocation: Joint v. Separate Decisions on Interacting Interventions. Med Decis Mak. 2018;38(4):476–86.

177. Grimm SE, Dixon S, Stevens JW. When Future Change Matters: Modeling Future Price and Diffusion in Health Technology Assessments of Medical Devices. Value Health. 2016;19(6):720–6.

178. Drummond MF, Sculphur MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. 4th ed. New York: Oxford University Press; 2015.

179. Johannesson M, Weinstein MC. On the decision rules of cost-effectiveness analysis. J Health Econ. 1993;12(4):459–67.

180. Persson U, Norlin JM. Multi-indication and Combination Pricing and Reimbursement of Pharmaceuticals: Opportunities for Improved Health Care through Faster Uptake of New Innovations. Appl Health Econ Health Policy. 2018;16(2):157–65.

181. Paulden M. Strategic behaviour and the cost-effectiveness threshold – a new conceptual model. In: Pandey H, Paulden M, McCabe C, editors. Theoretical models of the cost-effectiveness threshold, value assessment, and health care system sustainability. Edmonton (AB): Institute of Health Economics; 2018.

182. Ramaekers BLT, Joore MA, Grutters JPC. How Should We Deal with Patient Heterogeneity in Economic Evaluation: A Systematic Review of National Pharmacoeconomic Guidelines. Value Health. 2013;16(5):855–62.

183. National Centre for Pharmacoeconomics (NCPE). Full Submission Applicant Template [Internet]. 2019. Available at http://www.ncpe.ie/submission-process/submission-templates/format-of-full-submissions/ Date accessed 31 March 2020.

184. National Institute for Health and Care Excellence (NICE). Position statement on use of the EQ-5D-5L valuation set for England (updated November 2018) [Internet]. 2018. Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l Date accessed 31 March 2020.

185. Rowen D, Azzabi Zouraq I, Chevrou-Severac H, van Hout B. International Regulations and Recommendations for Utility Data for Health Technology Assessment. Pharmacoeconomics. 2017;35(1):11–9.

186. Longworth L, Rowen D. Mapping to Obtain EQ-5D Utility Values for Use in NICE Health Technology Assessments. Value Health. 2013;16(1):202–10.

187. Ara R, Rowen D, Mukuria C. The Use of Mapping to Estimate Health State Utility Values. Pharmacoeconomics. 2017;35(1):57–66.

188. Mukuria C, Rowen D, Harnan S, Rawdin A, Wong R, Ara R, et al. An Updated Systematic Review of Studies Mapping (or Cross-Walking) Measures of Health-Related Quality of Life to Generic Preference-Based Measures to Generate Utility Values. Appl Health Econ Health Policy. 2019;17(3):295–313.

189. Brazier J, Ara R, Azzabi I, Busschbach J, Chevrou-Séverac H, Crawford B, et al. Identification, Review, and Use of Health State Utilities in Cost-Effectiveness Models: An ISPOR Good Practices for Outcomes Research Task Force Report. Value Health. 2019;22(3):267–75.

190. Ara R, Peasgood T, Mukuria C, Chevrou-Severac H, Rowen D, Azzabi-Zouraq I, et al. Sourcing and Using Appropriate Health State Utility Values in Economic Models in Health Care. Pharmacoeconomics. 2017;35(1):7–9.

191. Bojke L, Manca A, Asaria M, Mahon R, Ren S, Palmer S. How to Appropriately Extrapolate Costs and Utilities in Cost-Effectiveness Analysis. Pharmacoeconomics. 2017;1–10.

192. Fogarty E, Walsh C, Adams R, McGuigan C, Barry M, Tubridy N. Relating health-related Quality of Life to disability progression in multiple sclerosis, using the 5-level EQ-5D. Mult Scler J. 2013;19(9):1190–6.

193. Adams R, Walsh C, Veale D, Bresnihan B, FitzGerald O, Barry M. Understanding the Relationship between the EQ-5D, SF-6D, HAQ and Disease Activity in Inflammatory Arthritis. Pharmacoeconomics. 2010;28(6):477–87.

194. Gillespie P, O'Shea E, Murphy AW, Byrne MC, Byrne M, Smith SM, et al. The cost-effectiveness of the SPHERE intervention for the secondary prevention of coronary heart disease. Int J Technol Assess Health Care. 2010;26(3):263–71.

195. Kenny RA, Whelan BJ, Cronin H, Kamiya Y, Kearney P, O' C, et al. The Design of the Irish Longitudinal Study on Ageing. Dublin: TILDA; 2010. Available at https://tilda.tcd.ie/publications/reports/pdf/Report_DesignReport.pdf. Date accessed13 February 2020.

196. Kearney PM, Cronin H, O'Regan C, Kamiya Y, Savva GM, Whelan B, et al. Cohort Profile: The Irish Longitudinal Study on Ageing. Int J Epidemiol. 2011;40(4):877–84.

197. The Irish Longitudinal Study on Agina (TILDA). Accessing the Data [Internet]. 2019. Available from: https://tilda.tcd.ie/data/accessing-data/ Date accessed 31 March 2020.

198. Longworth L, Rowen D. NICE DSU technical support document 10: the use of mapping methods to estimate health state utility values. Sheffield: ScHARR, University of Sheffield; 2011. Available from http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD-10-mapping-FINAL.pdf Date accessed 31 March 2020.

199. Dakin H, Abel L, Burns R, Yang Y. Review and critical appraisal of studies mapping from quality of life or clinical measures to EQ-5D: an online database and application of the MAPS statement. Health Qual Life Outcomes. 2018;16(1):31.

200. Wailoo AJ, Hernandez-Alava M, Manca A, Mejia A, Ray J, Crawford B, et al. Mapping to Estimate Health-State Utility from Non–Preference-Based Outcome Measures: An ISPOR Good Practices for Outcomes Research Task Force Report. Value Health. 2017;20(1):18–27.

201. Petrou S, Rivero-Arias O, Dakin H, Longworth L, Oppe M, Froud R, et al. Preferred Reporting Items for Studies Mapping onto Preference-Based Outcome Measures: The MAPS Statement. Pharmacoeconomics. 2015;33(10):985–91.

202. Petrou S, Rivero-Arias O, Dakin H, Longworth L, Oppe M, Froud R, et al. The MAPS Reporting Statement for Studies Mapping onto Generic Preference-Based Outcome Measures: Explanation and Elaboration. Pharmacoeconomics. 2015;33(10):993–1011.

203. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. Eur J Health Econ. 2010;11(2):215-25.

204. Dakin H, Abel L, Burns R, Yang Y. HERC Database of Mapping Studies, Version 7.0 [database]. Oxford, UK: HERC; 2019. Available from: https://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies Date accessed 31 March 2020.

205. Dakin H. Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database. Health Qual Life Outcomes. 2013;11(1):151.

206. Ghatnekar O, Eriksson M, Glader E-L. Mapping health outcome measures from a stroke registry to EQ-5D weights. Health Qual Life Outcomes. 2013;11(1):34.

207. Ara R, Kearns B, vanHout BA, Brazier JE. Predicting preference-based utility values using partial proportional odds models. BMC Res Notes. 2014;7:438.

208. Calvert MJ, Freemantle N, Yao G, Cleland JGF, Billingham L, Daubert J-C, et al. Cost-effectiveness of cardiac resynchronization therapy: results from the CARE-HF trial. Eur Heart J. 2005;26(24):2681–8.

209. Chen G, McKie J, Khan MA, Richardson JR. Deriving health utilities from the MacNew Heart Disease Quality of Life Questionnaire. Eur J Cardiovasc Nurs. 2014;14(5):405–15.

210. Edlin R, Tsuchiya A, Brazier J. Mapping the Minnesota living with heart failure questionnaire to the EQ-5D index [Unpublished Manuscript]. 2002.

211. Goldsmith KA, Dyer MT, Buxton MJ, Sharples LD. Mapping of the EQ-5D index from clinical outcome measures and demographic variables in patients with coronary heart disease. Health Qual Life Outcomes. 2010;8(1):54.

212. Longworth L, Buxton MJ, Sculpher M, Smith DH. Estimating utility data from clinical indicators for patients with stable angina. Eur J Health Econ. 2005;6(4):347–53.

213. Longworth L. Estimating Quality Adjusted Life Years Where Health-related Utility Data are Missing [dissertation]. London: Brunel University; 2007.

214. Oddershede L, Andreasen JJ, Ehlers L. Estimation of utility values from visual analog scale measures of health in patients undergoing cardiac surgery. Clinicoecon Outcomes Res. 2014;6:21–7.

215. Rivero-Arias O, Ouellet M, Gray A, Wolstenholme J, Rothwell PM, Luengo-Fernandez R. Mapping the modified Rankin scale (mRS) measurement into the generic EuroQol (EQ-5D) health outcome. Med Decis Making. 2010;30(3):341–54.

216. van Exel NJA, Scholte op Reimer WJM, Koopmanschap MA. Assessment of post-stroke quality of life in cost-effectiveness studies: the usefulness of the Barthel Index and the EuroQoL-5D. Qual Life Res. 2004 Mar;13(2):427–33.

217. Whynes DK, Sprigg N, Selby J, Berge E, Bath PM, Investigators E. Testing for differential item functioning within the EQ-5D. Med Decis Making. 2013;33(2):252–60.

218. Wijeysundera HC, Tomlinson G, Norris CM, Ghali WA, Ko DT, Krahn MD. Predicting EQ-5D utility scores from the Seattle Angina Questionnaire in coronary artery disease: a mapping algorithm using a Bayesian framework. Med Decis Making. 2011;31(3):481–93.

219. Buchholz I, Janssen MF, Kohlmann T, Feng YS. A Systematic Review of Studies Comparing the Measurement Properties of the Three-Level and Five-Level Versions of the EQ-5D. Pharmacoeconomics. 2018;36(6):645-661

220. Hobbins A, Barry L, Kelleher D, Shah K, Devlin N, Goni JMR, et al. Utility Values for Health States in Ireland: A Value Set for the EQ-5D-5L. Pharmacoeconomics. 2018;36(11):1345–53.

221. Szende A, Janssen B, Cabases J. Self-reported population health: an international perspective based on EQ-5D. Dordrecht: Springer; 2014

222. Reddy B, Walsh C, Barry M, Kind P, Adams R. Results of the EQ-5D-3L Valuation Project for Ireland. Value Health. 2015;18(7):A740.

223. Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: results from a general population study. Heal Econ. 1996;5:141–54.

224. Hernandez-Alava M, Wailoo AJ, Ara R. Tails from the Peak District: Adjusted Limited Dependent Variable Mixture Models of EQ-5D Questionnaire Health State Utility Values. Value Health. 2012;15(3):550–61.

225. Sullivan PW. Are utilities bounded at 1.0? Implications for statistical analysis and scale development. Med Decis Mak. 2011 Nov 8;31(6):787–9.

226. Pullenayegum EM, Tarride J-E, Xie F, Goeree R, Gerstein HC, O'Reilly D. Analysis of Health Utility Data When Some Subjects Attain the Upper Bound of 1: Are Tobit and CLAD Models Appropriate? Value Health. 2010;13(4):487–94.

227. Pullenayegum EM, Wong HS, Childs A. Generalized additive models for the analysis of EQ-5D utility data. Med Decis Making. 2013;33(2):244–51.

228. Hyde M, Wiggins RD, Higgs P, Blane DB. A measure of quality of life in early old age: The theory, development and properties of a needs satisfaction model (CASP-19). Aging Ment Health. 2003;7(3):186–94.

229. Radloff LS. The CES-D Scale. Appl Psychol Meas. 1977;1(3):385–401.

230. Snaith RP. The hospital anxiety and depression scale. Health Qual Life Outcomes. 2003;1:29.

231. Bowling A. Mode of questionnaire administration can have serious effects on data quality. J Public Health (Oxf). 2005;27(3):281–91.

232. Davison NJ, Thompson AJ, Turner AJ, Longworth L, McElhone K, Griffiths CEM, et al. Generating EQ-5D-3L Utility Scores from the Dermatology Life Quality Index: A Mapping Study in Patients with Psoriasis. Value Health. 2018;21(8):1010–8.

233. Sullivan PW, Lawrence WF, Ghushchyan V. A National Catalog of Preference-Based Scores for Chronic Conditions in the United States. Med Care. 2005;43(7):736–49.

234. 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.

235. Sullivan PW, Ghushchyan V. Preference-Based EQ-5D Index Scores for Chronic Conditions in the United States. Med Decis Mak. 2006;26(4):410–20.

236. Pullenayegum EM, Tarride JE, Xie F, O'Reilly D. Calculating utility decrements associated with an adverse event: Marginal tobit and CLAD coefficients should be used with caution. Med Decis Mak. 2011;31(6):790–9.

237. Austin PC, Escobar M, Kopec JA. The use of the Tobit model for analyzing measures of health status. Qual Life Res. 2000;9(8):901–10.

238. Gray LA, Hernández Alava M, Wailoo AJ. Development of Methods for the Mapping of Utilities Using Mixture Models: Mapping the AQLQ-S to the EQ-5D-5L and the HUI3 in Patients with Asthma. Value Health. 2018;21(6):748–57.

239. Gray LA, Hernández Alava M. A Command for Fitting Mixture Regression Models for Bounded Dependent Variables Using the Beta Distribution. Stata J. 2018;18(1):51–75.

240. Hernández-Alava M, Wailoo A. Fitting Adjusted Limited Dependent Variable Mixture Models to EQ-5D. Stata J. 2015;15(3):737–50.

241. Gray LA, Hernandez M, Wailoo AJ, Hernandez Alava M, Wailoo AJ. Development of methods for the mapping of utilities using mixture models: An application to asthma. 2017;(17):1–18.

242. Kent S, Gray A, Schlackow I, Jenkinson C, McIntosh E. Mapping from the Parkinson's Disease Questionnaire PDQ-39 to the Generic EuroQol EQ-5D-3L: The Value of Mixture Models. Med Decis Mak. 2015;35(7):902–11.

243. Hernandez-Alava M, Wailoo A, Wolfe F, Michaud K. A Comparison of Direct and Indirect Methods for the Estimation of Health Utilities from Clinical Outcomes. Med Decis Mak. 2014 Oct;34(7):919-30.

244. Hernández-Alava M, Pudney S. Econometric modelling of multiple self-reports of health states: The switch from EQ-5D-3L to EQ-5D-5L in evaluating drug therapies for rheumatoid arthritis. J Health Econ. 2017;55:139–52.

245. Robinson T, Oluboyede Y. Estimating CHU-9D Utility Scores from the WAItE: A Mapping Algorithm for Economic Evaluation. Value Health. 2019;22(2):239–46.

246. Hernandez Alava M, Wailoo A, Grimm S, Pudney S, Gomes M, Sadique Z, et al. EQ-5D-5L versus EQ-5D-3L: The Impact on Cost Effectiveness in the United Kingdom. Value Health. 2018;21(1):49–56.

247. ELSA. English Longitudinal Study on Aging (ELSA) [Internet]. 2020. Available from: https://www.elsa-project.ac.uk/about Date accessed 31 March 2020.

248. Institute for Social Research: University of Michigan. Health and Retirement Study (HRS) [Internet]. 2020. Available from: https://hrs.isr.umich.edu/about Date accessed 31 March 2020.

249. SHARE. SHARE: Survey of Health, Ageing and Retirement in Europe [Internet]. 2020 . Available from: http://www.share-project.org/home0.html Date accessed 31 March 2020.

250. Gray LA, Wailoo AJ, Hernandez Alava M. Mapping the FACT-B Instrument to EQ-5D-3L in Patients with Breast Cancer Using Adjusted Limited Dependent Variable Mixture Models versus Response Mapping. Value Health. 2018;21(12):1399-405.

251. Prosser LA, Wittenberg E. Advances in Methods and Novel Applications for Measuring Family Spillover Effects of Illness. Pharmacoeconomics. 2019;37(4):447-450.

252. Brouwer WBF. The Inclusion of Spillover Effects in Economic Evaluations: Not an Optional Extra. Vol. 37, Pharmacoeconomics. 2019;37(4):451-6.

253. Ara R, Brazier JE. Populating an Economic Model with Health State Utility Values: Moving toward Better Practice. Value Health. 2010;13(5):509–18.

254. Papaioannou D, Brazier J, Paisley S. NICE DSU Technical Support Document 9: The Identification, Review and Synthesis of Health State Utility Values From the Literature. Sheffield: ScHARR, University of Sheffield; 2010. Available from http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD9-HSUV-values_FINAL.pdf Date accessed 31 March 2020.

255. Ara R, Wailoo A. NICE DSU Technical Support Document 12: The Use of Health State Utility Values in Decision Models. Sheffield: ScHARR, University of Sheffield; 2011. Available from http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD12-Utilities-in-modelling-FINAL.pdf date accessed 31 March 2020.

256. Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: results from a general population study. Heal Econ. 1996;5:141–54.

257. Eurostat. Self-perceived health statistics - Statistics Explained [Internet]. 2020. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php/Self-perceived_health_statistics Date accessed 13 February 2020.

258. De Smedt D, Clays E, Doyle F, Kotseva K, Prugger C, Pająk A, et al. Validity and reliability of three commonly used quality of life measures in a large European population of coronary heart disease patients. Int J Cardiol. 2013;167(5):2294-9.

259. De Smedt D, Kotseva K, De Backer G, Wood D, Van Wilder L, De Bacquer D. EQ-5D in coronary patients: what are they suffering from? Results from the ESC EORP European Survey of Cardiovascular Disease Prevention and Diabetes (EUROASPIRE IV) Registry. Qual Life Res. 2020;29(4):1037-1046.

260. Ara R, Wailoo AJ. Estimating Health State Utility Values for Joint Health Conditions: a conceptual review and critique of the current evidence. Med Decis Mak. 2013;33(2):139–53.

261. Ara R, Brazier JE. Using Health State Utility Values from the General Population to Approximate Baselines in Decision Analytic Models when Condition-Specific Data are Not Available. Value Health. 2011;14(4):539–45.

262. Fryback DG, Lawrence WF. Dollars may not buy as many QALYs as we think: A problem with defining quality-of-life adjustments. Med Decis Mak. 1997;17(3):276–84.

263. Thompson A. Incorporating multiple conditions within decision-analytic modelling for use in economic evaluation [dissertation]. Manchester: The University of Manchester; 2019.

264. Thompson AJ, Sutton M, Payne K. Estimating Joint Health Condition Utility Values. Value Health. 2019;22(4):482-90.

265. Ara R, Brazier J. Estimating health state utility values for comorbidities. Pharmacoeconomics. 2017;35(Suppl_1):89–94.

266. Basu A, Dale W, Elstein A, Meltzer D. A linear index for predicting joint health-states utilities from single health-states utilities. Heal Econ. 2009;18:403–19.

267. Dale W, Basu A, Elstein A, Meltzer D. Predicting Utility Ratings for Joint Health States from Single Health States in Prostate Cancer: Empirical Testing of 3 Alternative Theories. Med Decis Mak. 2007;28(1):102–12.

268. Wee HL, Cheung YB, Li SC, Fong KY, Thumboo J. The impact of diabetes mellitus and other chronic medical conditions on health-related quality of life: Is the whole greater than the sum of its parts? Health Qual Life Outcomes. 2005;3(1):2.

269. Ara R, Hill H, Lloyd A, Woods HB, Brazier J. Are Current Reporting Standards Used to Describe Health State Utilities in Cost-Effectiveness Models Satisfactory? Value Health. 2020;23(3):397-405.

270. Bo Hu, Fu AZ. Predicting utility for joint health states: a general framework and a new nonparametric estimator. Med Decis Mak. 2010;30(5):E29-39.

271. Lewis EF, Li Y, Pfeffer MA, Solomon SD, Weinfurt KP, Velazquez EJ, et al. Impact of cardiovascular events on change in quality of life and utilities in patients after myocardial infarction. A VALIANT Study (Valsartan in acute myocardial infarction). JACC Hear Fail. 2014;2(2):159–65.

272. Stafford M, Soljak M, Pledge V, Mindell J. Socio-economic differences in the health-related quality of life impact of cardiovascular conditions. Eur J Public Health. 2012;22(3):301–5.

273. Munyombwe T, Hall M, Dondo TB, Alabas OA, Gerard O, West RM, et al. Quality of life trajectories in survivors of acute myocardial infarction: A national longitudinal study. Heart. 2020;106(1):33–9.

274. Smith DW, Davies EW, Wissinger E, Huelin R, Matza LS, Chung K. A systematic literature review of cardiovascular event utilities. Expert Rev Pharmacoecon Outcomes Res. 2013;13(6):767–90.

275. National Institute for Health and Care Excellence (NICE). NICE CG181. Lipid modification: cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease. London: NICE; 2014. Available from https://www.nice.org.uk/guidance/cg181 Date accessed 31 March 2020.

276. StataCorp. Stata 15 Survey Data manual. College Station, TX: Stata Press;2017. Available from https://www.stata.com/bookstore/survey-data-reference-manual/ Date accessed: 14 May 2019.

277. Torabi A, Cleland JGF, Rigby AS, Sherwi N. Development and course of heart failure after a myocardial infarction in younger and older people. J Geriatr Cardiol. 2014;11(1):1–12.

278. Luengo-Fernandez R, Gray AM, Bull L, Welch S, Cuthbertson F, Rothwell PM, et al. Quality of life after TIA and stroke: ten-year results of the Oxford Vascular Study. Neurology. 2013;81(18):1588–95.

279. De Smedt D, Clays E, Annemans L, Pardaens S, Kotseva K, De Bacquer D. Self-reported health status in coronary heart disease patients: A comparison with the general population. Eur J Cardiovasc Nurs. 2015;14(2):117–25.

280. Wu M, Brazier JE, Kearns B, Relton C, Smith C, Cooper CL. Examining the impact of 11 long-standing health conditions on health-related quality of life using the EQ-5D in a general population sample. Eur J Health Econ. 2015;16(2):141–51.

281. Hobbins A, Barry L, Kelleher D, O'Neill C. The health of the residents of Ireland: Population norms for Ireland based on the EQ-5D-5L descriptive system – a cross sectional study. HRB Open Res. 2018;1:22. doi: 10.12688/hrbopenres.12848.1

282. Matza LS, Stewart KD, Gandra SR, Delio PR, Fenster BE, Davies EW, et al. Acute and chronic impact of cardiovascular events on health state utilities. BMC Health Serv Res. 2015;15(173).

283. Ara R, Peasgood T, Mukuria C, Chevrou-Severac H, Rowen D, Azzabi-Zouraq I, et al. Sourcing and Using Appropriate Health State Utility Values in Economic Models in Health Care. Pharmacoeconomics. 2017 Dec 19;35(Suppl_1):7–9.

284. Wallace E, Moriarty F, McGarrigle C, Smith SM, Kenny R-A, Fahey T. Self-report versus electronic medical record recorded healthcare utilisation in older community-dwelling adults: Comparison of two prospective cohort studies. Arias SA, editor. PLoS One. 2018 13(10):e0206201.

285. Korman MJ, Retterstol K, Kristiansen IS, Wisloff T. Are PCSK9 Inhibitors Cost Effective? Pharmacoeconomics. 2018;36(9):1031–41.

286. Arrieta TF, Page TF, Veledar E, Nasir K. Economic evaluation of PCSK9 inhibitors in reducing cardiovascular risk from health system and private payer perspectives. PLoS One. 2017;12(1):1–14.

287. Fonarow GC, Keech AC, Pedersen TR, Giugliano RP, Sever PS, Lindgren P, et al. Cost-effectiveness of Evolocumab Therapy for Reducing Cardiovascular Events in Patients With Atherosclerotic Cardiovascular Disease. JAMA Cardiol. 2017;2(10):1069–78.

288. Gandra G, Fonarow G, Lothgren M, Lindgren P, Somaratne R, Van Hout B. Cost-effectiveness of LDL-C lowering with evolocumab in patients with high cardiovascular risk in the United States. Value Health. 2016;19(3):A3.

289. Korman M, Wisløff T. Modelling the cost-effectiveness of PCSK9 inhibitors vs. ezetimibe through LDL-C reductions in a Norwegian setting. Eur Hear J Cardiovasc Pharmacother. 2018;4(1):15-22.

290. Villa G, Lothgren M, Kutikova L, Lindgren P, Gandra SR, Fonarow GC, et al. Cost-effectiveness of Evolocumab in Patients With High Cardiovascular Risk in Spain. Value Health. 2017;39(4):771-786 e3.

291. Kazi DS, Penko J, Coxson PG, Moran AE, Ollendorf DA, Tice JA, et al. Updated cost-effectiveness analysis of PCSK9 inhibitors based on the results of the FOURIER trial. J Am Med Assoc. 2017;318(8):748–50.

