ECONOMIC ANALYSIS ALONGSIDE MULTINATIONAL STUDIES

By

RAYMOND AWUAH OPPONG

A thesis submitted to the

University of Birmingham

For the degree of

DOCTOR OF PHILOSOPHY

Health Economics Unit

Institute of Applied Health Research

University of Birmingham

April 2016

ABSTRACT

Conducting economic evaluations alongside multinational studies presents a range of diverse challenges which have contributed to a lack of consensus on how they should be approached particularly because of the difficulties of resolving between country differences. This thesis examines the implications of conducting economic evaluation alongside multinational studies and (i) explores different approaches to obtaining unit costs; (ii) investigates the impact of using different tariffs to value EQ-5D health state descriptions; and (iii) provides a systematic comparison of the pooled and split approaches to economic evaluation alongside multinational trials. This study documents challenges that have been reported in published studies and makes recommendations to help researchers undertake economic evaluations alongside multinational studies.

Results indicate that the main challenge related to dealing with the differences between countries. Collecting unit cost data in all participating countries proved a difficult task, but was most effectively done by collaborating/direct contact with project partners and researchers/health economists from participating countries. Applying different EQ-5D value sets within the context of multinational trials did not make a difference to the conclusions in most cases. However, it is recommended that results from various tariffs are compared within sensitivity analysis. This study also showed that the choice of whether to pool or split the data can lead to different conclusions and recommendations about the cost-effectiveness of interventions. The culmination of this work is a 10 point checklist to guide good practice in the design, conduct and analysis of multinational economic evaluation studies and also highlights many areas where further research is needed. The work provides researchers, policy makers and stakeholders with additional insight into the economic analysis of multinational studies.

ACKNOWLEDGEMENT

I would like to thank my supervisors Dr Sue Jowett and Professor Tracy Roberts for their guidance and support throughout this study. I really appreciate the time and efforts that they have put into helping me complete this work. Thanks to Professor Joanna Coast and Professor Richard Smith, who I worked with on the GRACE project. Many thanks to the GRACE governing board Professor Herman Goossens, Professor Paul Little, Professor Christopher Butler, Professor Theo Verheij and Professor Samuel Coenen for their advice and support and for allowing me to use data from the GRACE project for my PhD work. I also like to thank and acknowledge Kerry Hood, Jacqui Nuttall, Billingsley Kaambwa, Mark Jit, Michael Moore, Hasse Melbye and Sigvard Mölstad who have contributed to the work presented here.

To my colleagues at the Health Economics Unit at University of Birmingham, I would like to say thank you for all your help and support. Thanks to Jesse Kigozi, Lazaros Andronis, Hasan Basarir and Hareth Al-Janabi for their advice and experience which enabled me to put the PhD thesis together.

I would also like to say a big thank you to my wife Eunice and daughters Hannah and Ellie who have been supportive of me through this period. Thanks to Mum, Dad, Christie, Agnes, Felix, Pastor Ransford and Mama Joyce for their advice and prayers.

Finally, I would also like to thank God for his mercies and kindness and giving me the strength to undertake this study.

LIST OF PUBLICATIONS

Some of the work presented in this thesis (Chapters 3, 5, 6 and part of Chapter 7) has been published in the following peer reviewed papers:

- **Oppong R**, Coast J, Hood K, Nuttall J, Smith RD, Butler CC (2011) Resource use and costs of treating lower respiratory tract infections in 13 European countries: Challenges and Results *European Journal of Health Economics* 12 PP 319-329
- **Oppong R**, Kaambwa B, Nuttall J, Hood K, Smith RD, Coast J (2011) Assessment of the Construct Validity of EQ-5D in patients with acute cough/LRTI *Applied Research in Quality of life* 6(4) PP 411-423
- **Oppong R**, Kaambwa B, Nuttall J, Hood K, Smith RD, Coast J (2013) The Impact of Using Different Tariffs to Value EQ-5D health state descriptions: An example from a study of acute cough/LRTI in seven countries *European Journal of Health Economics* 14(2) PP 197-209
- **Oppong R**, Jit M, Smith RD, Butler CC, Melby H, Mölstad S, Coast J (2013) Cost-Effectiveness of Point of Care CRP Testing to Inform Antibiotic Prescribing Decisions *British Journal of General Practice* 63(612) PP e465- e471
- **Oppong R**, Jowett S, Roberts T (2015) Economic evaluation alongside multinational studies: A systematic review of empirical studies *PLoS ONE* 10(6)
- **Oppong R**, Smith RD, Little P, Verheij T, Butler CC, Goossens H, Coenen S, Moore M, Coast J (2016) Cost-effectiveness of Amoxicillin for LRTI in Primary Care: An economic evaluation accounting for the costs of antimicrobial resistance *British Journal of General Practice* ***DOI:*** *10.3399/bjgp16X686533*

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION 1
1.1 Background 1
1.2 Research purpose and objectives 4
1.3 Outline/structure of the thesis 5
CHAPTER 2 BACKGROUND: THEORETICAL FOUNDATIONS OF ECONOMIC EVALUATION AND MULTINATIONAL TRIALS 7
2.1 Introduction 7
2.2 Theoretical foundations of economic evaluation 7
2.2.1 Positive and normative economics 8
2.2.2 Welfare economics and Welfarism 9
2.2.3 Extra-welfarism 12
2.3 Types of economic evaluation techniques 14
2.4 Economic evaluation in practice 17
2.5 Economic evaluation alongside multinational trials 27
2.6 Analytical approaches to economic evaluation alongside multinational trials 28
2.6.1 Fully pooled analysis 28
2.6.2 Fully split analysis 29
2.6.3 Partially split analysis 30
2.7 Issues in the analysis of multinational trials 32
2.7.1 Costing 32
2.7.2 Outcomes in multinational trials 38
2.7.3 Generalisability and transferability 38
2.7.4 Multilevel modelling for the analysis of data from multinational trials 43
2.7.5 Some unresolved issues 45
2.8 Summary 45
CHAPTER 3 REVIEW OF THE LITERATURE ON ECONOMIC EVALUATION ALONGSIDE MULTINATIONAL TRIALS 47
3.1 Introduction 47
3.2 Background to the review 47
3.3 Methods 48
3.3.1 Search strategy 48
3.3.2 Selection of papers for review 49
3.3.3 Data extraction 51

3.4 Results 51
3.4.1 Summary of identified studies 53
3.4.2 Classification of studies based on geographical region and gross national income (GNI) 54
3.4.3 Health outcomes 55
3.4.4 Costing and study perspective 63
3.4.5 Analytical approach to economic evaluation 67
3.4.6 Methods for addressing the multinational nature of the data and ensuring the generalisability and transferability of results 67
3.4.7 Methodological approaches to dealing with the multinational trials 68
3.4.8 Challenges associated with the economic evaluation of multinational trials 69
3.4.9 Assessment of studies based on the Euroheed transferability checklist 71
3.5 Discussion 71
3.5.1 Summary of main findings 71
3.5.2 Strengths and limitations of the study 74
3.5.3 Implications for future research 75
3.6 Conclusion 76
3.7 Summary 77
CHAPTER 4 SUMMARY OF THE CASE STUDIES 78
4.1 Introduction 78
4.2 Lower respiratory tract infections and antimicrobial resistance 78
4.3 The Genomics to combat Resistance against Antibiotics in Community-acquired LRTI in Europe (GRACE) project 80
4.4 Workpackage 8 (Observational study on determinants of antibiotic use) 82
4.5 Workpackage 10 (Randomised controlled trials) 83
4.5.1 Overview of the WP10A trial design (Case study 1) 84
4.5.2 Overview of the WP10B/ GRACE INTRO trial design (Case study 2) 85
4.6 Application of the case studies 87
4