292. Kazi DS, Moran AE, Coxson PG, Penko J, Ollendorf DA, Pearson SD, et al. Cost-Effectiveness of PCSK9 inhibitor therapy in patients with heterozygous familial hypercholesterolemia or atherosclerotic cardiovascular disease. J Am Med Assoc. 2016;316(7):743–53.

293. Tice JA, Kazi DS, Pearson SD. Proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitors for treatment of high cholesterol levels effectiveness and value. JAMA Intern Med. 2016;176(1):107–8.

294. Institute for Clinical and Economic Review (ICER). PCSK9 Inhibitors for Treatment of High Cholesterol: Effectiveness, Value, and Value Based Price Benchmarks. Boston: ICER; 2015. Available from: https://icer-review.org/wp-content/uploads/2016/01/Final-Report-for-Posting-11-24-15-1.pdf Date accessed 31 March 2020.

295. Herrett E, Shah AD, Boggon R, Denaxas S, Smeeth L, Van Staa T, et al. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: Cohort study. BMJ. 2013;346:f2350

296. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. Value Health. 2012;15(6):812–20.

297. Williams C. Exploring multi-state modelling in an epidemiological and health economics context [dissertation]. Glasgow: University of Glasgow; 2018.

298. Rapsomaniki E, Shah A, Perel P, Denaxas S, George J, Nicholas O, et al. Prognostic models for stable coronary artery disease based on electronic health record cohort of 102 023 patients. Eur Heart J. 2014;35(13):844–52.

299. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med. 2007;26(11):2389–430.

300. Walker S, Asaria M, Manca A, Palmer S, Gale CP, Shah AD, et al. Long-term healthcare use and costs in patients with stable coronary artery disease: a population-based cohort using linked health records (CALIBER). Eur Hear J Qual Care Clin Outcomes. 2016;2(2):125–40.

301. Heldner MR, Li L, Lovett NG, Kubiak MM, Lyons S, Rothwell PM, et al. Long-Term Prognosis of Patients With Transient Ischemic Attack or Stroke and Symptomatic Vascular Disease in Multiple Arterial Beds. Stroke. 2018;49(7):1639–46.

302. Central Statistics Office (CSO). Irish Life Tables No. 16 2010-2012 [Internet]. 2020. Available from: https://www.cso.ie/en/releasesandpublications/er/ilt/irishlifetablesno162010-2012/ Date accessed 31 March 2020.

303. Central Statistics Office (CSO). VSA08: Deaths Occuring by Sex, Cause of Death, Age at Death and Year [Internet]. 2020. Available from: https://statbank.cso.ie/px/pxeirestat/Statire/SelectVarVal/Define.asp?maintable=VSA08&PLanguage=0# Date accessed 31 March 2020 .

304. Healthcare Pricing Office (HPO). ABF 2019 Admitted Price List. Dublin: HSE; 2019. Available from: http://www.hpo.ie/abf/ABF2019AdmittedPatientPriceList.pdf Date accessed 31 March 2020.

305. Health Service Executive (HSE) . HSE Ready Reckoner. Dublin: HSE; 2013. Available from: www.hse.ie/eng/services/list/1/schemes/cbd/Ready%20Reckoner.pdf Date accessed 14 September 2017.

306. National Centre for Pharmacoeconomics (NCPE). Guidelines for inclusion of drug costs in pharmacoeconomic evaluations Version 2.1. Dublin: NCPE; 2020. Available from: http://www.ncpe.ie/submission-process/hta-guidelines/guidelines-for-inclusion-of-drug-costs/ Date accessed 31 March 2020.

307. CTT Collaboration [Internet]. 2020. Available from: https://www.cttcollaboration.org/ Date accessed 31 March 2020.

308. Baigent C, Keech A, Kearney PM, Blackwell L, Buck G, Pollicino C, et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. Lancet. 2005;366(9493):1267–78.

309. Armitage J, Baigent C, Barnes E, John Betteridge D, Blackwell L, Blazing M, et al. Efficacy and safety of statin therapy in older people: a meta-analysis of individual participant data from 28 randomised controlled trials. Lancet. 2019;393(10170):407-415

310. Barton GR, Briggs AH, Fenwick EAL. 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.

311. Asaria M, Grasic K, Walker S. Using Linked Electronic Health Records to Estimate Healthcare Costs: Key Challenges and Opportunities. Pharmacoeconomics. 2016;34(2):155–60.

312. Wisløff T, Mundal LJ, Retterstøl K, Igland J, Kristiansen IS. Economic evaluation of lipid lowering with PCSK9 inhibitors in patients with familial hypercholesterolemia: Methodological aspects. Atherosclerosis. 2019;287:140–6.

313. Koerkamp BG, Stijnen T, Weinstein MC, Hunink MGM. The combined analysis of uncertainty and patient heterogeneity in medical decision models. Med Decis Mak. 2011;31(4):650–61.

314. Jalal H, Pechlivanoglou P, Krijkamp E, Alarid-Escudero F, Enns E, Myriam Hunink MG. An Overview of R in Health Decision Sciences. Med Decis Mak. 2017;37(7):735–46.

315. DARTH – Decision Analysis in R for Technologies in Health [Internet]. Available from: https://darthworkgroup.com/ Date accessed 28 March 2020.

316. R for HTA [Internet]. 2020. Available from: https://r-hta.org/ Date accessed 31 March 2020.

317. Hollman C, Paulden M, Pechlivanoglou P, McCabe C. A Comparison of Four Software Programs for Implementing Decision Analytic Cost-Effectiveness Models. Pharmacoeconomics. 2017;35(8):817-30.

318. Incerti D, Thom H, Baio G, Jansen JP. R You Still Using Excel? The Advantages of Modern Software Tools for Health Technology Assessment. Value Health. 2019;22(5):575–9.

319. Williams C, Lewsey JD, Briggs AH, Mackay DF. Cost-effectiveness Analysis in R Using a Multi-state Modeling Survival Analysis Framework: A Tutorial. Med Decis Mak. 2017;37(4):340–52.

320. 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.

APPENDICES

APPENDICES TABLE OF CONTENTS

1.1 Chapter 3 Appendices 3

1.1.1 Search Strategy 3

1.1.2 Excluded Studies 9

1.1.3 Risk of Bias Summary Graphs PCSK9 inhibitors versus Placebo for Study Primary Outcome Measures 11

1.2 Chapter 5 Appendices 14

1.2.1 Survey Questionnaire administered to the Independent Sample 14

1.2.2 Descriptive Statistics 28

1.2.3 Correlation Coefficients between TILDA Question Subset and EQ-5D-3L 29

1.2.4 Mapping model results 31

1.3 Chapter 6 Appendices 35

1.4 Chapter 7 Appendices 41

1.4.1 Costs 42

1.4.2 Population Characteristics 45

1.4.3 Results 49

1.4.4 Risk Equations and Parametric Survival Distributions as Estimated by Asaria et al 51

LIST OF APPENDIX TABLES

Table A.1 Excluded studies 9
Table A.2 Time to non-fatal myocardial infarction 11
Table A.3 Time to fatal or non-fatal ischemic stroke 11
Table A.4 Time to fatal or non-fatal haemorrhagic stroke 12
Table A.5 Time to cardiovascular death 12
Table A.6 Independent sample responses to EQ-5D-3L domains 28
Table A.7 Summary of independent sample responses to TILDA questions included in Mapping Model 28
Table A.8 Spearman correlation coefficient for all question subset variables considered for mapping model from Independent Sample 29
Table A.9 Model coefficient and standard errors for the six mapping models 31
Table A.10 Mapping goodness of fit results by observed utility subgroups 33
Table A.11 Variance-co-variance matrix for the chosen ALDV mapping model 34
Table A.12 Exemplar health state utilities reported in studies identified by Ara et al (1) 36
Table A.13 Population characteristics including initial predicted utility values 40
Table A.14 Table of cardiovascular cost parameters applied in the model per 90-day cycle 42
Table A.15 MI and stroke costs from the HPO activity based funding price list* 44
Table A.16a Modelled patient characteristics for the TILDA populations 45
Table A.16b Modelled patient characteristics for the TILDA populations 47
Table A.17 Total and incremental costs and outcomes for PCSK9 inhibitors for 23 subgroups from TILDA with a modelled of 4mmol/L under the indirect treatment effect approach 49
Table A.18 Total and incremental costs and outcomes for PCSK9 inhibitors versus standard of care for TILDA population 1 by age and gender assuming a direct treatment effect 50

LIST OF APPENDIX FIGURES

Figure A1 Efficacy of PCSK9 inhibitors versus control for serious adverse events. 13

1.1 Chapter 3 Appendices

1.1.1 Search Strategy

1.1.1.1 Central 3rd December 2019

#1	MeSH descriptor: [Antibodies, Monoclonal] explode all trees	10998
#2	monoclonal next antibod*	10118
#3	MAB*	3234
#4	evolocumab	269
#5	"amg 145" or amg145	88
#6	alirocumab	289
#7	"regn 727" or regn727 or "sar 236553" or sar236553 or "1D05-IgG2" or LGT209 or RG7652	59
#8	Bocoizumab	0
#9	"pf 04950615" or pf04950615 or "rn 316" or rn316	36
#10	#4 or #5 or #6 or #7 or #8 or #9	615
#11	MeSH descriptor: [Proprotein Convertases] explode all trees	1860
#12	proprotein next convertase*	422
#13	pro-protein next convertase*	4
#14	pcsk9	591
#15	serine next proteinase	274
#16	#11 or #12 or #13 or #14 or #15	2629
#17	MeSH descriptor: [Cardiovascular Diseases] explode all trees	99365
#18	cardio*	141600
#19	cardia*	66190
#20	heart*	151898
#21	coronary*	55721
#22	angina*	13863
#23	ventric*	31274
#24	myocard*	43600
#25	pericard*	1791
#26	isch?em*	43764
#27	emboli*	11191
#28	arrhythmi*	12636

#29	thrombo*	53981
#30	atrial next fibrillat*	12225
#31	tachycardi*	8906
#32	endocardi*	1209
#33	(sick next sinus)	337
#34	MeSH descriptor: [Stroke] explode all trees	8696
#35	(stroke or strokes)	66535
#36	cerebrovasc*	23520
#37	cerebral next vascular	659
#38	apoplexy	451
#39	(brain near/2 accident*)	259
#40	((brain* or cerebral or lacunar) near/2 infarct*)	5020
#41	MeSH descriptor: [Hyperlipidemias] explode all trees	5921
#42	hyperlipid*	6237
#43	hyperlip?emia	356
#44	hypercholesterol*	7900
#45	hypercholester?emia	86
#46	hyperlipoprotein?emia	971
#47	hypertriglycerid?emia	2115
#48	MeSH descriptor: [Arteriosclerosis] explode all trees	9307
#49	MeSH descriptor: [Cholesterol] explode all trees	9988
#50	cholesterol	35527
#51	"coronary risk factor"	133
#52	MeSH descriptor: [Cognition] explode all trees	9934
#53	MeSH descriptor: [Dementia] explode all trees	5426
#54	cognitive next function	8876
#55	dementia	22295
#56	alzheimer*	12113
#57	#17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41 or #42 or #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53 or #54 or #55 or #56	405480
#58	#10 AND #16 and #57 with Publication Year from 2005 to 2019, in Trials	400

1.1.1.2 Embase

#1	**'evolocumab'**/exp	1,667
#2	**'alirocumab'**/exp	1,506
#3	**evolocumab**:ti,ab OR **alirocumab**:ti,ab OR **repatha**:ti,ab OR **praluent**:ti,ab OR **'amg 145'**:ti,ab OR **amg145**:ti,ab	1,246
#4	((**pcsk9** OR **'proprotein convertase 9'**) NEAR/2 (**inhibit*** OR **supress***)):ti,ab	1,895
#5	#1 OR #2 OR #3 OR #4	3,221
#6	[controlled clinical trial]/lim OR [randomized controlled trial]/lim OR **'clinical article'**/de OR **'clinical trial'**/de OR **'clinical trial (topic)'**/de OR **'controlled clinical trial'**/de OR **'controlled study'**/de OR **'major clinical study'**/de OR **'multicenter study'**/de OR **'phase 3 clinical trial (topic)'**/de OR **'randomized controlled trial'**/de OR **'randomized controlled trial (topic)'**/de	10,960,217
#7	#5 AND #6 AND [2005-2020]/py	1,737

1.1.1.3 Web of Science -

1. TS=((evolocumab or alirocumab OR Repatha OR praluent OR 'amg 145' OR amg145)
2. TS=(random* or blind* or allocat* or assign* or trial* or placebo* or crossover* or cross-over*)
3. #2 AND #1

1.1.1.4 Medline Search Strategy -

1. exp antibodies, monoclonal/
2. monoclonal antibod*.tw.
3. MAB*.tw.
4. evolocumab.tw.
5. amg 145.tw.
6. amg145.tw.
7. alirocumab.tw.
8. regn 727.tw.
9. regn727.tw.
10. sar 236553.tw.
11. sar236553.tw.
12. 1D05-IgG2.tw.
13. LGT209.tw.

14. RG7652.tw.
15. Bococizumab.tw.
16. "pf 04950615".tw.
17. pf04950615.tw.
18. rn 316.tw.
19. rn316.tw.
20. or/1-19
21. exp Proprotein Convertases/
22. proprotein convertase*.tw.
23. pro-protein convertase*.tw.
24. pcsk9.tw.
25. serine proteinase*.tw.
26. or/21-25
27. exp Cardiovascular Diseases/
28. cardio*.tw.
29. cardia*.tw.
30. heart*.tw.
31. coronary*.tw.
32. angina*.tw.
33. ventric*.tw.
34. myocard*.tw.
35. pericard*.tw.
36. isch?em*.tw.
37. emboli*.tw.
38. arrhythmi*.tw.
39. thrombo*.tw.
40. atrial fibrillat*.tw.
41. tachycardi*.tw.
42. endocardi*.tw.
43. (sick adj sinus).tw.
44. exp Stroke/
45. (stroke or stokes).tw.

46. cerebrovasc*.tw.

47. cerebral vascular.tw.

48. apoplexy.tw.

49. (brain adj2 accident*).tw.

50. ((brain* or cerebral or lacunar) adj2 infarct*).tw.

51. exp Hyperlipidemias/

52. hyperlipid*.tw.

53. hyperlip?emia*.tw.

54. hypercholesterol*.tw.

55. hypercholester?emia*.tw.

56. hyperlipoprotein?emia*.tw.

57. hypertriglycerid?emia*.tw.

58. exp Arteriosclerosis/

59. exp Cholesterol/

60. cholesterol.tw.

61. "coronary risk factor* ".tw.

62. exp Cognition/

63. exp dementia/

64. cognitive function*.tw.

65. dementia.tw.

66. alzheimer*.tw.

67. or/27-66

68. 20 and 26 and 67

69. randomized controlled trial.pt.

70. controlled clinical trial.pt.

71. randomized.ab.

72. placebo.ab.

73. drug therapy.fs.

74. randomly.ab.

75. trial.ab.

76. groups.ab.

77. 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76

78. exp animals/ not humans.sh.

79. 77 not 78

80. 68 and 79

81. limit 80 to yr="2005 -2019"

1.1.2 Excluded Studies

Table A.1 Excluded studies

NCT	Name of Study	Acronym	Intervention	Reason for Exclusion
NCT01507831	Long-term Safety and Tolerability of Alirocumab (SAR236553/REGN727) Versus Placebo on Top of Lipid-Modifying Therapy in High Cardiovascular Risk Patients with Hypercholesterolemia	(ODYSSEY LONG TERM)	Alirocumab	Insufficient follow-up
NCT01516879	Durable Effect of PCSK9 Antibody Compared with placebo Study	DESCARTES	Evolocumab	Insufficient follow-up
NCT01588496	Trial Evaluating PCSK9 Antibody in Subjects with LDL Receptor Abnormalities	TESLA	Evolocumab	Insufficient population size
NCT01623115	Efficacy and Safety of Alirocumab (SAR236553/REGN727) Versus Placebo on Top of Lipid-Modifying Therapy in Patients with Heterozygous Familial Hypercholesterolemia Not Adequately Controlled with Their Lipid-Modifying Therapy	ODYSSEY FH I	Alirocumab	Insufficient population size
NCT01624142	Trial Assessing Long Term Use of PCSK9 Inhibition in Subjects with Genetic LDL Disorders	TAUSSIG	Evolocumab	Insufficient population size
NCT01644175	Efficacy and Safety of Alirocumab (SAR236553/REGN727) Versus Placebo on Top of Lipid-Modifying Therapy in Patients with High Cardiovascular Risk and Hypercholesterolemia	ODYSSEY COMBO I	Alirocumab	Insufficient population size
NCT01709500	Study of Alirocumab (REGN727/SAR236553) in Patients With HeFH Who Are Not Adequately Controlled with Their Lipid-Modifying Therapy	ODYSSEY FH II	Alirocumab	Insufficient population size
NCT01709513	Study of Alirocumab (REGN727/SAR236553) in Patients with Primary Hypercholesterolemia and Moderate, High, or Very High Cardiovascular Risk, Who Are Intolerant to Statins (ODYSSEY ALTERNATIVE)	ODYSSEY COMBO I	Alirocumab	Insufficient population size
NCT01730040	Study of the Efficacy and Safety of Alirocumab (REGN727/SAR236553) in Combination with Other Lipid-modifying Treatment	ODYSSEY OPTIONS I	Alirocumab	Insufficient population size
NCT01730053	Study of Alirocumab (REGN727/SAR236553) added-on to Rosuvastatin Versus Other Lipid Modifying Treatments	ODYSSEY OPTIONS II	Alirocumab	Insufficient population size
NCT01763827	Monoclonal Antibody Against PCSK9 to Reduce Elevated LDL-C in Subjects Currently Not Receiving Drug Therapy for Easing Lipid Levels-2	MENDEL-2	Evolocumab	Insufficient follow-up
NCT01763866	LDL-C Assessment with PCSK9 Monoclonal Antibody Inhibition Combined with Statin Therapy-2	LAPLACE-2	Evolocumab	Insufficient follow-up
NCT01763905	Goal Achievement After Utilizing an Anti-PCSK9 Antibody in Statin Intolerant Subjects -2	GAUSS-2	Evolocumab	Insufficient population size.
NCT01763918	Reduction of LDL-C With PCSK9 Inhibition in Heterozygous Familial Hypercholesterolemia Disorder Study-2	RUTHERFORD-2	Evolocumab	Insufficient population size.
NCT01813422	Global Assessment of Plaque regression With a PCSK9 antibody as Measured by intravascular Ultrasound	GLAGOV	Evolocumab	Insufficient follow-up
NCT01854918	Open-label Extension Study of Evolocumab (AMG 145) in Adults with Hyperlipidemia and Mixed Dyslipidemia	OSLER-2	Evolocumab	Open label
NCT01926782	Study to Evaluate the Efficacy and Safety of an Every Four Weeks Treatment Regimen of Alirocumab (REGN727/ SAR236553) in Patients With Primary Hypercholesterolemia	ODYSSEY CHOICE 1	Alirocumab	Insufficient follow-up
NCT01953328	Study of Low-Density Lipoprotein Cholesterol (LDL-C) Reduction Using Evolocumab (AMG 145) in Japanese Patients with Advanced Cardiovascular Risk	AMG145	Evolocumab	Insufficient population size.
NCT01954394	Open Label Study of Long Term Safety Evaluation of Alirocumab	ODYSSEY OLE	Alirocumab	Open label
NCT01984424	Goal Achievement After Utilizing an Anti-PCSK9 Antibody in Statin Intolerant Subjects-3	GAUSS-3	Evolocumab	Insufficient follow-up
NCT02023879	Phase III Study to Evaluate Alirocumab in Patients with Hypercholesterolemia Not Treated with a Statin	ODYSSEY CHOICE II	Alirocumab	Insufficient population size
NCT02107898	Efficacy and Safety Evaluation of Alirocumab in Patients with Heterozygous Familial Hypercholesterolemia or	ODYSSEY JAPAN	Alirocumab	Insufficient population size

NCT	Name of Study	Acronym	Intervention	Reason for Exclusion
	High Cardiovascular Risk Patients with Hypercholesterolemia on Lipid Modifying Therapy			
NCT02189837	Effects on Lipoprotein Metabolism from PCSK9 Inhibition Utilizing a Monoclonal Antibody	FLOREY	Evolocumab	Insufficient population size
NCT02289963	Evaluation of Alirocumab in Addition to Lipid-Modifying Therapy in Patients with High Cardiovascular Risk and Hypercholesterolemia in South Korea and Taiwan		Alirocumab	Insufficient population size
NCT02304484	Open-label Extension Study to Assess Safety and Efficacy of Evolocumab	OSLER-1	Evolocumab	Open label
NCT02392559	Trial Assessing Efficacy, Safety and Tolerability of PCSK9 Inhibition in Paediatric Subjects with Genetic LDL Disorders	HAUSER-RCT	Evolocumab	Wrong population
NCT02476006	Safety, Tolerability, and Effect of Alirocumab in High Cardiovascular Risk Patients with Severe Hypercholesterolemia Not Adequately Controlled with Conventional Lipid-modifying Therapies	ODYSSEY APPRISE	Alirocumab	Open label
NCT02585778	Efficacy and Safety of Alirocumab Versus Placebo on Top of Maximally Tolerated Lipid Lowering Therapy in Patients with Hypercholesterolemia Who Have Type 1 or Type 2 Diabetes and Are Treated With Insulin	ODYSSEY DM-Insulin	Alirocumab	Insufficient follow-up
NCT02634580	Goal Achievement After Utilizing an Anti-PCSK9 Antibody in Statin Intolerant Subjects-4	GAUSS-4	Evolocumab	Insufficient population size.
NCT02642159	Efficacy and Safety of Alirocumab Versus Usual Care on Top of Maximally Tolerated Statin Therapy in Patients with Type 2 Diabetes and Mixed Dyslipidemia	ODYSSEY DM-Dyslipidemia	Alirocumab	Insufficient population size
NCT02662569	Safety and Efficacy of Evolocumab in Combination with Statin Therapy in Adults with Diabetes and Hyperlipidemia or Mixed Dyslipidemia	BERSON	Evolocumab	Insufficient follow-up
NCT02715726	Evaluation of Alirocumab Versus Ezetimibe on Top of Statin in Asia in High Cardiovascular Risk Patients with Hypercholesterolemia	ODYSSEY EAST	Alirocumab	Insufficient follow-up
NCT02729025	Effects of Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9) Inhibition on Arterial Wall Inflammation in Patients with Elevated Lipoprotein(a) (Lp(a))	ANITSCHKOW	Evolocumab	Insufficient Population size.
NCT02739984	Evaluation of Evolocumab (AMG 145) Efficacy in Diabetic Adults with Hypercholesterolemia/Mixed Dyslipidemia	BANTING	Evolocumab	Insufficient population size.
NCT02867813	Further Cardiovascular Outcomes Research with PCSK9 Inhibition in Subjects with Elevated Risk Open-label Extension	FOURIER OLE	Evolocumab	Open label
NCT02984982	Evaluation of Effect of Alirocumab on Coronary Atheroma Volume in Japanese Patients Hospitalized for Acute Coronary Syndrome with Hypercholesterolemia	ODYSSEY J-IVUS	Alirocumab	Insufficient population size
NCT03080935	Fourier Open-label Extension Study in Subjects with Clinically Evident Cardiovascular Disease in Selected European Countries	FOURIER-OLE	Evolocumab	Open label
NCT03287609	Evolocumab for Early Reduction of LDL-cholesterol Levels in Patients with Acute Coronary Syndromes (EVOPACS)	EVOPACS	Evolocumab	Insufficient population size.

HeFH Heterozygous Familial Hypercholesterolemia LDL, Low Density Lipoprotein; LDL-C Low Density Lipoprotein Cholesterol; PCSK9, Proprotein Convertase Subtilisin/Kexin Type 9.

1.1.3 Risk of Bias Summary Graphs PCSK9 inhibitors versus Placebo for Study Primary Outcome Measures

Individuals trials are weighted by their contribution to the average effect estimate for each outcome

Table A.2 Time to non-fatal myocardial infarction

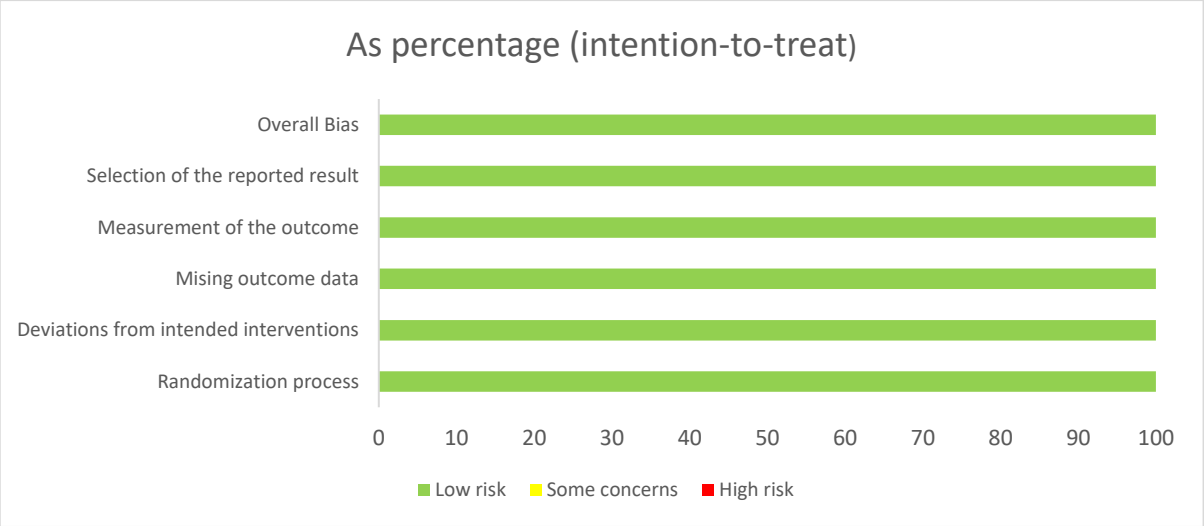


Table A.3 Time to fatal or non-fatal ischemic stroke

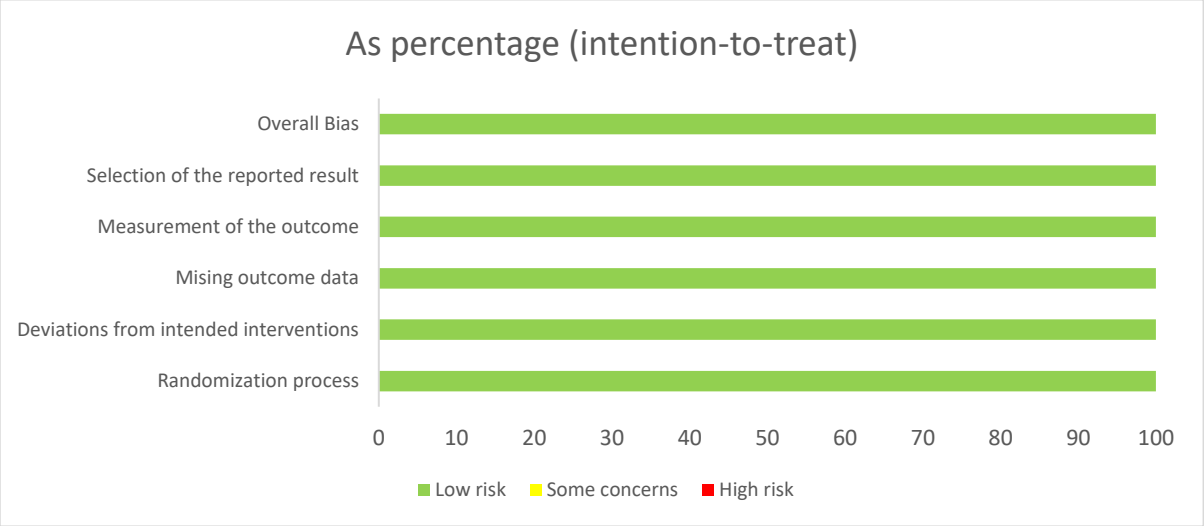


Table A.4 Time to fatal or non-fatal haemorrhagic stroke

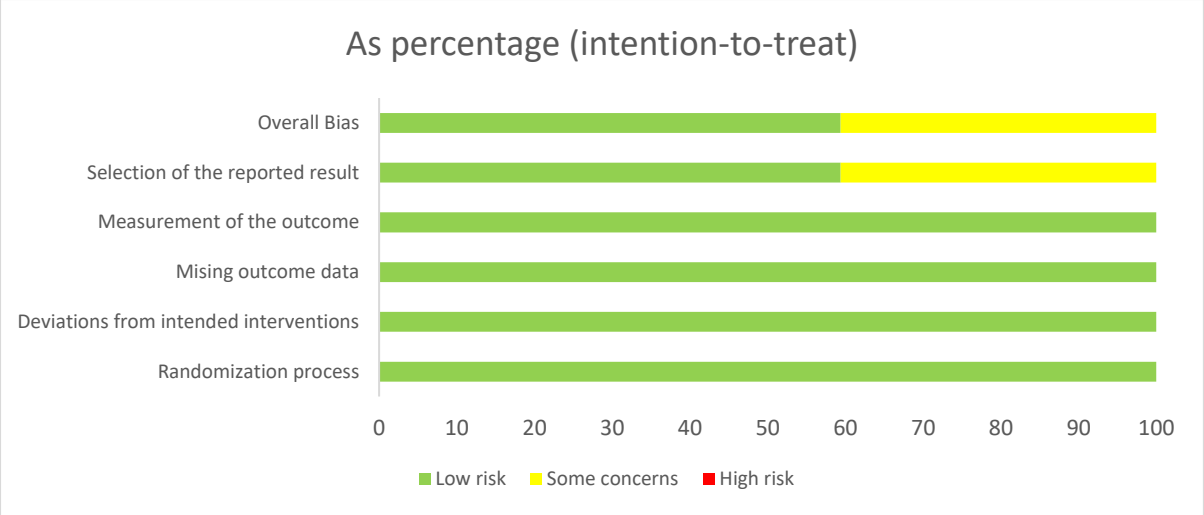
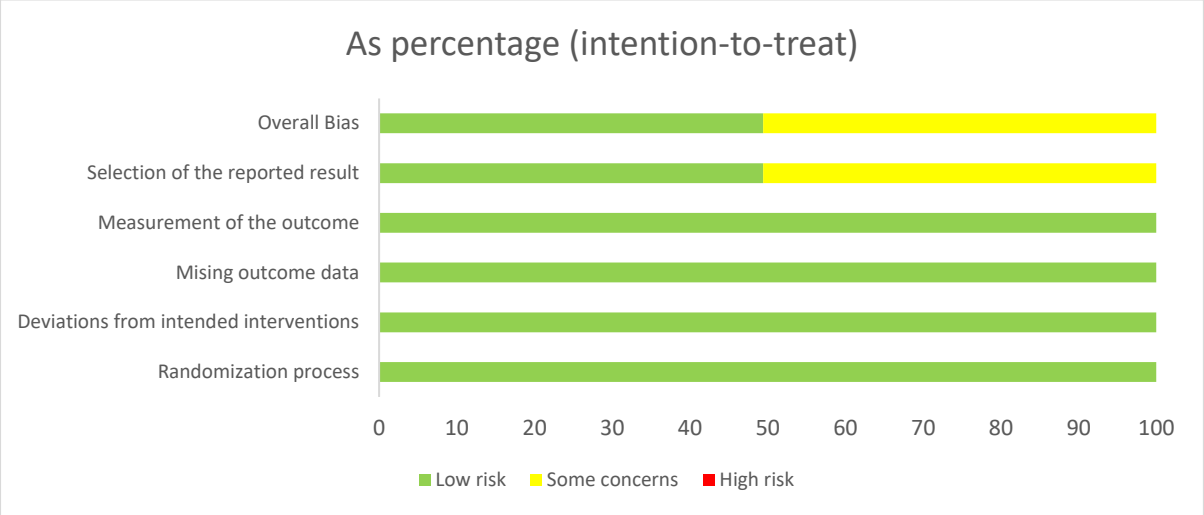


Table A.5 Time to cardiovascular death



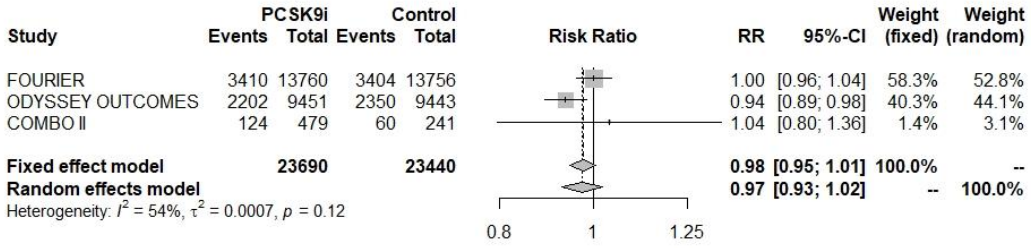


Figure A1 Efficacy of PCSK9 inhibitors versus control for serious adverse events.

1.2 Chapter 5 Appendices

1.2.1 Survey Questionnaire administered to the Independent Sample

D1: ID NUMBER:

Please circle the patients response or fill in blank where applicable.

D2: Date:_____

D3: Interviewer:_____

D4: Out Patient Clinic:_____

D5. Is this respondent male or female? Code without asking

1. Male

2. Female

Interview Starts Here

Abbreviations: DK = Don't Know; RF = Patient declines to answer

D6: What is your date of birth: _____

D7: Now I would like to ask some questions about your background. What is the highest level of education you have completed?

1. Some primary (not complete)

2. Primary or equivalent

3. Intermediate/junior/group certificate or equivalent

4. Leaving certificate or equivalent

5. Diploma/certificate

6. Primary degree

7. Postgraduate/higher degree

96. None

98. DK

99. RF

Section A Health Questions from TILDA

IWER: CODE THE ONE THAT APPLIES

PH001: Now I would like to ask you some questions about your health.

Would you say your health is..

IWER: READ OUT

1. excellent,

2. very good,

3. good,

4. fair,

5. or, poor?

98. DK

99. RF

IWER: CODE THE ONE THAT APPLIES

PH002: What about your emotional or mental health? Is it …

IWER: READ OUT

1. excellent,

2. very good,

3. good,

4. fair,

5. or, poor?

98. DK

99. RF

PH003: Some people suffer from chronic or long-term health problems. By long-term we mean it has troubled you over a period of time or is likely to affect you over a period of time. Do you have any long-term health problems, illness, disability or infirmity?

NOTE: INCLUDING MENTAL HEALTH PROBLEMS

1. Yes

5. No

98. DK

99. RF

PH004: Does this illness or disability limit your activities in any way?

1. Yes

5. No **Go to PH009**

98. DK **Go to PH009**

99. RF **Go to PH009**

PH005: For the past six months or more, to what extent have you been limited because of a health problem in activities people usually do?

IWER: READ OUT

1. Severely limited

2. Limited, but not severely

3. Not limited

98. DK

99. RF

PH009: In general, compared to other people your age, would you say your health is ……

IWER: READ OUT

1. excellent,

2. very good,

3. good,

4. fair,

5. or, poor?

98. DK

99. RF

Heart Disease Section

INTRO: READ OUT **We are interested in finding out more information about heart problems people may suffer from.**

IWER: SHOW CARD PH1

PH201: Please look at card PH1. Has a doctor ever told you that you have any of the conditions on this card?

INTERVIEWER: PROBE - 'WHAT OTHERS?' CODE ALL THAT APPLY.

1.	High blood pressure or hypertension		**[ph201_1]**
2.	Angina	**GO TO PH203**	**[ph201_2]**
3.	A heart attack (including myocardial infarction or coronary thrombosis)	**GO TO PH205**	**[ph201_3]**
4.	Congestive heart failure	**GO TO PH212**	**[ph201_4]**
5.	Diabetes or high blood sugar		**[ph201_5]**
6.	A stroke (cerebral vascular disease)	**GO TO PH218**	**[ph201_6]**
7.	Ministroke or TIA	**GO TO PH221**	**[ph201_7]**
8.	High cholesterol		**[ph201_8]**
9.	A heart murmur		**[ph201_9]**
10.	An abnormal heart rhythm		**[ph201_10]**
11.	Peripheral Arterial Disease (PAD)		**[HOD1]**
95.	Any other heart trouble (specify) **[ph201a]**	**GO TO PH224**	**[ph201_15]**
96.	None of these		**[ph201_14]**
98.	DK		**[ph201_12]**
99.	RF		**[ph201_13]**

If Angina = Yes

PH203: Approximately how old were you when you were first told by a doctor that you had angina?

Age in Years _____

98. DK

99. RF

PH204: Are you limiting your usual activities because of your angina?

1. Yes

5. No

98. DK

99. RF

****

*****

If Heart Attack = Yes

PH205: Approximately how old were you when you were first told by a doctor that you had a heart attack (including myocardial infarction or coronary thrombosis)?

Age in Years _____

98. DK

99. RF

PH206: In what year/month was your (most recent) heart attack?

(MM/YYYY)

_____/_____

_____ DK RF MONTH

_____ DK RF YEAR

PH207: According to the doctor how many heart attacks have you had?

Number _____

98. DK

99. RF

PH208: Have you ever had an angioplasty or Stent?

1. Yes Go to PH209

5. No Go to PH210

98. DK

99. RF

PH209: In what year/month was your last angioplasty or Stent? [ph209, ph209a]

(MM/YYYY)

_____/_____

_____ DK RF MONTH

_____ DK RF YEAR

PH210: Have you ever had open heart surgery?

1. Yes Go to PH211

5. No

98. DK

99. RF

PH211: In what year/month was your last heart surgery? [ph211, ph211a]

(MM/YYYY)

_____/_____

_____ DK RF MONTH

_____ DK RF YEAR

If heart failure = yes

PH212: Approximately how old were you when you were first told by a doctor that you had congestive heart failure?

Age in Years _____

98. DK

99. RF

If stroke = Yes

PH218: Approximately how old were you when you were first told by a doctor that you had a stroke?

Age in Years _____

98. DK

99. RF

PH219: How many strokes have you had?

Number _____

98. DK

99. RF

PH220: In what year was your most recent stroke? [ph220, ph220a]

(MM/YYYY)

_____/_____

_____ DK RF MONTH

_____ DK RF YEAR

If TIA = Yes

PH221: Approximately how old were you when you were first told by a doctor that you had a TIA, ministroke, or transient ischemic attack?

Age in Years _____

98. DK

99. RF

PH222: How many TIA's or ministrokes have you had?

Number _____

98. DK

99. RF

PH223: In what year was your most recent TIA or ministrokes? [ph223, ph223a]

(MM/YYYY)

_____/_____

_____ DK RF MONTH _____ DK RF YEAR

If PAD = Yes

HOD2: Approximately how old were you when you were first told by a doctor that you had peripheral arterial disease (PAD)

Age in Years _____

98. DK

99. RF

If any other heart trouble = Yes

PH224: Approximately how old were you when you were first told by a doctor that you had any other heart trouble?

1….120

98. DK

99. RF

(TILDA)

Other Comorbidities

IWER: SHOW CARD PH2

PH301: Please look at card PH2. Has a doctor ever told you that you have any of the following conditions?

IWER: PROBE - 'WHAT OTHERS?' CODE ALL THAT APPLY.

1. Chronic lung disease such as chronic bronchitis emphysema or asthma?

2. Arthritis (including osteoarthritis, or rheumatism)

3. Cancer or a malignant tumour

(including leukaemia or lymphoma but excluding minor skin cancers)

4. Any emotional, nervous or psychiatric problems,

such as depression or anxiety

5. Chronic Kidney Disease

98. DK

99. RF

Pain Section

PH501: Are you often troubled with pain?

1. Yes **GO TO PH502**

5. No

98. DK

99. RF

IWER: CODE THE ONE THAT APPLIES

PH502: How bad is the pain most of the time? Is it...

IWER: READ OUT

1 mild,

2 moderate,

3 or, severe

98. DK

99. RF

****

Difficulties with Activities of Daily Life

INTRO: We need to understand the difficulties people may have with various activities

IWER: IF R IS CONFINED TO BED OR A WHEELCHAIR, READ THE FOLLOWING STATEMENT:

'I AM REQUIRED TO ASK ABOUT ALL OF THESE ACTIVITIES. I REALIZE THAT YOU MAY NOT BE ABLE TO DO SOME OF THEM, BUT I WOULD APPRECIATE IT IF YOU WOULD JUST CONFIRM THAT WITH ME AS WE GO THROUGH THE LIST.'

IWER: SHOW CARD FL1

FL001. Please look at card FL1. Because of a physical or mental health problem, do you have difficulty doing any of the activities on this card? Exclude any difficulties that you expect to last less than three months.

IWER: READ EACH ONE OF THE DIFFICULTIES

IWER: PROBE: ANY OTHERS? CODE ALL THAT APPLY

1. Walking 100 meters (100 yards)		**[fl001_01]**
2. Running or jogging about 1.5 kilometers (1 mile)		**[fl001_02]**
3. Sitting for about two hours		**[fl001_03]**
4. Getting up from a chair after sitting for long periods		**[fl001_04]**
5. Climbing several flights of stairs without resting		**[fl001_05]**
6. Climbing one flight of stairs without resting		**[fl001_06]**
7. Stooping, kneeling, or crouching		**[fl001_07]**
8. Reaching or extending your arms above shoulder level		**[fl001_08]**
9. Pulling or pushing large objects like a living room chair		**[fl001_09]**
10. Lifting or carrying weights over 10 pounds/5 kilos, like a heavy bag of groceries		**[fl001_10]**
11. Picking up a small coin from a table		**[fl001_11]**
96. None of these	**Questionaire Finished**	**[fl001_13]**
99 RF	**Questionarire Finished**	**[fl001_12]**

IWER: IF RESPONDENT SELECTS MORE THAN ONE DIFFICULTY ENTER THE CODE FOR THAT CONDITION AND GO TO THE ROUTED QUESTIONS. ONCE COMPLETE, ENTER THE CODE FOR THE NEXT CONDITION AND GO TO THE ROUTED QUESTIONS UNTIL ALL DIFFICULTIES ARE ENTERED. WHEN ALL CONDITIONS ARE ENTERED OR PATIENT HAS NO DIFFICULTIES GO TO BEHAVIOURAL HEALTH SECTIONHAS IWER: SHOW CARD FL2.

FL002. Please look at card FL2. Because of a health or memory problem, do you have difficulty doing any of the activities on this card? Again exclude any difficulties you expect to last less than three months.

IWER: READ EACH ONE OF THE DIFFICULTIES

IWER: CODE ALL THAT APPLY

1. Dressing, including putting on shoes and socks **GO TO FL003** **[fl002_1]**

2. Walking across a room **[fl002_2]**

3. Bathing or showering **[fl002_3]**

4. Eating, such as cutting up your food **[fl002_4]**

5. Getting in or out of bed **[fl002_5]**

6. Using the toilet, including getting up or down **[fl002_6]**

96. None of these **Go to Behavioural Health** **[fl002_9]**

98. DK **Go to Behavioural Health** **[fl002_7]**

99. RF **Go to Behavioural Health** **[fl002_8]**

IWER: PROBE: ANY OTHERS?

IF respondent states difficulty with any of the items from 1-3, please ask the follow up questions below. If there are no difficulties or refuse or don't know then section of questionnaire is finished. Go to Behavioural Health Section

******If Problems with Dressing*******

FL003: Do you ever use equipment or devices to help you get dressed?

1. Yes

5. No **GO TO FL005**

98. DK **GO TO FL005**

99. RF **GO TO FL005**

FL005: Does anyone ever help you with dressing including putting on shoes and socks?

1. Yes

5. No

98. DK

99. RF

******If Problem walking across a room************************

FL006: Do you ever use equipment or devices such as a walking stick or frame when crossing a room?

1. Yes

5. No **GO TO FL008**

98. DK **GO TO FL008**

99. RF **GO TO FL008**

(HRS/SHARE/ELSA)

FL007: Which equipment is that? ***(code response)***

1. Walking stick	**[fl007_01]**
2. Walking frame	**[fl007_02]**
3. Crutches	**[fl007_03]**
4. Railing	**[fl007_04]**
5. Orthopaedic shoes	**[fl007_05]**
6. Brace (leg or back)	**[fl007_06]**
7. Limb prosthesis	**[fl007_07]**
8. Oxygen/Respirator	**[fl007_08]**
9. Furniture or walls	**[fl007_09]**
10. Wheelchair or cart	**[fl007_10]**
95. Other (specify)	**[fl007_11]**
98. DK	**[fl007_12]**
99. RF	**[fl007_13]**

FL008: Does anyone ever help you with walking across a room?

1. Yes

5. No

98. DK

99. RF

************************If problem bathing or showering******************************

FL009: Do you ever use equipment or devices such as a shower seat, grab rails, hand-held shower when bathing or showering?

1. Yes

5. No **GO TO FL011**

98. DK

99. RF **GO TO FL011**

FL010: Which equipment is that?

1. Shower seat		**[fl010_1]**
2. Grab rails		**[fl010_2]**
3. Hand-held shower		**[fl010_3]**
4. Walking frame or stick		**[fl010_4]**
5. Rubber mat		**[fl010_5]**
95. Others (specify)	**[fl010_oth]**	**[fl010_6]**
98. DK		**[fl010_7]**
99. RF		**[fl010_8]**

FL011: Does anyone ever help you with bathing or showering?

1. Yes

5. No

98. DK

99. RF

Behavioural Health Smoking

BH001: Have you ever smoked cigarettes, cigars, cigarillos or a pipe daily for a period of at least one year?

1. Yes **GO TO BH002**

5. No **Questionnaire Finished**

98. DK **Questionnaire Finished**

99. RF **Questionnaire Finished.**

BH002: Do you smoke at the present time?

IWER: IF RESPONDENT SMOKED IN THE PAST 3 MONTHS CODE 1

1. Yes **GO TO BH004**

5. No, I have stopped **GO TO BH003**

98. DK **GO TO BH003**

99. RF **GO TO BH003**

BH003: How old were you when you stopped smoking?

1 … 100 _____

98. DK

99. RF

Questionnaire Finished

1.2.2 Descriptive Statistics

Table A.6 Independent sample responses to EQ-5D-3L domains

	EQ-5D-3L Level					
	1		2		3	
	n	%	n	%	n	%
Mobility	69	57.02%	52	42.98%	0	0.00%
Self-Care	107	88.43%	13	10.74%	1	0.83%
Usual Activities	75	61.98%	38	31.40%	8	6.61%
Pain/Discomfort	63	52.07%	52	42.98%	6	4.96%
Anxiety/Depression	94	77.69%	27	22.31%	0	0.00%

Table A.7 Summary of independent sample responses to TILDA questions included in Mapping Model

	Number	%
Mobility		
No Difficulty Walking 100m	85	70.25%
Difficulty Walking 100m	36	29.75%
Self-Care		
No Difficulty Bathing or Showering	103	85.12%
Reported Difficulty Bathing or Showering	8	6.61%
Receives help when bathing or showering	10	8.26%
Usual Activities		
No Long-term Health Problem /Not Limited because of a health problem in activities people usually do	60	49.59%
Limited, but not severely	43	35.54%
Severely Limited	18	14.88%
Pain		
No Pain most of the time	68	56.20%
Mild Moderate most of the time	42	34.71%
Severe most of the time	11	9.09%
Anxiety/Depression		
Excellent/Very Good/Good Emotional or Mental Health	108	89.26%
Fair/Poor Emotional or Mental Health	13	10.74%

1.2.3 Correlation Coefficients between TILDA Question Subset and EQ-5D-3L

Table A.8 Spearman correlation coefficient for all question subset variables considered for mapping model from Independent Sample

Table A.8.1 Mobility

			Comparable TILDA Question Subset Variable			
Spearman Correlation Coefficient		0.28***	0.53***	0.61***	0.67***	0.67***
EQ-5D Response Levels	1. I have no problems in walking about	No difficulty walking across a room	No difficulty climbing several flights of stairs	No difficulty climbing one flight of stairs	No Difficulty Walking 100m,	**No Difficulty Walking 100m,**
	2. I have some problems in walking about	Difficulty walking across a room	Difficulty climbing several flights of stairs	Difficulty Climbing One flight of stairs	Difficulty Walking 100m,	**Difficulty Walking 100m** Or **Difficulty Walking Across a Room**
	3. I am confined to bed	Needs equipment walking across a room			Difficulty Walking Across a Room	

Table A.8.2 Self Care

		Comparable TILDA Question Subset Variable			
Spearman Correlation Coefficient		0.39***	0.41***	0.81***	0.80***
EQ-5D Response Levels	1. I have no problems with self-care	No difficulty dressing including shoes and socks	No difficulty dressing including shoes and socks	No Difficulty Bathing or Showering	**No Difficulty Bathing or Showering**
	2. I have some problems washing or dressing myself	Reported difficulty dressing including shoes and socks	Reported difficulty dressing including shoes and socks	Reported Difficulty Bathing or Showering	**Reported Difficulty Bathing or Showering**
	3. I am unable to wash or dress myself	Uses equipment dressing including shoes and socks	Needs Help Dressing including shoes and socks	Uses Equipment when bathing or showering	**Receives help when bathing or showering**

Table A.8.3 Usual Activities

			Comparable TILDA Question Subset Variable
Spearman Correlation Coefficient			0.62***
EQ-5D Response Levels	1.	I have no problems with performing my usual activities	**No Long term Health Problem /Not Limited because of a health problem in activities people usually do**
	2.	I have some problems with performing my usual activities	**Limited, but not severely**
	3.	I am unable to perform my usual activities	**Severely Limited**

Table A.8.4 Pain Discomfort

			Comparable TILDA Question Subset Variable	
Correlation Coefficient			0.70***	0.67***
EQ-5D Response Levels	1.	I have no pain or discomfort	No Pain most of the time	**No Pain most of the time**
	2.	I have moderate pain or discomfort	Mild most of the time	**Mild or Moderate most of the time**
	3.	I have extreme pain or discomfort	Moderate most of the time	
			Severe most of the time	**Severe most of the time**

Table A.8.5 Anxiety/Depression

			Comparable TILDA Question Subset Variable	
Correlation Coefficient			0.48***	0.46***
EQ-5D Response Levels			How would you rate your emotional or mental health?	**How would you rate your emotional or mental health?**
	1.	I am not anxious or depressed	Excellent	**Excellent, Very Good or Good**
	2.	I am moderately anxious or depressed	Very Good	
	3.	I am extremely anxious or depressed	Good	
			Fair	**Fair or Poor**
			Poor	

1.2.4 Mapping model results

Table A.9 Model coefficient and standard errors for the six mapping models

	OLS	Tobit	ALDV	Beta	BetaInf	BetaInfTrun
Mobility						
No difficulty walking 100m						
Difficulty walking 100m	-0.116*	-0.152**	-0.124**	-1.008**	-0.267	-0.428*
	-0.0447	-0.0524	-0.0457	-0.38	-0.141	-0.212
Self Care						
No difficulty bathing or showering						
Reported difficulty bathing or showering	-0.152	-0.155	-0.149	-0.658	-0.432	-0.581
	-0.0808	-0.0832	-0.0788	-0.376	-0.237	-0.312
Receives help bathing or showering	-0.240***	-0.233***	-0.234***	-0.817**	-0.689***	-0.801***
	-0.0632	-0.0622	-0.0616	-0.253	-0.165	-0.19
Usual Activities						
No long term health problem/Not Limited because of a health problem in activities people usually do						
Limited, but not severely	-0.107**	-0.194***	-0.145***	-0.687*	-0.259	-0.882*
	-0.0341	-0.0483	-0.0387	-0.335	-0.15	-0.365
Severely Limited	-0.133*	-0.204**	-0.161**	-0.748	-0.336	-0.912*
	-0.0545	-0.0625	-0.0548	-0.409	-0.195	-0.381
Pain/Discomfort						
No pain most of the time						
Mild/Moderate Pain	-0.0716*	-0.133**	-0.0937*	-0.765*	-0.0978	-0.193
	-0.0341	-0.0477	-0.0383	-0.337	-0.12	-0.232
Severe Pain	-0.340***	-0.420***	-0.366***	-1.841***	-0.793***	-1.115***
	-0.0781	-0.0791	-0.0766	-0.341	-0.218	-0.31
Anxiety Depression						
Excellent/Very Good/Good Emotional or Mental Health						
Fair/Poor Emotional or Mental Health	-0.109	-0.0999	-0.104	-0.45	-0.33	-0.337
	-0.0569	-0.0613	-0.0572	-0.28	-0.177	-0.233
cons	0.965***	1.119***	1.016***	3.783***	1.942***	3.140***
	-0.012	-0.0355	-0.028	-0.147	-0.118	-0.344
lns_1						
_cons		-1.670***	-1.862***			

	OLS	Tobit	ALDV	Beta	BetaInf	BetaInfTrun
		-0.0723	-0.0782			
C1_lnphi						
_cons				1.576***	3.258***	2.556***
				-0.132	-0.166	-0.189
PM_ub						
_cons					-0.385*	-0.385*
					-0.186	-0.186
Standard errors in parentheses	* p<0.05		** p<0.01		*** p<0.001	

ALDV, Adjusted Limited Dependent Variable Model; BetaInf, Beta inflated; BetaInfTrun, Beta Inflated Truncated; cons, constant; m, metres; OLS, Ordinary Least Squares.

Table A.10 Mapping goodness of fit results by observed utility subgroups

	MAE	RMSE
Observed EQ-5D: <0.33 n=11		
OLS	0.2101	0.2275
Tobit	0.2018	**0.2258**
ALDV	0.2076	0.2265
Beta	**0.1921**	0.2310
Beta Inflated	0.4760	0.4796
Beta Inflated Truncated	0.4723	0.4774
Observed EQ-5D : 0.33 - <0.66 n= 17		
OLS	**0.1437**	**0.1673**
Tobit	0.1551	0.1802
ALDV	0.1486	0.1720
Beta	0.1959	0.2414
Beta Inflated	0.1552	0.1758
Beta Inflated Truncated	0.1566	0.1768
Observed Eq-5D 0.66- <=1 n=93		
OLS	**0.0809**	**0.1055**
Tobit	0.0823	0.1073
ALDV	0.0844	0.1068
Beta	0.0849	0.1079
Beta Inflated	0.1043	0.1128
Beta Inflated Truncated	0.1011	0.1084

The model with the lowest error is highlighted in bold.
ALDV, Adjusted Limited Dependent Variable Model; MAE, Mean Absolute Error; OLS, Ordinary Least Squares; RMSE, Root Mean Squared Error.

Table A.11 Variance-co-variance matrix for the chosen ALDV mapping model

Variance -Covariance Matrix										
	Comp_1: Difficulty walking 100m	**Comp_1:** Reported difficulty bathing or showering	**Comp_1:** Receives help bathing or showering	**Comp_1:** Usual Activities: Limited, but not severely	**Comp_1:** Usual Activities Severely Limited	**Comp_1:** Mild/Moderate Pain	**Comp_1:** Severe Pain	**Comp_1:** Fair/Poor Mental Health	**Comp_1:** _cons	**lns_1:** _cons
Difficulty walking 100m	0.002091									
Reported difficulty bathing or showering	-0.00063	0.006215								
Receives help bathing or showering	-0.00054	0.000674	0.0038							
Usual Activities: Limited, but not severely	-0.00048	0.000483	3.67E-05	0.001497						
Usual Activities Severely Limited	-0.00141	1.73E-05	-0.00032	0.001014	0.003001					
Mild/Moderate Pain	9.35E-05	-0.00125	0.000168	-0.00045	-0.00054	0.001471				
Severe Pain	0.000245	0.000391	-0.00147	-0.00031	-0.00069	0.000706	0.005871			
Fair/Poor Mental Health	-0.00036	0.000874	6.20E-06	-0.00019	0.000142	-0.00046	-0.00072	0.003273		
Comp_1:_cons	-0.00025	6.84E-05	-8.8E-05	-0.00046	-0.00025	-0.00036	-0.00027	0.000127	0.000783	
lns_1:_cons	-0.00167	0.000864	-0.00051	0.000209	0.001501	-0.00079	-4.5E-05	0.000767	0.000948	0.006117

ALDV, Adjusted Limited Dependent Variable Model; m, metre.

1.3 Chapter 6 Appendices

Table A.12 Exemplar health state utilities reported in studies identified by Ara et al (1)

Author Year	Pop	Baseline	Angina Year 1(≥2)	MI Year 1(≥2)	Stroke Year 1(≥2)	CHD plus Stroke	Acute/ Chronic	Details of how HSUs are combined
Aarnio 2015 (2)	primary	Age-adjust: 45-54=0.876 55-64=0.821 65-74=0.781	NM	-0.092(-0.011)	NM	NR	Yes	Apply a constant absolute decrement on age-adjusted values Unclear how HSU for more than one different event is modelled
Amirsardi 2015 (3)	primary	Age-adjust NR	NM	0.760(0.880)	NM	NA	Yes	Unclear how MI HSUs are combined with age-adjusted baseline as no description provided
Amirsardi 2017 (4)	primary	Age-adjust NR	NM	0.760(0.880)	NM	NA	Yes	Unclear how MI HSUs are combined with age-adjusted baseline as no description provided
Ferket 2017 (5)	primary	Age-adjust: HSU=0.95698 + 0.02465*male - 0.00085*age - 0.00002*Age^2	-0.181 (-0.018)	-0.105 (-0.060)	-0.206 (-0.126)	NM	Yes	Assume the values reported are absolute decrements applied to the age-adjusted baseline. In addition to different acute and chronic HSUs a transition reward representing hospitalisation is applied in the first year, the duration of this decrement, and the proportion who receive this is not reported. More than one event not modelled
Galper 2015 (6)	primary	Constant 1.000	NM	0.829(0.865)	NM	NA	Yes	Baseline HSU is not described thus assumed full health Acute HSU (0.829) applied for 8 days only
Heller 2017 (7)	primary	Not possible to determine either the health states or corresponding HSUs from the article or cited references: 'Each health state and event has an annual cost and quality-of-life adjustment. More details are published elsewhere'.						
Lin 2015 (8)	primary	Constant 1.000	NM	M: 0.760(0.760) F: 0.680(0.680)	M: 0.660(0.72) F: 0.620(0.680)	NR	No	Constant absolute decrements applied Unclear how HSU for more than one different event is modelled
Pandya 2015 (9)	primary	Constant 1.000	0.768(0.768)	0.778(0.778)	0.768(0.768)	Minimum	No	Constant absolute decrements applied For comorbidities states: 'if an individual has had multiple CVD events, the individual remains in the health state of the more severe event' (assumed minimum HSU)
Roberts 2015 (10)	primary	Age-adjust: 50=0.840 60=0.820 70=0.790 80=0.740 90=0.680	0.835(0.835	0.835(0.835)	0.827(0.827)	NR	No	CV HSUs applied as constant relative decrements on the age-adjusted Unclear how the HSU for more than one different event is modelled
Shiffman 2016 (11)	primary	Constant 1.000	NM	0.778(0.778)	0.768(0.768)	0.605 (0.605)	No	Constant absolute decrements applied

Author Year	Pop	Baseline	Angina Year 1(≥2)	MI Year 1(≥2)	Stroke Year 1(≥2)	CHD plus Stroke	Acute/ Chronic	Details of how HSUs are combined
Mixed								
Arrieta 2017 (12)	mixed	Constant 0.790	0.630-0.790 over 5 year	0.580-0.790 over 5 year	0.460-0.790 over 5 year	NR	Yes	HSUs for CV events increase to baseline value in non-linear fashion over 5 year period (no additional details) Unclear how HSU for more than one different event is modelled
Chen 2015 (13)	HeFH (P&S)	Constant 0.996	0.680(0.680)	0.680(0.680)	0.680(0.680)	NA	No	Have one event health state: 'CVD/stroke' and state the reported HSU was adjusted by age group for use in the model (data not provided)
Gandra 2016 (14)	HeFH & second ary	Constant 0.824	0.672(0.824)	0.672(0.824)	0.327(0.524)	minimum	Yes	Baseline for primary HeFH=0.824 Baseline for secondary (established CVD) is not provided For comorbidities, the minimum method is used e.g. if had stroke then MI, the HSUs for stroke would be applied to the MI health state NB: evnet =0.80; IS = 1
Kazi 2016 (15)	HeFH	Constant 1.000	0.008(0.906)	0.008(0.965)	0.011(0.884)	NR (0.852)	Yes	Transient toll is applied for initial event but unclear what the duration is, assume the same toll is applied for more than one different event
	second ary	NR	0.008(0.906)	0.008(0.965)	0.0113(0.884)	NR (0.852)	Yes	Transient toll is applied for initial event but unclear what the duration is; assumed one year. Assume the same toll is applied if more than one different event. Assume baseline for entering the model is the full health as not described in the text
Korman 2017 (16)	primar y	Age-adjust: 40-49=0.940; 50-59=0.930; 60-69=0.900; 70-79=0.860; 80+=0.700	NA	0.710(0.800)	Minor 0.740(0.740) Moderate 0.650(0.650) Severe 0.410(0.410)	NR	some	Initial acute HSU for MI applied for 11 days Unclear how HSU for events are combined with age adjusted baseline – assume additive
	second ary	NR	NA	0.71 (0.80)	Minor 0.740(0.740) Moderate 0.650(0.650) Severe 0.410(0.410)		some	Baseline HSU: unclear what HSUs are assigned on entering the model; assumed 0.80 HSU for chronic MI
Laires 2015 (17)	primar y	Constant 1.000	0.770(0.808)	0.760(0.808)	NM	NA	Yes	Constant absolute decrements applied
	second ary	Constant 0.808	0.770(0.808)	0.760(0.808)	NM	NA	Yes	Constant absolute decrements applied
Ribeiro 2015(18)	primar y	Constant 0.800	0.740(0.740)	0.740(0.740)	0.600(0.600)	NR	No	Constant absolute decrements applied Unclear how HSU for more than one different event is modelled

Author Year	Pop	Baseline	Angina Year 1(≥2)	MI Year 1(≥2)	Stroke Year 1(≥2)	CHD plus Stroke	Acute/ Chronic	Details of how HSUs are combined
	secondary	NR	0.740(0.740)	0.74 (0.74)	0.600(0.600)		No	Unclear what baseline HSUs are for secondary patients Unclear how HSU for more than one different event is modelled
Secondary Prevention								
Almalki 2017 (19)	secondary	Constant 0.820	0.780(0.780)	0.800(0.800)	IS:0.610(0.610) HS:0.390(0.390)	NR	No	Apply a constant decrement
Barrios 2017(20)	secondary	Constant 0.836	0.760(0.836)	0.760(0.836)	0.629(0.692)	NR	Yes	Apply a constant absolute decrement Unclear how HSU for more than one different event is modelled
Becarra 2015 (21)	secondary	Constant 0.836	0.760(0.836)	0.760(0.836)	0.629(0.692)	NR	No	Apply a constant decrement Unclear how HSU for more than one different event is modelled
Ito 2015 (22)	secondary	Constant 0.850	NM	0.770(0.850)	0.600(0.650)	multiply	Yes	Constant relative decrement applied for more than one different event
Stam-Slob 2017(23)	secondary	Constant 0.780	NM	0.650(0.650)	0.640(0.640)	0.620	No	Constant absolute decrements applied MACE HSU (0.62) is for patients who experience two or more of the events within a given year, assume if not within same year, HSU for specific event is applied irrespective of history

"Reprinted with adaptions from Ara et al. 2020. Appendix. Copyright (2020), with permission from Elsevier

The study by Davies et al was removed from this table as the data in the table referring to Davies 2017 did not match the reference in the paper which was published in 2006.

CHD, Coronary Heart Disease; CV, cardiovascular; CVD, Cardiovascular Disease, HeFH,, Heterozygous Familial Hypercholesterolemia, HSU, Health State Utility; MACE, Major Adverse Cardiovascular Event; MI Myocardial Infarction, NA, Not applicable; NM, Not modelled; NR, Not reported; Pop, Population.

1. Ara R, Hill H, Lloyd A, Woods HB, Brazier J. Are Current Reporting Standards Used to Describe Health State Utilities in Cost-Effectiveness Models Satisfactory? Value Health. 2020 Feb;0(0).
2. Aarnio E, Korhonen MJ, Huupponen R, Martikainen J. Cost-effectiveness of statin treatment for primary prevention in conditions of real-world adherence–estimates from the Finnish prescription register. Atherosclerosis. 2015;239(1):240–7.
3. Amirsadri M, Hassani A. Cost-effectiveness and cost-utility analysis of OTC use of simvastatin 10 mg for the primary prevention of myocardial infarction in Iranian men. DARU J Pharm Sci. 2015;23(1):56.
4. Amirsadri M, Sedighi MJ. Cost-effectiveness evaluation of aspirin in primary prevention of myocardial infarction amongst males with average cardiovascular risk in Iran. Res Pharm Sci. 2017;12(2):144.
5. Ferket BS, Hunink MGM, Khanji M, Agarwal I, Fleischmann KE, Petersen SE. Cost-effectiveness of the polypill versus risk assessment for prevention of cardiovascular disease. Heart. 2017 Jan 11;heartjnl-2016-310529.
6. Galper BZ, Wang YC, Einstein AJ. Strategies for primary prevention of coronary heart disease based on risk stratification by the ACC/AHA Lipid Guidelines, ATP III Guidelines, Coronary Calcium Scoring, and C-Reactive Protein, and a Global Treat-All Strategy: a comparative--effectiveness m. PLoS One. 2015;10(9).
7. Heller DJ, Coxson PG, Penko J, Pletcher MJ, Goldman L, Odden MC, et al. Evaluating the impact and cost-effectiveness of statin use guidelines for primary prevention of coronary heart disease and stroke. Circulation. 2017;136(12):1087–98.
8. Lin L, Teng M, Zhao YJ, Khoo AL, Seet RCS, Yong QW, et al. Long-term cost-effectiveness of statin treatment for primary prevention of cardiovascular disease in the elderly. Cardiovasc drugs Ther. 2015;29(2):187–97.
9. Pandya A, Sy S, Cho S, Weinstein MC, Gaziano TA. Cost-effectiveness of 10-year risk thresholds for initiation of statin therapy for primary prevention of cardiovascular disease. Jama. 2015;314(2):142–50.
10. Roberts ET, Horne A, Martin SS, Blaha MJ, Blankstein R, Budoff MJ, et al. Cost-Effectiveness of Coronary Artery Calcium Testing for Coronary Heart and Cardiovascular Disease Risk Prediction to Guide

Statin Allocation: The Multi-Ethnic Study of Atherosclerosis (MESA). Steyerberg EW, editor. PLoS One. 2015 Mar 18;10(3):e0116377.
11. Shiffman D, Arellano AR, Caulfield MP, Louie JZ, Bare LA, Devlin JJ, et al. Use of low density lipoprotein particle number levels as an aid in statin treatment decisions for intermediate risk patients: a cost-effectiveness analysis. BMC Cardiovasc Disord. 2016;16(1):251.
12. Arrieta A, Page TF, Veledar E, Nasir K. Economic evaluation of PCSK9 inhibitors in reducing cardiovascular risk from health system and private payer perspectives. PLoS One. 2017;12(1 PG-).
13. Chen CX, Hay JW. Cost-effectiveness analysis of alternative screening and treatment strategies for heterozygous familial hypercholesterolemia in the United States. Int J Cardiol. 2015;181:417–24.
14. Gandra SR. Cost-Effectiveness of LDL-C Lowering With Evolocumab in Patients With High Cardiovascular Risk in the United States (vol 39, pg 313, 2016). Clin Cardiol. 2016;39(10):621.
15. Kazi DS, Moran AE, Coxson PG, Penko J, Ollendorf DA, Pearson SD, et al. Cost-Effectiveness of PCSK9 inhibitor therapy in patients with heterozygous familial hypercholesterolemia or atherosclerotic cardiovascular disease. JAMA - J Am Med Assoc. 2016;316(7 PG-743–753):743–53.
16. Korman T. M. W, Korman M, Wisløff T. Modelling the cost-effectiveness of PCSK9 inhibitors vs. ezetimibe through LDL-C reductions in a Norwegian setting. Eur Hear J Cardiovasc Pharmacother. 2017/04/27. 2017;
17. Laires PA, Ejzykowicz F, Hsu T-Y, Ambegaonkar B, Davies G. Cost–effectiveness of adding ezetimibe to atorvastatin vs switching to rosuvastatin therapy in Portugal. J Med Econ. 2015;18(8):565–72.
18. Ribeiro RA, Duncan BB, Ziegelmann PK, Stella SF, Vieira JL da C, Restelatto LMF, et al. Cost-effectiveness of high, moderate and low-dose statins in the prevention of vascular events in the Brazilian public health system. Arq Bras Cardiol. 2015;104(1):32–44.
19. Almalki ZS, Guo JJ, Alahmari A, Alotaibi N, Thaibah H. Cost-effectiveness of simvastatin plus ezetimibe for cardiovascular prevention in patients with a history of acute coronary syndrome: analysis of results of the IMPROVE-IT trial. Hear Lung Circ. 2018;27(6):656–65.
20. Barrios V, Kaskens L, Castellano JM, Cosin-Sales J, Ruiz JE, Zsolt I, et al. Usefulness of a cardiovascular polypill in the treatment of secondary prevention patients in Spain: a cost-effectiveness study. Rev Española Cardiol (English Ed. 2017;70(1):42–9.
21. Becerra V, Gracia A, Desai K, Abogunrin S, Brand S, Chapman R, et al. Cost-effectiveness and public health benefit of secondary cardiovascular disease prevention from improved adherence using a polypill in the UK. BMJ Open. 2015 May 1;5(5):e007111.
22. Ito K, Avorn J, Shrank WH, Toscano M, Spettel C, Brennan T, et al. Long-term cost-effectiveness of providing full coverage for preventive medications after myocardial infarction. Circ Cardiovasc Qual Outcomes. 2015;8(3):252–9.
23. Stam-Slob MC, van der Graaf Y, Greving JP, Dorresteijn JAN, Visseren FLJ. Cost-Effectiveness of Intensifying Lipid-Lowering Therapy With Statins Based on Individual Absolute Benefit in Coronary Artery Disease Patients. J Am Heart Assoc. 2017;6(2):e004648.

Table A.13 Population characteristics including initial predicted utility values

	Population	Predicted Utility		Male		Age	
		Mean (SE)	95% CI	Proportion (SE)	95% CI	Median (SE)	95% CI
1	Any Angina, MI, TIA or Stroke (Base-case)	0.7788 (0.008)	0.7631-0.7945	0.5711 (0.0168)	0.5711-0.0168	71 (0.5)	70.02-71.98
2	Any Angina MI TIA, Stroke or other CVD	0.7847 (0.0073)	0.7705-0.799	0.5719 (0.0153)	0.5719-0.0153	70 (0.5)	69.02-70.98
3	Any Angina, MI or other CVD	0.7834 (0.0082)	0.7673-0.7994	0.6023 (0.0174)	0.6023-0.0174	70 (0.5)	69.02-70.98
4	Any Angina or MI	0.7764 (0.0095)	0.7578-0.795	0.6117 (0.0195)	0.6117-0.0195	71 (0.75)	69.53-72.47
5	Any TIA or Stroke	0.7496 (0.0154)	0.7195-0.7798	0.4679 (0.0302)	0.4679-0.0302	72 (1)	70.04-73.96
6	Any Angina	0.7533 (0.0125)	0.7287-0.7779	0.5519 (0.0243)	0.5519-0.0243	71 (0.75)	69.53-72.47
7	Only Angina (No MI, TIA or stroke)	0.7736 (0.0149)	0.7443-0.8028	0.4962 (0.0319)	0.4962-0.0319	71 (1)	69.04-72.96
8	Any MI	0.7952 (0.0115)	0.7727-0.8178	0.7061 (0.0245)	0.7061-0.0245	70 (0.75)	68.53-71.47
9	Only MI (No Angina, TIA, or stroke)	0.8313 (0.0134)	0.805-0.8577	0.7575 (0.0322)	0.7575-0.0322	69 (1.25)	66.55-71.45
10	>1 MI	0.7422 (0.0243)	0.6944-0.79	0.7391 (0.0496)	0.7391-0.0496	67 (2)	63.08-70.92
11	Any TIA	0.777 (0.0181)	0.7416-0.8125	0.4241 (0.0392)	0.4241-0.0392	73 (1.25)	70.55-75.45
12	Only TIA (other heart excluded)	0.8348 (0.0175)	0.8005-0.8691	0.3724 (0.0489)	0.3724-0.0489	73 (1.25)	70.55-75.45
13	Any Stroke	0.7052 (0.0248)	0.6565-0.754	0.5258 (0.0464)	0.5258-0.0464	71 (1)	69.04-72.96
14	Only Stroke (No angina, MI, TIA, or other CVD)	0.7364 (0.0276)	0.6822-0.7906	0.5251 (0.0551)	0.5251-0.0551	71 (1.25)	68.55-73.45
15	>1 Stroke	0.5483 (0.0635)	0.4235-0.673	0.5785 (0.1084)	0.5785-0.1084	72 (1.5)	69.06-74.94
16	Other CVD	0.8153 (0.016)	0.7839-0.8467	0.5761 (0.0371)	0.5761-0.0371	65 (1.25)	62.55-67.45
17	Only other CVD	0.7964 (0.0131)	0.7707-0.822	0.6093 (0.0302)	0.6093-0.0302	67 (1.25)	64.51-69.45
18	MI or Stroke	0.7793 (0.0107)	0.7582-0.8004	0.6605 (0.0231)	0.6605-0.0231	70 (0.75)	68.53-71.47
19	>1 MI or > 1 Stroke	0.7672 (0.0224)	0.7233-0.8111	0.7475 (0.0508)	0.7475-0.0508	67 (2)	63.08-70.92
20	Disease in two vascular beds i.e. (MI or Angina) and (TIA or stroke)	0.6362 (0.0378)	0.5620-0.7104	0.5370 (0.0613)	0.5370-0.0613	73 (1.5)	70.06-75.94
21	As population 1 but excluding Population 20	0.791 (0.0079)	0.7755-0.8066	0.574 (0.0177)	0.574-0.0177	70 (0.5)	77.02-78.98
22	Diabetes and CVD (Angina, MI, TIA or Stroke)	0.7352 (0.0232)	0.6897-0.7808	0.6197 (0.0445)	0.5322-0.7071	71 (1.75)	67.5701-74.4299
23	Diabetes and MI	0.7883 (0.0238)	0.7414-0.8351	0.8289 (0.0565)	0.7179-0.9399	71 (2.25)	66.5901-75.4099

CI, Confidence Interval; CVD, cardiovascular disease; MI, myocardial infarction; SE, Standard Error; TIA, Transient Ischemic Attack.

1.4 Chapter 7 Appendices

1.4.1 Costs

Table A.14 Table of cardiovascular cost parameters applied in the model per 90-day cycle

Parameter	Definition	Cost (€)	SOURCE
Event costs			
fatalCVD	Cost of a fatal cardiovascular event	2878.90	(1)
fatalNONCVD	Cost of a fatal non cardiovascular event	2414.46	(1)
firsteventMI	Cost of myocardial infarction in first quarter following event	6762.46	(2)
MIdiabetes	Additional cost of myocardial infarction in first quarter for patients with diabetes	0	(1)
firsteventMI2	Cost of myocardial infarction in second quarter following event	1680.79	(1)
MIdiabetes2	Additional cost of myocardial infarction in second quarter for patients with diabetes	1447.81	(1)
firsteventMI3	Cost of myocardial infarction in third quarter following event	889.33	(1)
MIdiabetes3	Additional cost of myocardial infarction in third quarter for patients with diabetes	917.55	(1)
firsteventMI4	Cost of myocardial infarction in fourth quarter following event	937.98	(1)
MIdiabetes4	Additional cost of myocardial infarction in fourth quarter for patients with diabetes	560.44	(1)
feMI	Cost of myocardial infarction in all subsequent quarters following event	669.08	(1)
feMIdiabetes	Additional cost of myocardial infarction in all subsequent for patients with diabetes	388.55	(1)
firsteventStroke_I	Cost of ischemic stroke in first quarter following event	9001.32	(2)
firsteventStroke_I2	Cost of ischemic stroke in second quarter following event	1599.38	(1)
firsteventStroke_I3	Cost of ischemic stroke in third quarter following event	938.53	(1)
firsteventStroke_I4	Cost of ischemic stroke in fourth quarter following event	749.09	(1)
feSTROKE_I	Cost of ischemic stroke in all subsequent quarters following event	622.61	(1)
firsteventStroke_H	Cost of haemorrhagic stroke in first quarter following event	9001.32	(1)
firsteventStroke_H2	Cost of haemorrhagic stroke in second quarter following event	2109.26	(1)
firsteventStroke_H3	Cost of haemorrhagic stroke in third quarter following event	813.67	(1)
firsteventStroke_H4	Cost of haemorrhagic stroke in fourth quarter following event	546.32	(1)
feSTROKE_H	Cost of haemorrhagic stroke in all subsequent quarters following event	930.74	(1)
Background cost coefficients for quarter costs			
age0	Baseline age - 70	8.61	(1)
timeperiod	Additional cost to be added to background cost with each model cycle	9.08	(1)
diabetes	History of diabetes	269.54	(1)
hist_liver	History of liver disease	387.74	(1)
hist_hf	History of heart failure	344.34	(1)
hist_af	History of atrial fibrillation	306.51	(1)
hist_pad	History of peripheral artery disease	337.04	(1)
hist_copd	History of chronic obstructive pulmonary disease	196.88	(1)
hist_cancer	History of cancer	213.66	(1)
hist_renal	History of renal disease	580.62	(1)

Parameter	Definition	Cost (€)	SOURCE
sex	Female	-31.63	(1)
CHD	Other Coronary Heart Disease	3.34	(1)
NSTEMI	NSTEMI – Non ST elevated Myocardial Infarction	202.22	(1)
STEMI	STEMI - ST elevated Myocardial Infarction	39.70	(1)
UA	Unstable Angina	174.16	(1)
_cons	Constant	311.61	(1)
Consultant Cost	Pro-rata cost of annual consultant visit for each 90 day cycle	€33.66	(3)
PCSK9 Initiation Costs			(1)
Consultant 1 and 2	Outpatient consultant appointment cost Assumed additional outpatient consultant appoint ment required at initiation and after three months follow-up	136.50	(3)
Training Costs	Nurse Training Cost – Assumed cost equivalent to outpatient visit	136.50	(3)

All costs are presented in €2019.

1. Asaria M, Walker S, Palmer S, Gale CP, Shah AD, Abrams KR, et al. Using electronic health records to predict costs and outcomes in stable coronary artery disease. Heart. 2016;102(10):755–62.
2. Healthcare Pricing Office (HPO). ABF 2019 Admitted Price List. Dublin: HSE; 2019. Available from: http://www.hpo.ie/abf/ABF2019AdmittedPatientPriceList.pdf Date accessed 31 March 2020.
3. Health Service Executive (HSE) . HSE Ready Reckoner. Dublin: HSE; 2013. Available from: www.hse.ie/eng/services/list/1/schemes/cbd/Ready%20Reckoner.pdf Date accessed 14 September 2017.

Baseline cost

The baseline cost is calculated as follows

= "_cons" + (70-age)* "age0"+ proportion of the population reporting each characteristic by each of the remaining background cost-coefficients.

The baseline cost is increased by time period coefficient each cycle.

Event Costs

In the cycle directly following an MI, the value of the firsteventMI parameter is added to the baseline cost. For cycles 2-4 following the event, the firsteventMI2-4 parameters are added respectively. For each subsequent cycle following the event, the feMI parameter is added. As there is an interaction between MI and diabetes, additional event parameters are added for MI and diabetes, weighted by the proportion of patients reporting diabetes.

Event costs for ischemic stroke and haemorrhagic stroke are accounted for in the same way.

Once off costs for fatal CVD and non-fatal CVD are applied.

*Table A.15 MI and stroke costs from the HPO activity based funding price list**

	Myocardial Infarction DRGs	**Weighted Average Cost**	**€6,762.46**
DRG	DRG Description	Indicative inpatient cases	inpatient inlier price
F10	INTERVENT CRNRY PR + AMI, MAJC	270	€15,147
F10B	INTERVENT CRNRY PR + AMI, MINC	2,145	€8,085
F41A	CRC DSRD+AMI+INVA INV PR, MAJC	182	€9,463
F41b	CRC DSRD+AMI+INVA INV PR, MINC	448	€5,167
F60A	CIRC DIS+AMI-INVA INV PR	2,279	€5,770
F60n	CIRC DIS+AMI-INVA INV PR,T<5D	558	2,075

Stroke DRGs		**Weighted Average Costs**	**€9,001.32**
DRG	DRG Description	Indicative inpatient cases	inpatient inlier price
B70A	STROKE & OTH CEREB DIS, MAJC	805	€27,598
B70B	STROKE & OTH CEREB DIS, INTC	2267	€8,245
B70C	STROKE & OTH CEREB DIS, MINC	2645	€4,831
B70D	STROKE & OTH CEREB DIS, TR<5D	321	€2,069

4. * Healthcare Pricing Office (HPO). ABF 2019 Admitted Price List. Dublin: HSE; 2019. Available from: http://www.hpo.ie/abf/ABF2019AdmittedPatientPriceList.pdf Date accessed 31 March 2020.

1.4.2 Population Characteristics

Table A.16a Modelled patient characteristics for the TILDA populations

Population		Hist Renal disease (Prp.)	Hist COPD (Prp.)	Hist Cancer (Prp.)	hist liver (Prp.)	Depre-ssion (Prp.)	hist_ anxiety (Prp.)	Age (yrs)	pulse_ rate bpm	HDL-C mmol /L	TCHOL mmol /L	CREAT mmol /L	WCC 10^9/L	HGB g/100 ml	hist Stroke TIA only	Male age (yrs)	Femal-e Age (yrs)
1	Any Angina, MI, TIA or Stroke (Base-case)	0.07	0.07	0.08	0.01	0.06	0.03	71	72	1.56	6.17	100	7.46	13.58	0.22	70	72
2	Any Angina MI TIA, Stroke or other CVD	0.07	0.07	0.07	0.01	0.06	0.03	70	72	1.56	6.17	100	7.46	13.58	0.19	69	71
3	Any Angina, MI or other CVD	0.07	0.07	0.07	0.01	0.06	0.03	70	72	1.56	6.17	100	7.46	13.58	0	69	71
4	Any Angina or MI	0.07	0.07	0.07	0.01	0.06	0.03	71	72	1.56	6.17	100	7.46	13.58	0	70	72
5	Any TIA or Stroke	0.07	0.08	0.09	0	0.06	0.03	72	72	1.56	6.17	100	7.46	13.58	0.69	72	73
6	Any Angina	0.07	0.08	0.08	0.01	0.06	0.03	71	72	1.56	6.17	100	7.46	13.58	0	70	73
7	Only Angina (No MI, TIA or stroke)	0.07	0.07	0.08	0.02	0.06	0.03	71	72	1.56	6.17	100	7.46	13.58	0	68	73
8	Any MI	0.07	0.08	0.06	0	0.05	0.02	70	72	1.56	6.17	100	7.46	13.58	0	70	71
9	Only MI (No Angina, TIA, or stroke)	0.07	0.06	0.04	0	0.04	0.02	69	72	1.56	6.17	100	7.46	13.58	0	70	69
10	>1 MI	0.07	0.13	0.06	0.01	0.06	0.03	67	72	1.56	6.17	100	7.46	13.58	0	67	68
11	Any TIA	0.07	0.07	0.09	0.01	0.06	0.03	73	72	1.56	6.17	100	7.46	13.58	0.68	70	74
12	Only TIA (other heart excluded)	0.07	0.06	0.09	0	0.05	0.02	73	72	1.56	6.17	100	7.46	13.58	1	69	75
13	Any Stroke	0.07	0.09	0.09	0.01	0.07	0.03	71	72	1.56	6.17	100	7.46	13.58	0.69	72	69
14	Only Stroke (No angina, MI, TIA, or other CVD)	0.07	0.08	0.09	0	0.07	0.03	71	72	1.56	6.17	100	7.46	13.58	1	72	69
15	>1 Stroke	0.07	0.06	0.08	0.04	0.1	0.05	72	72	1.56	6.17	100	7.46	13.58	0.5	74	68
16	Other CVD	0.07	0.05	0.06	0.01	0.08	0.04	65	72	1.56	6.17	100	7.46	13.58	0	65	65
17	Only other CVD	0.07	0.05	0.07	0.01	0.09	0.04	67	72	1.56	6.17	100	7.46	13.58	0	65	69
18	MI or Stroke	0.07	0.08	0.07	0	0.05	0.03	70	72	1.56	6.17	100	7.46	13.58	0.19	71	70
19	>1 MI or > 1 Stroke	0.07	0.12	0.07	0.01	0.07	0.03	67	72	1.56	6.17	100	7.46	13.58	0	67	66
20	Disease in two vascular beds i.e. (MI	0.07	0.11	0.07	0.01	0.05	0.02	73	72	1.56	6.17	100	7.46	13.58	0	73	73

Population		Hist Renal disease (Prp.)	Hist COPD (Prp.)	Hist Cancer (Prp.)	hist liver (Prp.)	Depre-ssion (Prp.)	hist_ anxiety (Prp.)	Age (yrs)	pulse_ rate bpm	HDL-C mmol /L	TCHOL mmol /L	CREAT mmol /L	WCC 10^9/L	HGB g/100 ml	hist Stroke TIA only	Male age (yrs)	Femal-e Age (yrs)
	or Angina) and (TIA or stroke)																
21	As population 1 but excluding Population 20	0.07	0.07	0.08	0.01	0.06	0.03	70	72	1.56	6.17	100	7.46	13.58	0.24	70	72
22	Diabetes and CVD (Angina, MI, TIA or Stroke)	0.07	0.06	0.06	0.01	0.09	0.04	71	72	1.56	6.17	100	7.46	13.58	0.5	67	72
23	Diabetes and MI	0.07	0.07	0.07	0	0.04	0.02	71	72	1.56	6.17	100	7.46	13.58	0	68	80

COPD Chronic Obstructive Pulmonary Disease; CREAT, creatinine; CVD, Cardiovascular Disease; HDL-C, High Density Lipoprotein Cholesterol; g/100ml; grams /100 millilitre; HGB, Haemoglobin; MI, myocardial infarction; Prp. , proportion; mmol/L, millimole/L; TCHOL, Total Cholesterol; TIA, Transient Ischemic Attack; TILDA , The Irish Longitudinal Study on Aging; WCC, White cell count; Yrs, years.

Table A.16b Modelled patient characteristics for the TILDA populations

Population		Sex	IMD5	CHD	NSTE-MI	STEMI	UA	early PCI	early CABG	Recur-ent_mi	nitrates_long	Smc. Curr-ent	Smc ex	HT	Diab-etes	hist_HF	hist_PAD	hist_AF	Hist stroke
											Proportion								
1	Any Angina, MI, TIA or Stroke (Base-case)	0.43	0.2	0.09	0.24	0.17	0.06	0.09	0.04	0.09	0.28	0.18	0.48	0.86	0.14	0.06	0.08	0.2	0.32
2	Any Angina MI TIA, Stroke or other CVD	0.43	0.2	0.08	0.2	0.14	0.05	0.09	0.04	0.08	0.28	0.18	0.48	0.84	0.15	0.06	0.08	0.19	0.27
3	Any Angina, MI or other CVD	0.4	0.2	0.1	0.25	0.18	0.06	0.09	0.04	0.1	0.28	0.18	0.49	0.86	0.16	0.07	0.08	0.2	0.1
4	Any Angina or MI	0.39	0.2	0.12	0.32	0.23	0.08	0.09	0.04	0.13	0.28	0.18	0.51	0.89	0.15	0.08	0.08	0.21	0.1
5	Any TIA or Stroke	0.53	0.2	0.02	0.09	0.07	0.01	0.09	0.04	0.07	0.28	0.19	0.46	0.79	0.13	0.04	0.08	0.18	1
6	Any Angina	0.45	0.2	0.18	0.19	0.13	0.11	0.09	0.04	0.11	0.28	0.17	0.5	0.92	0.16	0.07	0.08	0.25	0.11
7	Only Angina (No MI, TIA or stroke)	0.51	0.2	0.27	0	0	0.17	0.09	0.04	0	0.28	0.18	0.45	0.91	0.13	0.03	0.08	0.24	0
8	Any MI	0.29	0.2	0	0.59	0.41	0	0.09	0.04	0.23	0.28	0.18	0.55	0.88	0.16	0.11	0.08	0.19	0.13
9	Only MI (No Angina, TIA, or stroke)	0.24	0.2	0	0.59	0.41	0	0.09	0.04	0.17	0.28	0.21	0.52	0.85	0.13	0.08	0.08	0.12	0
10	>1 MI	0.26	0.2	0	0.59	0.41	0	0.09	0.04	1	0.28	0.2	0.59	0.92	0.19	0.23	0.08	0.34	0.22
11	Any TIA	0.58	0.2	0.02	0.09	0.06	0.01	0.09	0.04	0.06	0.28	0.18	0.44	0.81	0.12	0.03	0.08	0.21	1
12	Only TIA (other heart excluded)	0.63	0.2	0	0	0	0	0.09	0.04	0	0.28	0.18	0.4	0.75	0.06	0.01	0.08	0.17	1
13	Any Stroke	0.47	0.2	0.02	0.11	0.08	0.01	0.09	0.04	0.09	0.28	0.21	0.49	0.77	0.15	0.06	0.08	0.18	1
14	Only Stroke (No angina, MI, TIA, or other CVD)	0.47	0.2	0	0	0	0	0.09	0.04	0	0.28	0.21	0.48	0.72	0.15	0.03	0.08	0.12	1
15	>1 Stroke	0.42	0.2	0.01	0.24	0.17	0.01	0.09	0.04	0.25	0.28	0.18	0.54	0.87	0.13	0.06	0.08	0.38	1
16	Other CVD	0.42	0.2	0	0	0	0	0.09	0.04	0	0.28	0.16	0.44	0.72	0.18	0.03	0.08	0.14	0
17	Only other CVD	0.39	0.2	0.04	0.08	0.06	0.03	0.09	0.04	0.05	0.28	0.16	0.48	0.8	0.16	0.07	0.08	0.17	0.09
18	MI or Stroke	0.34	0.2	0.01	0.45	0.32	0	0.09	0.04	0.18	0.28	0.19	0.53	0.85	0.16	0.09	0.08	0.18	0.32
19	>1 MI or > 1 Stroke	0.25	0.2	0	0.59	0.41	0	0.09	0.04	1	0.28	0.19	0.58	0.93	0.2	0.23	0.08	0.33	0.17
20	Disease in two vascular beds i.e. (MI or Angina) and (TIA or stroke)	0.46	0.2	0.09	0.4	0.28	0.05	0.09	0.04	0.27	0.28	0.15	0.57	0.92	0.23	0.13	0.08	0.29	1

Population		Sex	IMD5	CHD	NSTE-MI	STEMI	UA	early PCI	early CABG	Recur-ent_mi	nitrates_long	Smc. Curr-ent	Smc ex	HT	Diab-etes	hist_HF	hist_PAD	hist_AF	Hist stroke
											Proportion								
21	As population 1 but excluding Population 20	0.43	0.2	0.09	0.23	0.16	0.06	0.09	0.04	0.08	0.28	0.19	0.48	0.85	0.13	0.06	0.08	0.19	0.27
22	Diabetes and CVD (Angina, MI, TIA or Stroke)	0.38	0.2	0.1	0.28	0.19	0.06	0.09	0.04	0.13	0.28	0.16	0.58	0.94	1	0.1	0.08	0.23	0.31
23	Diabetes and MI	0.17	0.2	0	0.59	0.41	0	0.09	0.04	0.27	0.28	0.14	0.66	0.96	1	0.13	0.08	0.29	0.17

AF, atrial fibrillation; CABG, coronary artery bypass graft, CHD, Coronary Heart disease; CVD; cardiovascular disease; HF, heart failure ; HT, hypertension; IMD5, most deprived quintile; MI, Myocardial Infarction; NSTEMI, Non-ST elevation Myocardial Infarction; PAD, peripheral arterial disease; PCI, percutaneous coronary intervention; Smc, smoker; STEMI, ST elevation myocardial infarction; UA, Unstable Angina; TIA, transient ischemic attack; TILDA, The Irish Longitudinal Study on Aging.

1.4.3 Results

Table A.17 Total and incremental costs and outcomes for PCSK9 inhibitors for 23 subgroups from TILDA with a modelled of 4mmol/L under the indirect treatment effect approach

		Costs			QALYs			Deterministic
		PCSK9	SoC	Inc	PCSK9	Soc	Inc	ICER
1	Any Angina, MI, TIA or Stroke (Base-case)	€102,447	€10,167	€92,280	7.30	7.15	0.14	€636,931
2	Any Angina MI TIA, Stroke or other CVD	€104,795	€9,952	€94,843	7.59	7.45	0.14	€692,882
3	Any Angina, MI or other CVD	€105,463	€10,388	€95,075	7.55	7.41	0.14	€655,773
4	Any Angina or MI	€103,310	€10,860	€92,450	7.24	7.08	0.16	€585,271
5	Any TIA or Stroke	€102,960	€12,506	€90,454	7.06	6.90	0.16	€565,992
6	Any Angina	€103,606	€9,942	€93,663	7.10	6.96	0.14	€676,448
7	Only Angina (No MI, TIA or stroke)	€106,844	€8,683	€98,161	7.70	7.58	0.12	€787,385
8	Any MI	€106,453	€14,405	€92,049	7.41	7.20	0.20	€456,061
9	Only MI (No Angina, TIA, or stroke)	€108,081	€13,764	€94,316	7.95	7.76	0.19	€484,367
10	>1 MI	€118,176	€20,622	€97,555	7.20	6.98	0.23	€433,103
11	Any TIA	€100,563	€12,247	€88,316	7.14	6.98	0.16	€539,762
12	Only TIA (other heart excluded)	€99,118	€10,733	€88,385	7.82	7.68	0.14	€620,274
13	Any Stroke	€101,895	€12,278	€89,616	6.53	6.39	0.14	€620,606
14	Only Stroke (No angina, MI, TIA, or other CVD)	€100,362	€10,786	€89,576	6.92	6.79	0.13	€692,569
15	>1 Stroke	€101,468	€14,164	€87,304	4.79	4.65	0.15	€590,933
16	Other CVD	€118,124	€8,648	€109,477	9.20	9.09	0.11	€1,038,287
17	Only CVD	€112,337	€9,449	€102,888	8.32	8.20	0.12	€856,076
18	MI or Stroke	€105,594	€12,443	€93,150	7.35	7.18	0.17	€540,785
19	>1 MI or > 1 Stroke	€118,385	€20,343	€98,041	7.48	7.25	0.23	€427,773
20	Disease in two vascular beds i.e. (MI or Angina) and (TIA or stroke)	€99,683	€16,858	€82,824	5.30	5.10	0.20	€407,032
21	As population 1 but excluding Population 20	€104,873	€9,985	€94,887	7.64	7.50	0.14	€686,448
22	Diabetes and CVD (Angina, MI, TIA or Stroke)	€110,183	€13,563	€96,620	5.61	5.47	0.14	€690,170
23	Diabetes and MI	€110,254	€21,094	€89,160	6.72	6.47	0.25	€352,041

CAD, coronary artery disease; CVD, Cardiovascular disease, ICER incremental cost-effectiveness ratio; Inc, Incremental; MI myocardial infarction; PAD, peripheral artery disease, QALYs, Quality Adjusted Life Year; SoC, standard of care; TIA, Transient ischemic attack, TILDA, The Irish Longitudinal Study on Aging.

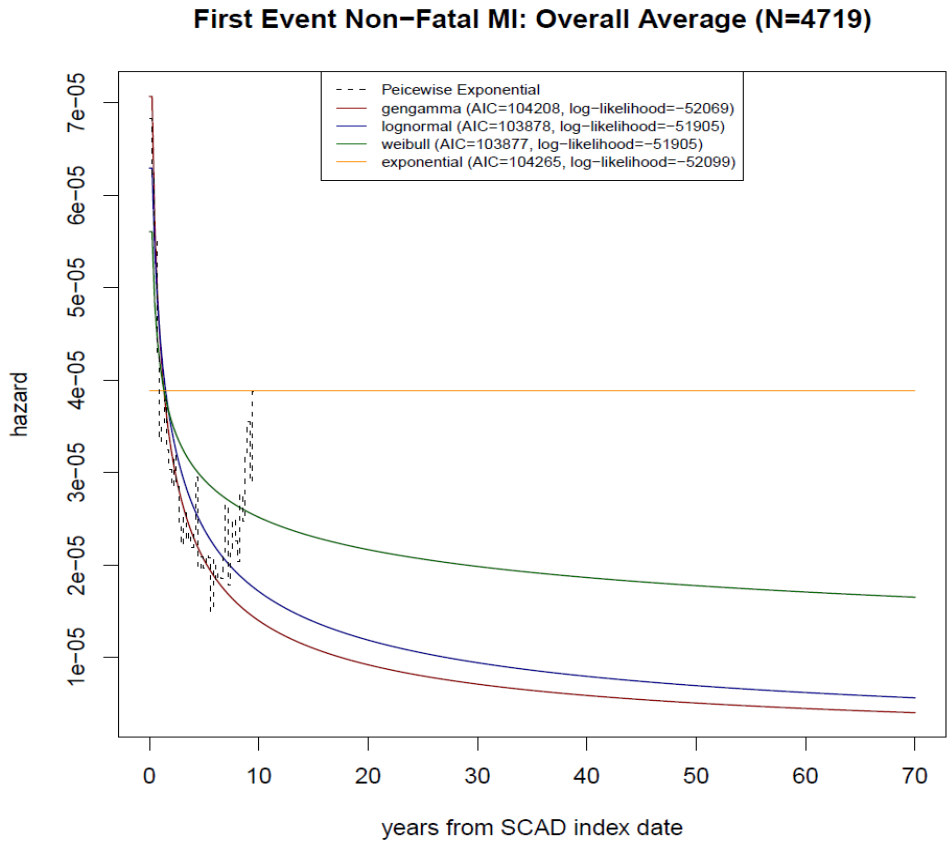
Table A.18 Total and incremental costs and outcomes for PCSK9 inhibitors versus standard of care for TILDA population 1 by age and gender assuming a direct treatment effect

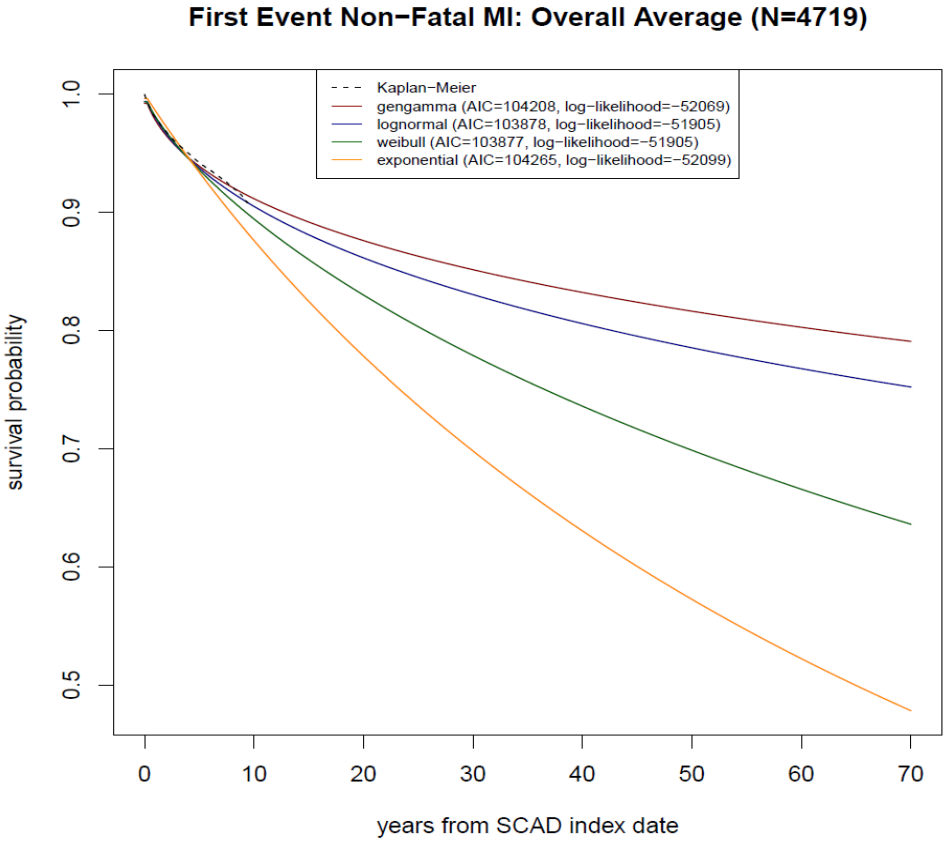
	Male							Female						
	Costs			QALYs			ICER	Costs			QALYs			ICER
	PCSK9i	SoC	Incremental	PCSK9i	Soc	Incremental		PCSK9i	SoC	Incremental	PCSK9	Soc	Incremental	
50	€150,151	€15,430	€134,720	11.67	11.63	0.04	€3,608,581	€161,822	€14,287	€147,535	12.44	12.41	0.04	€4,052,380
60	€122,693	€13,477	€109,216	9.46	9.42	0.05	€2,287,186	€135,708	€13,358	€122,350	10.31	10.27	0.04	€2,920,285
70	€93,966	€10,654	€83,312	7.13	7.07	0.06	€1,469,924	€104,604	€10,756	€93,848	8.05	8.00	0.05	€1,932,343
80	€74,902	€8,585	€66,317	5.50	5.41	0.09	€747,504	€84,344	€8,677	€75,667	6.33	6.25	0.08	€947,621

ICER, Incremental Cost Effectiveness Ratio; PCSK9i Proprotein Convertase Subtilisin/Kexin Type 9 inhibitor; QALY, Quality Adjusted Life Year; SOC, Standard of Care; TILDA, The Irish Longitudinal Study on Aging.

1.4.4 Risk Equations and Parametric Survival Distributions as Estimated by Asaria et al

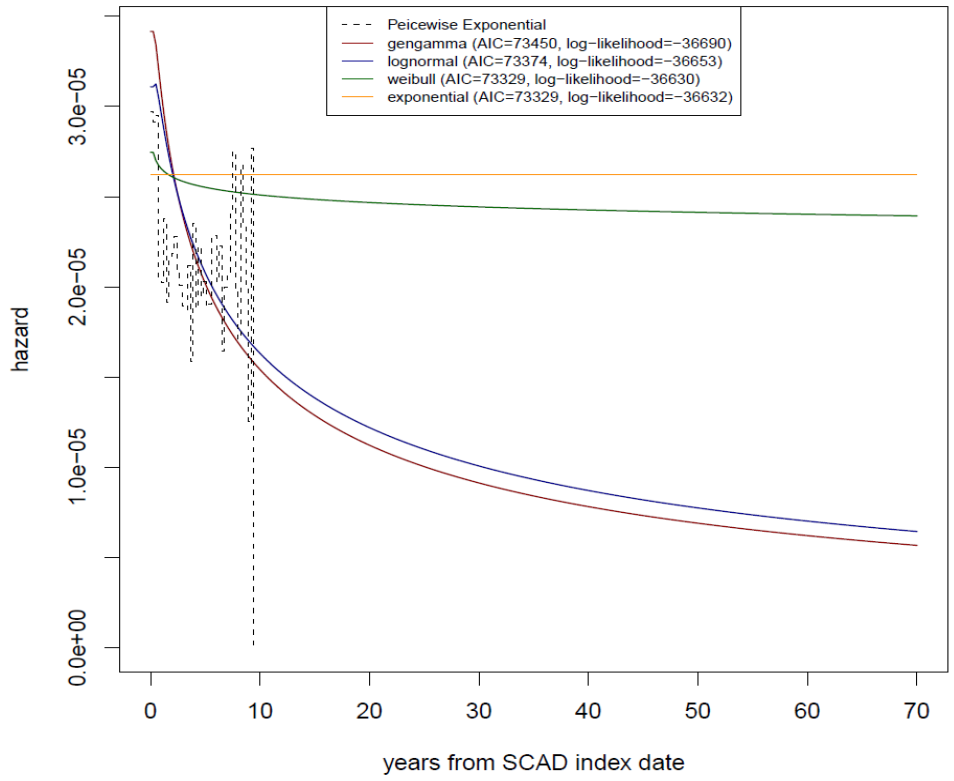
Equation 1: FE MI	GenGamma	LogNormal	Weibull	Exponential
Sociodemographic characteristics				
Age in men	0.98 (0.96-0.99)	0.98 (0.98-0.99)	0.98 (0.98-0.98)	0.98 (0.98-0.99)
Age in women	0.98 (0.97-1.00)	0.98 (0.98-0.99)	0.98 (0.98-0.99)	0.99 (0.98-0.99)
Women vs men	1.59 (1.19-2.14)	1.49 (1.32-1.68)	1.44 (1.29-1.59)	1.35 (1.24-1.46)
Most deprived quintile, yes vs. no	0.89 (0.62-1.27)	0.81 (0.73-0.89)	0.85 (0.78-0.93)	0.88 (0.82-0.94)
SCAD diagnosis and severity				
Other CHD vs. stable angina	0.89 (0.69-1.13)	0.85 (0.75-0.96)	0.79 (0.71-0.89)	0.84 (0.77-0.92)
NSTEMI vs. stable angina	0.19 (0.14-0.26)	0.19 (0.17-0.22)	0.23 (0.20-0.26)	0.31 (0.28-0.34)
STEMI vs stable angina	0.26 (0.19-0.36)	0.26 (0.22-0.32)	0.29 (0.25-0.35)	0.37 (0.33-0.43)
Unstable angina vs. stable angina	0.58 (0.47-0.71)	0.57 (0.50-0.65)	0.56 (0.50-0.63)	0.64 (0.58-0.70)
PCI in last 6 months	1.20 (0.77-1.87)	1.13 (0.97-1.32)	1.11 (0.97-1.27)	1.05 (0.95-1.17)
CABG in last 6 months	3.81 (1.90-7.62)	3.05 (2.39-3.91)	2.88 (2.28-3.65)	2.37 (1.97-2.85)
Previous/recurrent MI	0.53 (0.46-0.62)	0.57 (0.51-0.63)	0.62 (0.56-0.68)	0.69 (0.64-0.74)
Use of nitrates	0.62 (0.52-0.73)	0.64 (0.59-0.71)	0.69 (0.64-0.75)	0.75 (0.70-0.80)
CVD risk factors				
Current smoker vs. never	0.85 (0.65-1.10)	0.80 (0.69-0.92)	0.85 (0.75-0.97)	0.91 (0.83-1.01)
Ex-smoker vs. never	0.91 (0.68-1.21)	0.91 (0.79-1.05)	0.92 (0.81-1.05)	0.94 (0.85-1.04)
Hypertension	1.34 (0.85-2.10)	1.19 (1.07-1.32)	1.15 (1.05-1.27)	1.12 (1.04-1.21)
Diabetes mellitus	0.63 (0.44-0.91)	0.60 (0.54-0.67)	0.63 (0.57-0.70)	0.69 (0.64-0.75)
Total cholesterol, per 1 mmol/L increase	0.88 (0.76-1.03)	0.91 (0.86-0.96)	0.92 (0.87-0.97)	0.94 (0.90-0.98)
HDL, per 0.5 mmol/L increase	1.20 (0.98-1.47)	1.11 (1.04-1.18)	1.10 (1.04-1.17)	1.08 (1.03-1.13)
CVD co-morbidities				
Heart failure	0.86 (0.70-1.05)	0.86 (0.77-0.95)	0.86 (0.79-0.95)	0.88 (0.82-0.94)
Peripheral arterial disease	0.72 (0.41-1.28)	0.62 (0.54-0.71)	0.64 (0.57-0.71)	0.69 (0.63-0.75)
Atrial fibrillation	1.13 (0.84-1.51)	1.03 (0.92-1.17)	1.00 (0.90-1.11)	0.98 (0.90-1.06)
Stroke	0.77 (0.60-0.99)	0.79 (0.69-0.90)	0.80 (0.71-0.90)	0.82 (0.75-0.90)
Non-CVD co-morbidities				
Chornic kidney disease	0.84 (0.41-1.70)	0.97 (0.81-1.16)	0.90 (0.77-1.05)	0.84 (0.74-0.94)
Chronic obstructive pulmonary disease	0.91 (0.73-1.13)	0.85 (0.77-0.94)	0.86 (0.78-0.94)	0.88 (0.82-0.94)
Cancer	0.93 (0.70-1.25)	0.96 (0.83-1.11)	0.96 (0.84-1.09)	0.95 (0.86-1.05)
Chronic liver disease	0.83 (0.26-2.59)	0.77 (0.51-1.15)	0.75 (0.53-1.07)	0.78 (0.59-1.02)
Psychosocial characteristics				
Depression at diagnosis	1.16 (1.00-1.35)	1.15 (1.02-1.29)	1.11 (0.99-1.24)	1.06 (0.97-1.15)
Anxiety at diagnosis	1.11 (0.79-1.58)	1.04 (0.88-1.22)	1.04 (0.90-1.21)	1.02 (0.91-1.15)
Biomarkers				
Heart rate, per 10 b.p.m. increase	1.02 (0.96-1.08)	1.00 (0.95-1.05)	0.99 (0.95-1.04)	0.99 (0.95-1.03)
Creatinine, per 30 micromol/L increase	0.88 (0.77-1.00)	0.89 (0.85-0.94)	0.91 (0.87-0.95)	0.93 (0.90-0.96)
White cell count, per 1.5 10^9/L increase	0.89 (0.82-0.97)	0.89 (0.85-0.93)	0.90 (0.87-0.93)	0.92 (0.89-0.94)
Haemoglobin, per 1.5 g/dL increase	1.21 (1.09-1.34)	1.17 (1.10-1.25)	1.15 (1.08-1.21)	1.12 (1.07-1.17)
Generalised gamma model parameters				
mu	13.24 (12.32-14.16)	12.79 (12.59-12.99)	11.83 (11.65-12.00)	10.90 (10.80-11.00)
sigma	3.67 (2.27-5.95)	2.91 (2.84-2.97)	1.28 (1.24-1.31)	1
Q	-0.23 (-0.77-0.31)	0	1	1
Model Fit				
Log-likelihood	-52068.90	-51904.95	-51904.60	-52099.34
AIC	104207.81	103877.90	103877.20	104264.68


First Event Non−Fatal MI: Overall Average (N=4719)
Peicewise Exponential
gengamma (AIC=104208, log−likelihood=−52069)
lognormal (AIC=103878, log−likelihood=−51905)
weibull (AIC=103877, log−likelihood=−51905)
exponential (AIC=104265, log−likelihood=−52099)
hazard
1e−05
2e−05
3e−05
4e−05
5e−05
6e−05
7e−05
0
10
20
30
40
50
60
70
years from SCAD index date

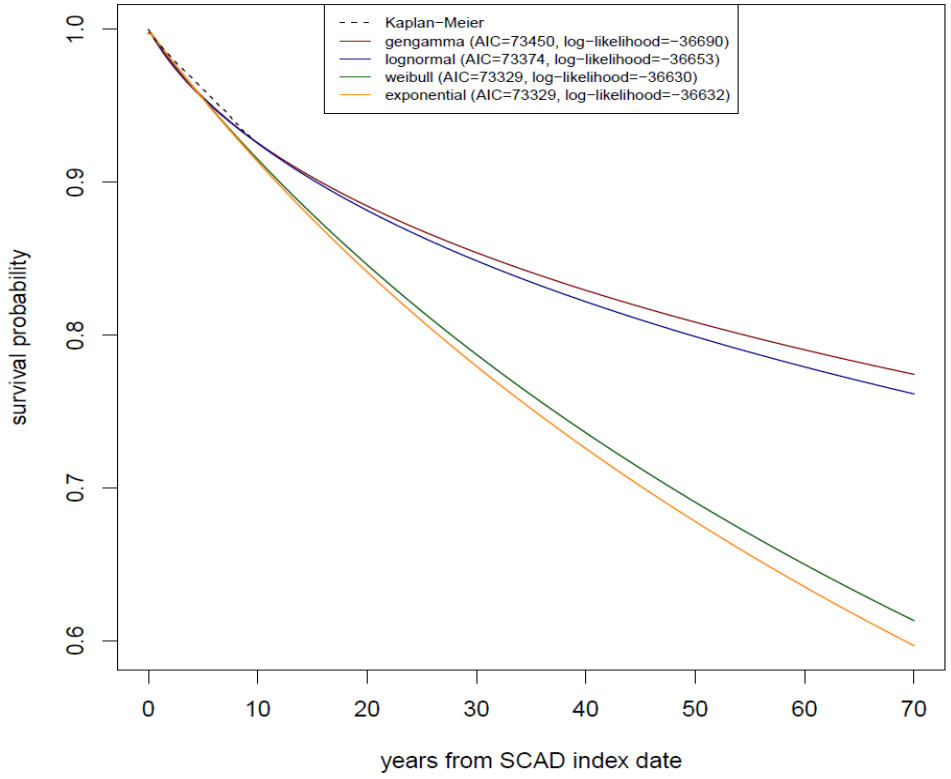

First Event Non−Fatal MI: Overall Average (N=4719)
Kaplan−Meier
gengamma (AIC=104208, log−likelihood=−52069)
lognormal (AIC=103878, log−likelihood=−51905)
weibull (AIC=103877, log−likelihood=−51905)
exponential (AIC=104265, log−likelihood=−52099)
survival probability
0.5
0.6
0.7
0.8
0.9
1.0
0
10
20
30
40
50
60
70
years from SCAD index date

Equation 2: FE Stroke I	GenGamma	LogNormal	Weibull	Exponential
Sociodemographic characteristics				
Age in men	0.95 (0.95-0.96)	0.95 (0.95-0.96)	0.96 (0.95-0.96)	0.96 (0.95-0.96)
Age in women	1.01 (1.00-1.02)	1.01 (1.00-1.01)	1.00 (1.00-1.01)	1.00 (1.00-1.01)
Women vs men	1.12 (0.98-1.28)	1.13 (1.02-1.25)	1.12 (1.02-1.23)	1.12 (1.02-1.23)
Most deprived quintile, yes vs. no	0.77 (0.68-0.87)	0.78 (0.71-0.86)	0.81 (0.74-0.88)	0.81 (0.75-0.88)
SCAD diagnosis and severity				
Other CHD vs. stable angina	0.99 (0.86-1.14)	1.01 (0.90-1.14)	1.00 (0.90-1.11)	1.00 (0.91-1.11)
NSTEMI vs. stable angina	1.00 (0.84-1.18)	0.92 (0.78-1.08)	0.93 (0.81-1.08)	0.93 (0.81-1.07)
STEMI vs stable angina	1.22 (0.68-2.18)	1.06 (0.84-1.34)	1.05 (0.84-1.31)	1.04 (0.84-1.30)
Unstable angina vs. stable angina	0.91 (0.75-1.11)	0.88 (0.77-0.99)	0.88 (0.79-0.98)	0.88 (0.79-0.98)
PCI in last 6 months	1.09 (0.80-1.49)	1.13 (0.94-1.37)	1.14 (0.95-1.36)	1.13 (0.95-1.35)
CABG in last 6 months	1.21 (0.96-1.52)	1.19 (0.95-1.48)	1.15 (0.94-1.41)	1.15 (0.95-1.40)
Previous/recurrent MI	0.87 (0.69-1.08)	0.88 (0.78-0.99)	0.90 (0.81-1.00)	0.90 (0.82-1.00)
Use of nitrates	0.97 (0.84-1.14)	0.96 (0.88-1.05)	0.97 (0.89-1.05)	0.97 (0.89-1.04)
CVD risk factors				
Current smoker vs. never	0.74 (0.59-0.92)	0.74 (0.65-0.84)	0.79 (0.71-0.88)	0.80 (0.72-0.89)
Ex-smoker vs. never	0.99 (0.81-1.21)	1.01 (0.89-1.14)	1.01 (0.90-1.14)	1.01 (0.91-1.13)
Hypertension	1.02 (0.85-1.22)	1.04 (0.93-1.15)	1.02 (0.93-1.13)	1.02 (0.93-1.12)
Diabetes mellitus	0.69 (0.60-0.80)	0.72 (0.64-0.80)	0.74 (0.67-0.82)	0.75 (0.68-0.82)
Total cholesterol, per 1 mmol/L increase	0.94 (0.86-1.02)	0.93 (0.89-0.99)	0.95 (0.90-1.00)	0.95 (0.91-1.00)
HDL, per 0.5 mmol/L increase	1.00 (0.90-1.11)	1.00 (0.91-1.10)	0.99 (0.91-1.07)	0.99 (0.91-1.07)
CVD co-morbidities				
Heart failure	0.85 (0.75-0.96)	0.86 (0.78-0.95)	0.90 (0.83-0.98)	0.90 (0.83-0.98)
Peripheral arterial disease	0.80 (0.67-0.96)	0.84 (0.73-0.96)	0.87 (0.78-0.98)	0.87 (0.78-0.98)
Atrial fibrillation	0.59 (0.51-0.68)	0.62 (0.56-0.69)	0.66 (0.60-0.72)	0.67 (0.61-0.73)
Stroke	0.23 (0.19-0.28)	0.22 (0.20-0.25)	0.30 (0.27-0.33)	0.31 (0.28-0.33)
Non-CVD co-morbidities				
Chornic kidney disease	1.23 (0.96-1.57)	1.11 (0.92-1.34)	1.05 (0.88-1.25)	1.03 (0.87-1.23)
Chronic obstructive pulmonary disease	1.07 (0.96-1.20)	1.08 (0.98-1.19)	1.07 (0.98-1.16)	1.06 (0.98-1.16)
Cancer	1.08 (0.89-1.31)	1.03 (0.90-1.19)	1.04 (0.92-1.17)	1.03 (0.92-1.17)
Chronic liver disease	0.72 (0.41-1.24)	0.79 (0.52-1.19)	0.78 (0.55-1.12)	0.79 (0.55-1.12)
Psychosocial characteristics				
Depression at diagnosis	0.90 (0.76-1.06)	0.90 (0.81-1.01)	0.89 (0.81-0.98)	0.89 (0.81-0.98)
Anxiety at diagnosis	0.96 (0.80-1.16)	0.94 (0.81-1.09)	0.94 (0.82-1.07)	0.94 (0.83-1.07)
Biomarkers				
Heart rate, per 10 b.p.m. increase	1.00 (0.93-1.07)	0.99 (0.94-1.04)	0.99 (0.95-1.03)	0.99 (0.95-1.03)
Creatinine, per 30 micromol/L increase	0.97 (0.90-1.03)	0.96 (0.90-1.02)	0.97 (0.92-1.02)	0.97 (0.92-1.02)
White cell count, per 1.5 10^9/L increase	0.92 (0.88-0.97)	0.93 (0.90-0.97)	0.94 (0.91-0.97)	0.94 (0.91-0.97)
Haemoglobin, per 1.5 g/dL increase	1.04 (0.99-1.09)	1.04 (0.98-1.09)	1.03 (0.99-1.08)	1.03 (0.99-1.08)
Generalised gamma model parameters				
mu	12.56 (12.32-12.8)	12.37 (12.17-12.57)	11.28 (11.11-11.45)	11.19 (11.07-11.31)
sigma	2.72 (2.29-3.23)	2.47 (2.40-2.54)	1.02 (0.99-1.06)	1
Q	-0.09 (-0.36-0.17)	0	1	1
Model Fit				
Log-likelihood	-36689.77	-36652.82	-36630.25	-36631.56
AIC	73449.54	73373.64	73328.51	73329.11

First Event Non−Fatal Ischaemic Stroke: Overall Average (N=3222)

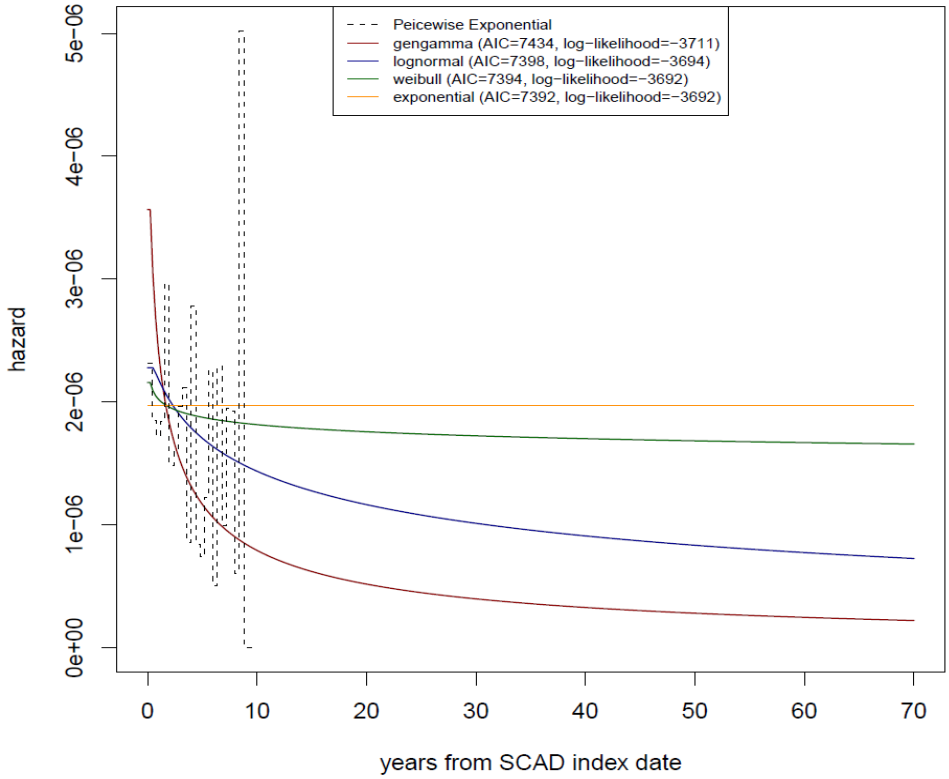


First Event Non−Fatal Ischaemic Stroke: Overall Average (N=3222)

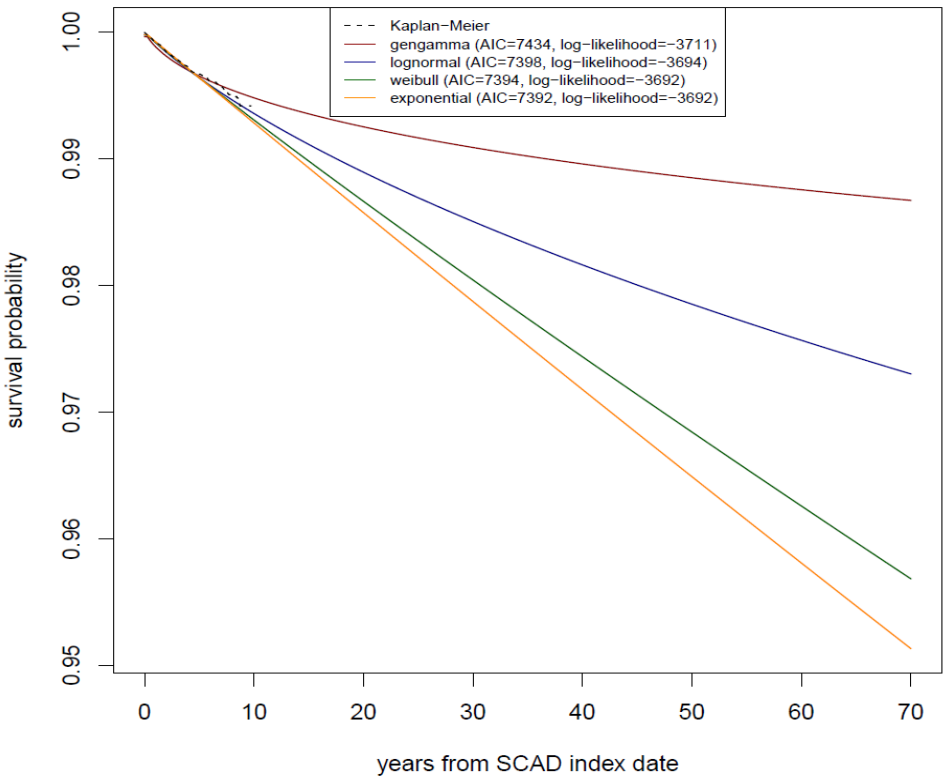


Equation 3: FE Stroke H	GenGamma	LogNormal	Weibull	Exponential
Sociodemographic characteristics				
Age in men	0.94 (0.92-0.96)	0.95 (0.93-0.97)	0.95 (0.94-0.97)	0.95 (0.94-0.97)
Age in women	1.02 (0.98-1.05)	1.01 (0.98-1.03)	1.01 (0.99-1.03)	1.01 (0.99-1.03)
Women vs men	1.83 (1.24-2.70)	1.49 (1.11-2.00)	1.41 (1.07-1.86)	1.39 (1.07-1.81)
Most deprived quintile, yes vs. no				
SCAD diagnosis and severity				
Other CHD vs. stable angina				
NSTEMI vs. stable angina				
STEMI vs stable angina				
Unstable angina vs. stable angina				
PCI in last 6 months				
CABG in last 6 months				
Previous/recurrent MI				
Use of nitrates				
CVD risk factors				
Current smoker vs. never				
Ex-smoker vs. never				
Hypertension				
Diabetes mellitus				
Total cholesterol, per 1 mmol/L increase				
HDL, per 0.5 mmol/L increase				
CVD co-morbidities				
Heart failure				
Peripheral arterial disease				
Atrial fibrillation				
Stroke				
Non-CVD co-morbidities				
Chornic kidney disease				
Chronic obstructive pulmonary disease				
Cancer				
Chronic liver disease				
Psychosocial characteristics				
Depression at diagnosis				
Anxiety at diagnosis				
Biomarkers				
Heart rate, per 10 b.p.m. increase				
Creatinine, per 30 micromol/L increase				
White cell count, per 1.5 10^9/L increase				
Haemoglobin, per 1.5 g/dL increase				
Generalised gamma model parameters				
mu	23.07 (21.60-24.54)	16.59 (15.65-17.53)	13.36 (12.73-13.99)	13.09 (12.93-13.25)
sigma	14.52 (12.56-16.78)	3.41 (3.09-3.77)	1.05 (0.94-1.17)	1
Q	-2.76 (-3.32--2.19)	0	1	1
Model Fit				
Log-likelihood	-3711.20	-3694.11	-3691.75	-3692.16
AIC	7434.41	7398.23	7393.50	7392.32

First Event Non−Fatal Hemorrhagic Stroke: Overall Average (N=262)

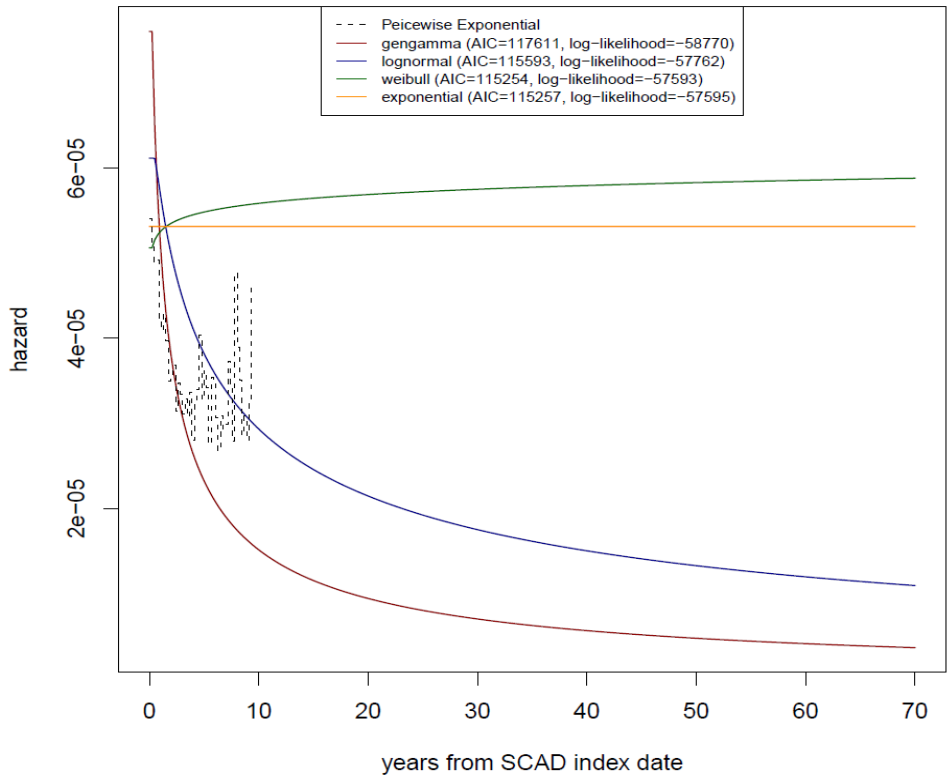


First Event Non−Fatal Hemorrhagic Stroke: Overall Average (N=262)

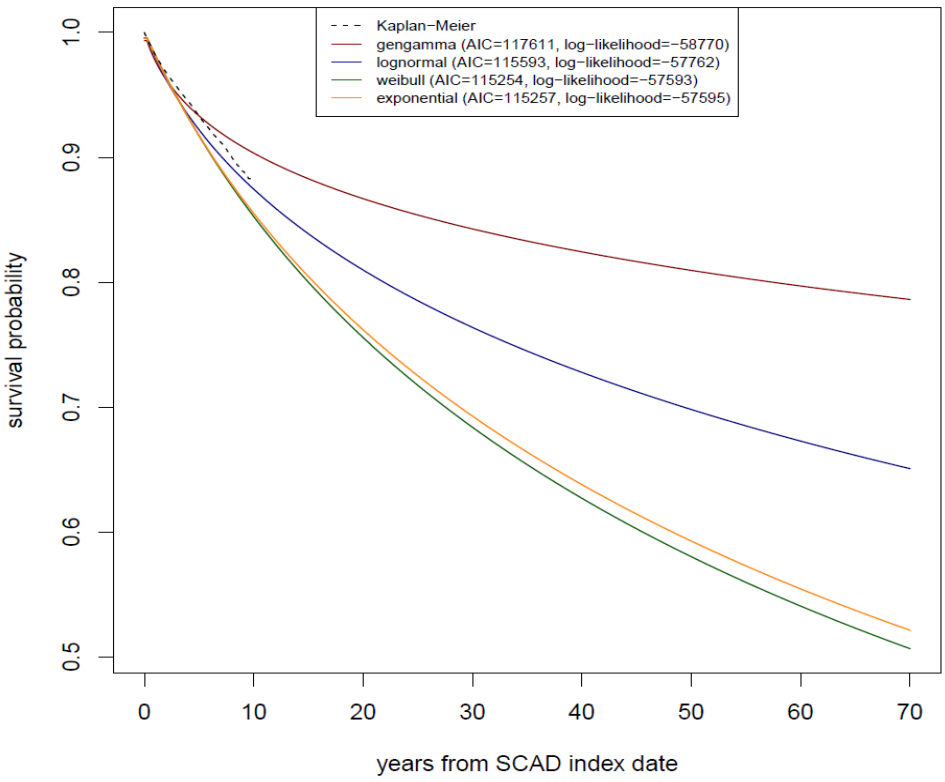


Equation 4: FE Fatal CVD	GenGamma	LogNormal	Weibull	Exponential
Sociodemographic characteristics				
Age in men	0.94 (0.92-0.96)	0.94 (0.93-0.94)	0.94 (0.94-0.94)	0.94 (0.94-0.94)
Age in women	0.97 (0.95-1.00)	0.97 (0.97-0.98)	0.97 (0.97-0.98)	0.97 (0.97-0.98)
Women vs men	1.82 (0.95-3.46)	2.04 (1.86-2.24)	1.97 (1.81-2.16)	2.00 (1.83-2.19)
Most deprived quintile, yes vs. no	0.90 (0.51-1.58)	0.85 (0.79-0.92)	0.90 (0.84-0.96)	0.90 (0.84-0.96)
SCAD diagnosis and severity				
Other CHD vs. stable angina	0.84 (0.57-1.24)	0.85 (0.78-0.93)	0.85 (0.79-0.92)	0.85 (0.78-0.91)
NSTEMI vs. stable angina	0.54 (0.33-0.88)	0.54 (0.48-0.60)	0.57 (0.53-0.63)	0.57 (0.52-0.62)
STEMI vs stable angina	0.73 (0.29-1.84)	0.74 (0.62-0.87)	0.77 (0.65-0.90)	0.77 (0.65-0.90)
Unstable angina vs. stable angina	0.91 (0.61-1.34)	0.90 (0.82-1.00)	0.89 (0.81-0.97)	0.89 (0.81-0.97)
PCI in last 6 months	1.42 (0.56-3.56)	1.71 (1.46-2.00)	1.82 (1.55-2.13)	1.85 (1.58-2.18)
CABG in last 6 months	1.58 (0.46-5.46)	2.09 (1.73-2.51)	1.98 (1.65-2.36)	2.00 (1.67-2.41)
Previous/recurrent MI	0.68 (0.52-0.90)	0.72 (0.66-0.78)	0.76 (0.72-0.82)	0.76 (0.71-0.81)
Use of nitrates	0.71 (0.59-0.85)	0.70 (0.65-0.74)	0.75 (0.71-0.79)	0.74 (0.70-0.79)
CVD risk factors				
Current smoker vs. never	0.60 (0.34-1.04)	0.76 (0.68-0.84)	0.80 (0.73-0.87)	0.79 (0.72-0.86)
Ex-smoker vs. never	0.77 (0.51-1.17)	0.95 (0.86-1.05)	0.96 (0.87-1.05)	0.96 (0.87-1.05)
Hypertension	0.93 (0.71-1.22)	0.98 (0.90-1.06)	0.98 (0.91-1.06)	0.98 (0.91-1.06)
Diabetes mellitus	0.79 (0.57-1.11)	0.73 (0.68-0.80)	0.75 (0.70-0.80)	0.75 (0.70-0.80)
Total cholesterol, per 1 mmol/L increase	0.97 (0.84-1.13)	0.96 (0.91-1.00)	0.97 (0.93-1.01)	0.97 (0.92-1.01)
HDL, per 0.5 mmol/L increase	1.05 (0.81-1.36)	1.05 (0.99-1.12)	1.03 (0.98-1.08)	1.03 (0.98-1.09)
CVD co-morbidities				
Heart failure	0.47 (0.38-0.59)	0.52 (0.48-0.55)	0.58 (0.54-0.61)	0.57 (0.53-0.60)
Peripheral arterial disease	0.70 (0.44-1.12)	0.72 (0.66-0.79)	0.75 (0.70-0.81)	0.75 (0.69-0.81)
Atrial fibrillation	0.69 (0.59-0.82)	0.73 (0.68-0.79)	0.76 (0.72-0.81)	0.76 (0.71-0.81)
Stroke	0.63 (0.35-1.14)	0.66 (0.60-0.71)	0.72 (0.67-0.77)	0.71 (0.66-0.77)
Non-CVD co-morbidities				
Chornic kidney disease	1.04 (0.33-3.33)	0.98 (0.86-1.11)	0.94 (0.84-1.04)	0.95 (0.85-1.05)
Chronic obstructive pulmonary disease	1.22 (0.78-1.91)	1.05 (0.98-1.13)	1.03 (0.97-1.10)	1.04 (0.97-1.10)
Cancer	1.29 (0.83-2.00)	1.10 (1.00-1.22)	1.12 (1.03-1.22)	1.12 (1.03-1.23)
Chronic liver disease	0.44 (0.09-2.09)	0.64 (0.48-0.85)	0.76 (0.59-0.99)	0.76 (0.59-0.99)
Psychosocial characteristics				
Depression at diagnosis	0.89 (0.58-1.37)	0.90 (0.83-0.98)	0.89 (0.83-0.96)	0.89 (0.83-0.96)
Anxiety at diagnosis	0.71 (0.39-1.28)	0.85 (0.77-0.95)	0.88 (0.80-0.97)	0.88 (0.80-0.97)
Biomarkers				
Heart rate, per 10 b.p.m. increase	0.89 (0.81-0.98)	0.90 (0.87-0.93)	0.92 (0.89-0.95)	0.92 (0.89-0.94)
Creatinine, per 30 micromol/L increase	0.89 (0.81-0.97)	0.89 (0.86-0.92)	0.91 (0.89-0.93)	0.90 (0.88-0.93)
White cell count, per 1.5 10^9/L increase	0.85 (0.77-0.95)	0.89 (0.86-0.93)	0.91 (0.88-0.94)	0.91 (0.88-0.94)
Haemoglobin, per 1.5 g/dL increase	1.30 (1.12-1.50)	1.28 (1.23-1.32)	1.23 (1.19-1.26)	1.23 (1.19-1.27)
Generalised gamma model parameters				
mu	12.33 (11.37-13.29)	11.49 (11.35-11.63)	10.9 (10.77-11.02)	10.98 (10.88-11.09)
sigma	4.24 (3.86-4.66)	2.09 (2.05-2.14)	0.97 (0.95-1.00)	1
Q	-1.30 (-1.78--0.81)	0	1	1
Model Fit				
Log-likelihood	-58770.27	-57762.41	-57592.89	-57595.48
AIC	117610.54	115592.82	115253.77	115256.97

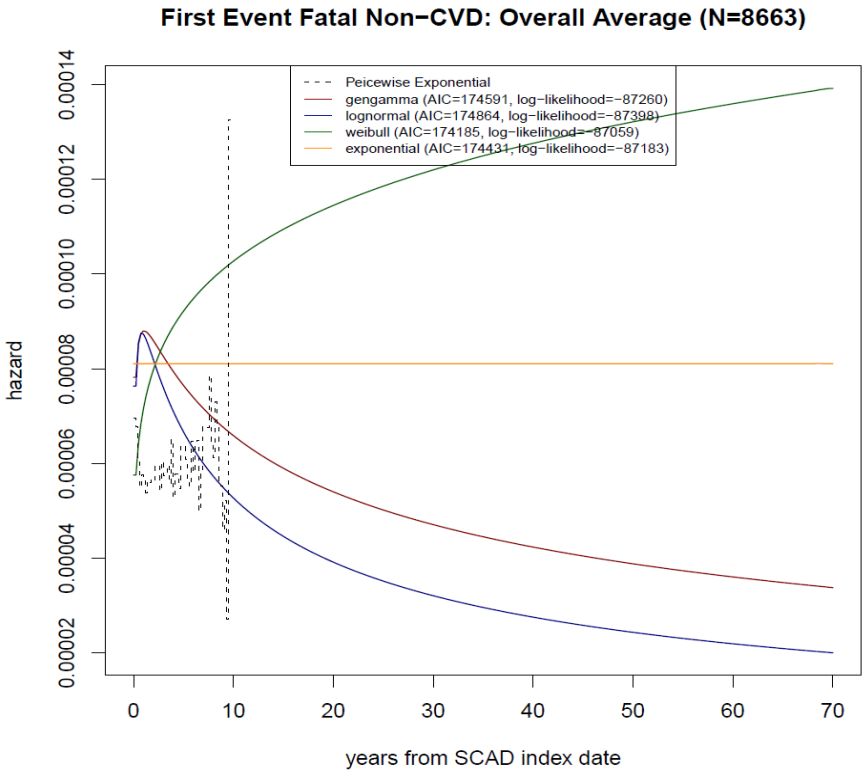
First Event Fatal CVD: Overall Average (N=5536)

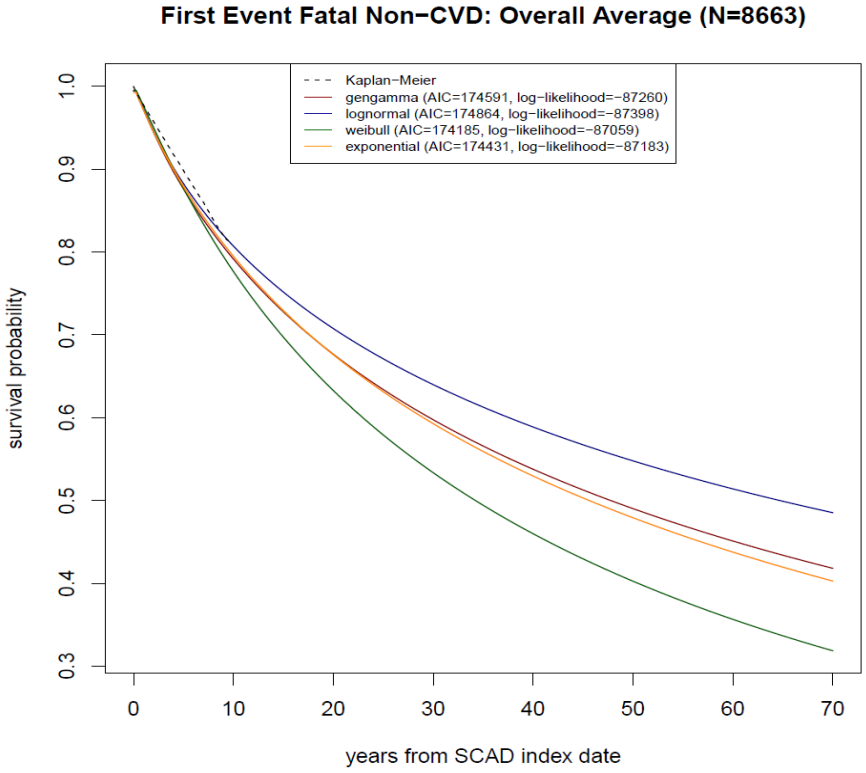


First Event Fatal CVD: Overall Average (N=5536)



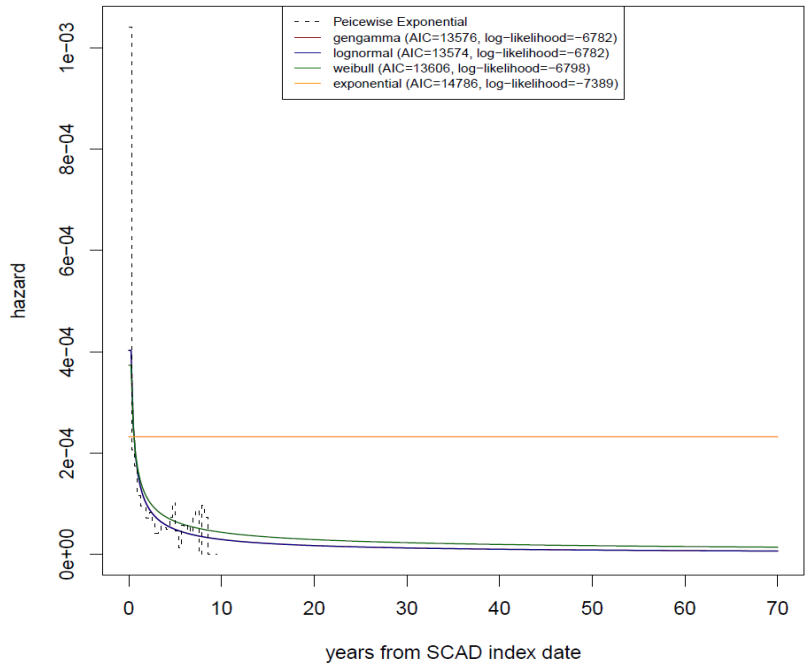
Equation 5: FE Fatal non-CVD	GenGamma	LogNormal	Weibull	Exponential
Sociodemographic characteristics				
Age in men	0.94 (0.94-0.95)	0.94 (0.94-0.95)	0.94 (0.94-0.95)	0.94 (0.93-0.94)
Age in women	0.99 (0.98-0.99)	0.99 (0.98-0.99)	0.99 (0.98-0.99)	0.99 (0.98-0.99)
Women vs men	1.71 (1.60-1.84)	1.75 (1.65-1.86)	1.65 (1.56-1.74)	1.75 (1.64-1.87)
Most deprived quintile, yes vs. no	0.90 (0.74-1.08)	0.84 (0.79-0.88)	0.86 (0.82-0.90)	0.85 (0.80-0.89)
SCAD diagnosis and severity				
Other CHD vs. stable angina	1.03 (0.90-1.18)	1.04 (0.98-1.11)	1.02 (0.97-1.08)	1.02 (0.96-1.08)
NSTEMI vs. stable angina	0.91 (0.70-1.19)	0.89 (0.82-0.96)	0.90 (0.84-0.96)	0.90 (0.84-0.98)
STEMI vs stable angina	1.05 (0.62-1.77)	0.93 (0.83-1.04)	0.90 (0.81-1.00)	0.91 (0.80-1.02)
Unstable angina vs. stable angina	1.07 (0.91-1.26)	1.03 (0.96-1.10)	1.02 (0.96-1.09)	1.02 (0.95-1.10)
PCI in last 6 months	1.31 (1.07-1.59)	1.34 (1.21-1.50)	1.32 (1.19-1.47)	1.40 (1.24-1.58)
CABG in last 6 months	1.75 (1.23-2.50)	1.85 (1.62-2.12)	1.74 (1.53-1.99)	1.86 (1.60-2.17)
Previous/recurrent MI	1.02 (0.90-1.16)	1.01 (0.95-1.08)	1.00 (0.95-1.06)	1.00 (0.94-1.07)
Use of nitrates	0.90 (0.85-0.96)	0.89 (0.85-0.93)	0.92 (0.88-0.96)	0.91 (0.87-0.95)
CVD risk factors				
Current smoker vs. never	0.73 (0.65-0.81)	0.70 (0.65-0.76)	0.74 (0.69-0.80)	0.69 (0.64-0.75)
Ex-smoker vs. never	0.86 (0.79-0.94)	0.84 (0.78-0.91)	0.85 (0.80-0.91)	0.83 (0.77-0.89)
Hypertension	1.15 (1.08-1.22)	1.11 (1.05-1.18)	1.11 (1.06-1.17)	1.13 (1.07-1.19)
Diabetes mellitus	0.90 (0.81-1.00)	0.91 (0.86-0.97)	0.91 (0.86-0.95)	0.90 (0.85-0.95)
Total cholesterol, per 1 mmol/L increase	1.00 (0.96-1.04)	1.01 (0.98-1.04)	1.01 (0.98-1.04)	1.00 (0.97-1.04)
HDL, per 0.5 mmol/L increase	0.97 (0.94-1.00)	0.98 (0.95-1.01)	0.98 (0.95-1.01)	0.98 (0.95-1.01)
CVD co-morbidities				
Heart failure	0.72 (0.68-0.77)	0.71 (0.67-0.74)	0.76 (0.73-0.80)	0.74 (0.70-0.77)
Peripheral arterial disease	0.84 (0.76-0.93)	0.81 (0.75-0.86)	0.83 (0.78-0.87)	0.81 (0.76-0.87)
Atrial fibrillation	0.88 (0.75-1.03)	0.84 (0.80-0.90)	0.88 (0.84-0.93)	0.88 (0.83-0.93)
Stroke	0.86 (0.75-0.98)	0.84 (0.79-0.90)	0.87 (0.82-0.92)	0.86 (0.81-0.92)
Non-CVD co-morbidities				
Chornic kidney disease	0.92 (0.81-1.05)	0.91 (0.83-0.99)	0.89 (0.82-0.96)	0.93 (0.85-1.02)
Chronic obstructive pulmonary disease	0.74 (0.63-0.86)	0.73 (0.70-0.77)	0.76 (0.73-0.79)	0.74 (0.70-0.77)
Cancer	0.49 (0.42-0.57)	0.41 (0.39-0.44)	0.56 (0.53-0.58)	0.51 (0.49-0.54)
Chronic liver disease	0.44 (0.27-0.72)	0.43 (0.36-0.52)	0.53 (0.46-0.62)	0.50 (0.42-0.59)
Psychosocial characteristics				
Depression at diagnosis	0.81 (0.73-0.89)	0.80 (0.75-0.84)	0.82 (0.78-0.86)	0.80 (0.76-0.85)
Anxiety at diagnosis	0.83 (0.55-1.25)	0.78 (0.72-0.84)	0.83 (0.78-0.89)	0.82 (0.76-0.88)
Biomarkers				
Heart rate, per 10 b.p.m. increase	0.90 (0.86-0.95)	0.89 (0.87-0.92)	0.91 (0.89-0.93)	0.90 (0.88-0.93)
Creatinine, per 30 micromol/L increase	0.98 (0.93-1.04)	1.00 (0.98-1.02)	0.99 (0.97-1.01)	0.99 (0.97-1.01)
White cell count, per 1.5 10^9/L increase	0.87 (0.82-0.92)	0.85 (0.84-0.87)	0.89 (0.87-0.90)	0.87 (0.86-0.89)
Haemoglobin, per 1.5 g/dL increase	1.38 (1.33-1.42)	1.41 (1.37-1.44)	1.33 (1.30-1.36)	1.38 (1.35-1.41)
Generalised gamma model parameters				
mu	10.1 (10.01-10.19)	10.25 (10.15-10.34)	9.95 (9.87-10.03)	10.32 (10.24-10.40)
sigma	1.37 (1.11-1.70)	1.73 (1.7-1.76)	0.86 (0.85-0.88)	1
Q	0.46 (0.19-0.73)	0	1	1
Model Fit				
Log-likelihood	-87260.46	-87397.88	-87058.62	-87182.58
AIC	174590.93	174863.76	174185.23	174431.16


First Event Fatal Non−CVD: Overall Average (N=8663)
Peicewise Exponential
gengamma (AIC=174591, log−likelihood=−87260)
lognormal (AIC=174864, log−likelihood=−87398)
weibull (AIC=174185, log−likelihood=−87059)
exponential (AIC=174431, log−likelihood=−87183)
hazard
0.00002
0.00004
0.00006
0.00008
0.00010
0.00012
0.00014
0
10
20
30
40
50
60
70
years from SCAD index date

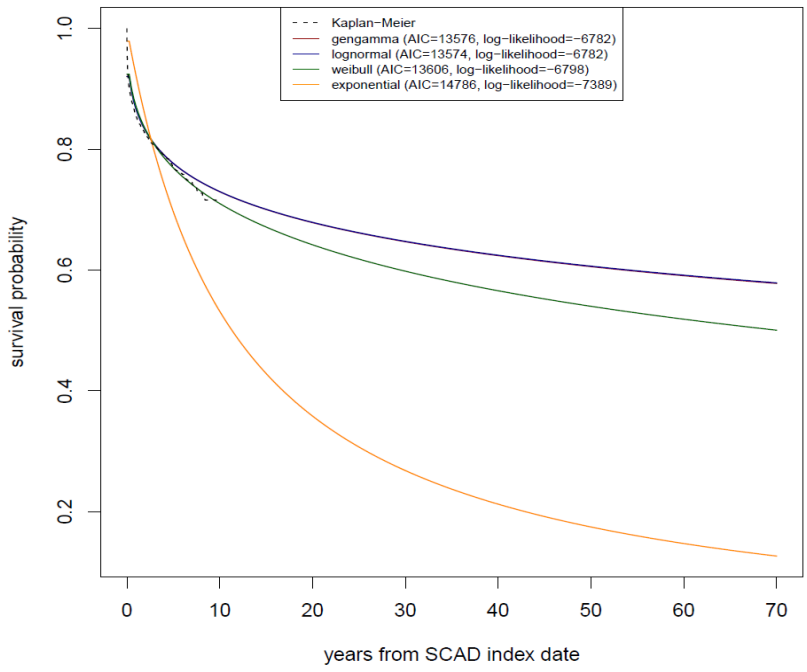

First Event Fatal Non−CVD: Overall Average (N=8663)
Kaplan−Meier
gengamma (AIC=174591, log−likelihood=−87260)
lognormal (AIC=174864, log−likelihood=−87398)
weibull (AIC=174185, log−likelihood=−87059)
exponential (AIC=174431, log−likelihood=−87183)
survival probability
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0
10
20
30
40
50
60
70
years from SCAD index date

Equation 6: Post MI Fatal CVD	GenGamma	LogNormal	Weibull	Exponential
Sociodemographic characteristics				
Age in men	0.85 (0.83-0.87)	0.85 (0.83-0.87)	0.85 (0.83-0.87)	0.92 (0.91-0.93)
Age in women	0.98 (0.95-1.02)	0.98 (0.95-1.02)	0.98 (0.95-1.02)	0.99 (0.97-1.00)
Women vs men	1.87 (1.12-3.11)	1.87 (1.12-3.11)	1.90 (1.12-3.22)	1.39 (1.11-1.74)
Generalised gamma model parameters				
mu	11.14 (10.71-11.56)	11.13 (10.74-11.51)	11.27 (10.89-11.65)	8.85 (8.74-8.96)
sigma	4.21 (3.60-4.93)	4.24 (4.02-4.48)	2.39 (2.25-2.54)	1
Q	0.02 (-0.32-0.36)	0	1	1
Model Fit				
Log-likelihood	-6781.95	-6781.95	-6798.12	-7388.97
AIC	13575.89	13573.90	13606.23	14785.94

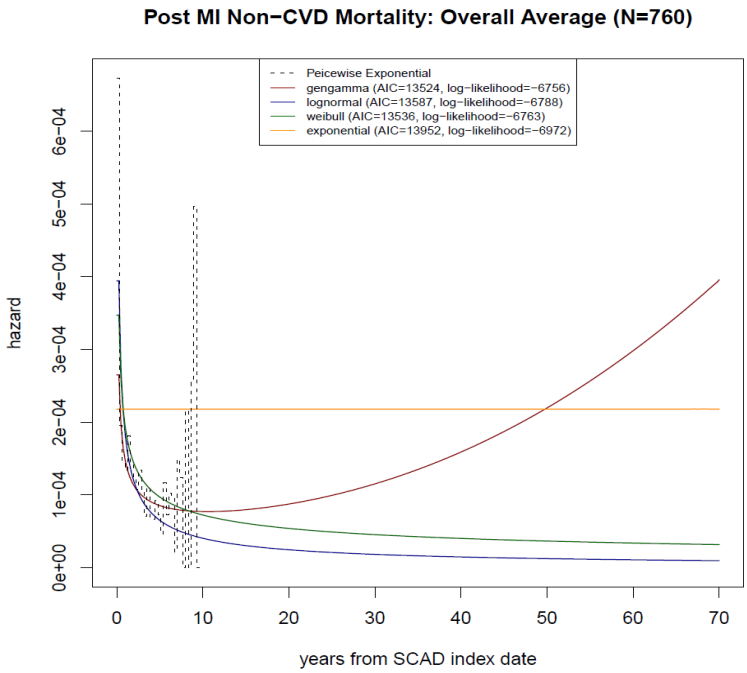
Post MI CVD Mortality: Overall Average (N=813)

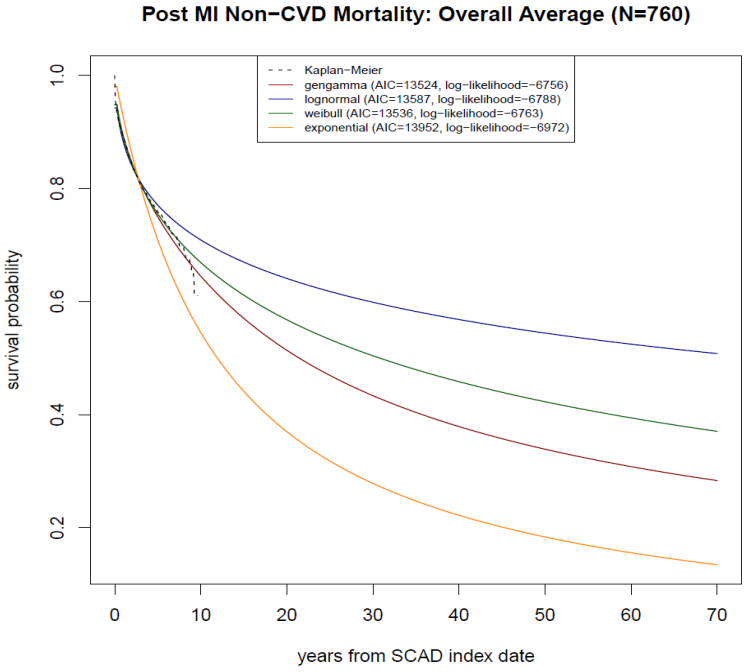


Post MI CVD Mortality: Overall Average (N=813)

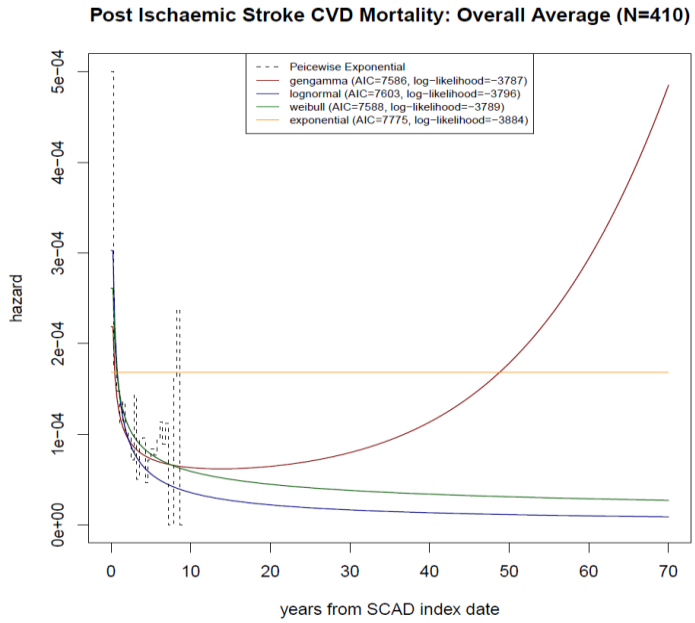


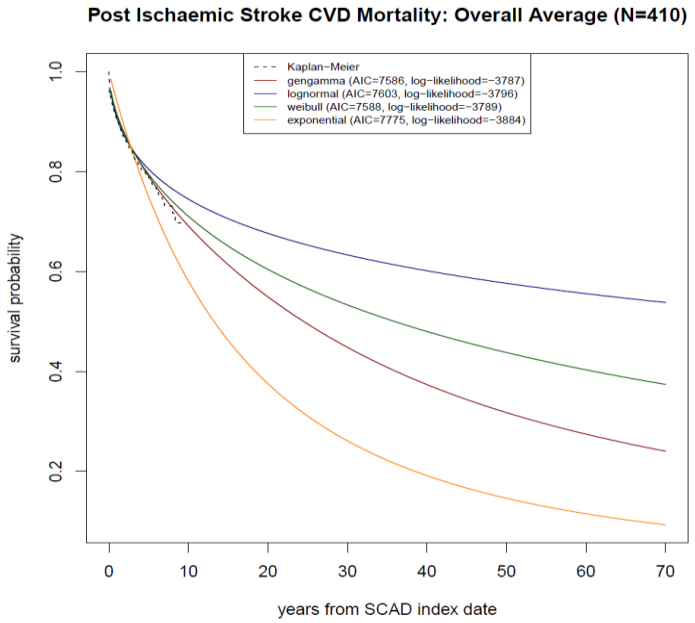
Equation 7: Post MI Fatal Non-CVD	GenGamma	LogNormal	Weibull	Exponential
Sociodemographic characteristics				
Age in men	0.88 (0.86-0.89)	0.87 (0.85-0.88)	0.87 (0.86-0.89)	0.91 (0.90-0.92)
Age in women	1.03 (1.00-1.05)	1.02 (0.99-1.05)	1.03 (1.00-1.05)	1.01 (1.00-1.03)
Women vs men	0.91 (0.64-1.30)	1.04 (0.71-1.52)	0.95 (0.66-1.37)	1.00 (0.81-1.24)
Generalised gamma model parameters				
mu	10.38 (10.02-10.75)	10.38 (10.07-10.69)	10.32 (10.03-10.6)	8.99 (8.87-9.11)
sigma	0.77 (0.37-1.61)	3.27 (3.09-3.45)	1.73 (1.63-1.84)	1
Q	2.47 (0.63-4.3)	0	1	1
Model Fit				
Log-likelihood	-6755.79	-6788.29	-6762.82	-6972.03
AIC	13523.58	13586.59	13535.64	13952.06



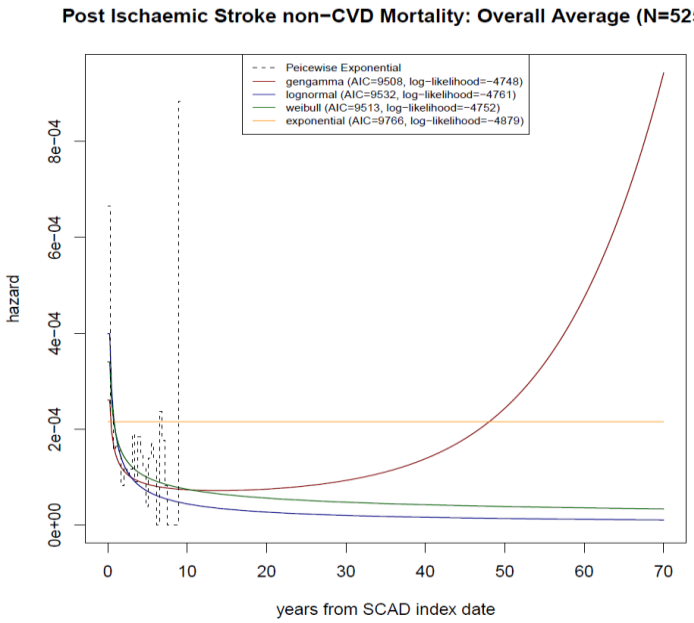


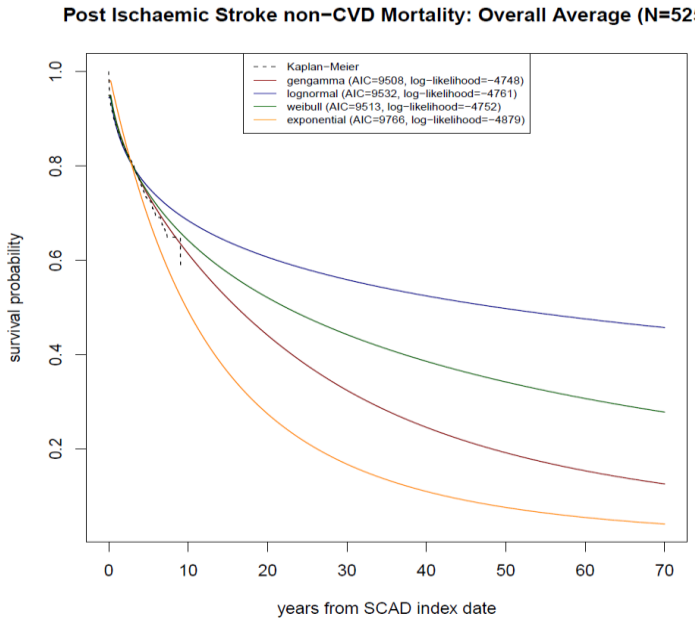
Equation 8: Post Ischaemic Stroke Fatal CVD	GenGamma	LogNormal	Weibull	Exponential
Sociodemographic characteristics				
Age in men	0.91 (0.89-0.93)	0.91 (0.89-0.94)	0.91 (0.89-0.93)	0.94 (0.92-0.95)
Age in women	0.99 (0.96-1.03)	0.99 (0.95-1.03)	0.99 (0.96-1.03)	0.99 (0.97-1.01)
Women vs men	1.52 (0.90-2.54)	1.54 (0.90-2.62)	1.54 (0.91-2.59)	1.35 (0.98-1.86)
Generalised gamma model parameters				
mu	10.42 (9.45-11.39)	10.68 (10.22-11.14)	10.4 (9.98-10.81)	9.08 (8.89-9.27)
sigma	0.59 (0.04-9.78)	3.30 (3.07-3.56)	1.67 (1.54-1.81)	1
Q	3.00 (-5.42-11.42)	0	1	1
Model Fit				
Log-likelihood	-3786.80	-3796.47	-3789.03	-3883.70
AIC	7585.61	7602.95	7588.07	7775.40



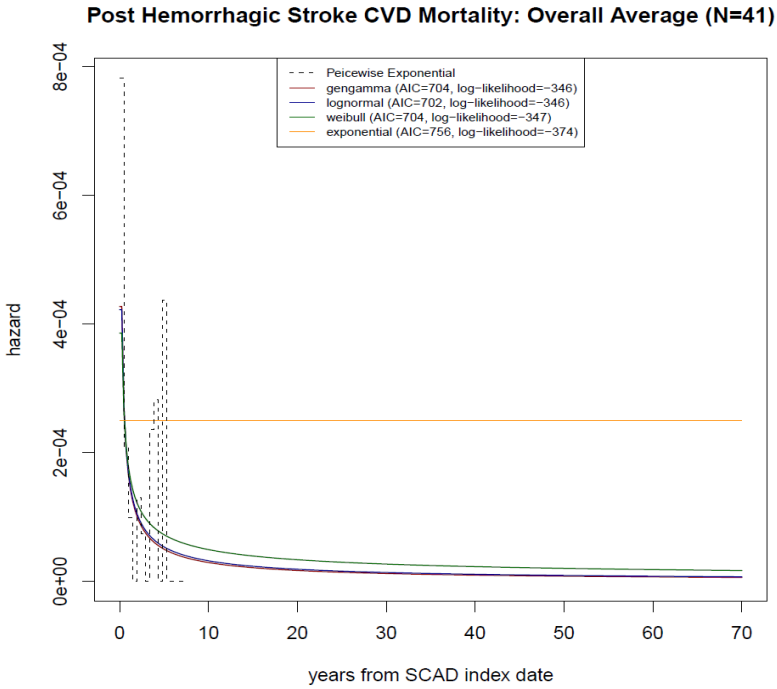


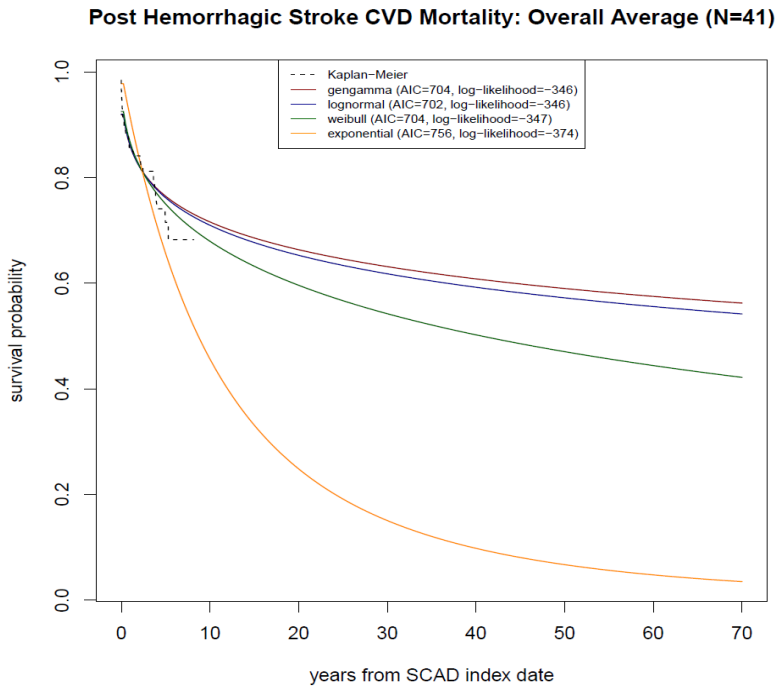
Equation 9: Post Ischaemic Stroke Fatal Non-CVD	GenGamma	LogNormal	Weibull	Exponential
Sociodemographic characteristics				
Age in men	0.93 (0.91-0.95)	0.93 (0.91-0.95)	0.93 (0.91-0.95)	0.95 (0.94-0.96)
Age in women	0.99 (0.97-1.03)	1.01 (0.97-1.04)	1.00 (0.97-1.03)	1.00 (0.98-1.02)
Women vs men	1.48 (0.97-2.26)	1.59 (1.02-2.49)	1.50 (0.97-2.31)	1.32 (1.02-1.71)
Generalised gamma model parameters				
mu	9.92 (8.53-11.3)	9.86 (9.49-10.23)	9.80 (9.47-10.13)	8.70 (8.54-8.85)
sigma	0.53 (0.01-38.25)	3.23 (3.02-3.45)	1.69 (1.57-1.82)	1
Q	3.40 (-11.15-17.96)	0	1	1
Model Fit				
Log-likelihood	-4747.78	-4760.80	-4751.72	-4879.00
AIC	9507.57	9531.60	9513.44	9765.99





Equation 10: Post Hemorrhagic Stroke Fatal CVD	GenGamma	LogNormal	Weibull	Exponential
Sociodemographic characteristics				
Age in men	0.88 (0.80-0.96)	0.88 (0.80-0.96)	0.89 (0.81-0.97)	0.94 (0.90-0.97)
Age in women	1.02 (0.87-1.20)	1.02 (0.87-1.19)	1.04 (0.90-1.19)	1.02 (0.96-1.08)
Women vs men	0.79 (0.09-6.85)	0.85 (0.11-6.58)	1.06 (0.15-7.27)	1.19 (0.52-2.76)
Generalised gamma model parameters				
mu	10.95 (8.99-12.9)	11.02 (9.31-12.74)	10.79 (9.18-12.41)	8.58 (8.06-9.09)
sigma	4.60 (2.04-10.36)	4.14 (3.27-5.25)	2.25 (1.73-2.92)	1
Q	-0.26 (-2.4-1.89)	0	1	1
Model Fit				
Log-likelihood	-346.21	-346.24	-346.99	-373.82
AIC	704.42	702.48	703.98	755.64





Equation 11: Post Hemorrhagic Stroke Fatal Non-CVD	GenGamma	LogNormal	Weibull	Exponential
Sociodemographic characteristics				
Age in men	0.91 (0.84-0.99)	0.91 (0.84-0.99)	0.92 (0.85-0.99)	0.95 (0.91-0.99)
Age in women	0.95 (0.81-1.12)	0.95 (0.81-1.11)	0.94 (0.8-1.09)	0.97 (0.91-1.04)
Women vs men	5.81 (0.68-49.96)	5.64 (0.62-51.43)	5.70 (0.57-57.48)	2.59 (0.88-7.65)
Generalised gamma model parameters				
mu	9.86 (7.32-12.39)	10.24 (8.70-11.79)	10.22 (8.81-11.63)	8.51 (8.03-9.00)
sigma	4.55 (2.38-8.72)	3.68 (2.85-4.75)	2.01 (1.53-2.66)	1
Q	-0.61 (-2.86-1.64)	0	1	1
Model Fit				
Log-likelihood	-305.31	-305.49	-306.60	-323.30
AIC	622.63	620.97	623.20	654.61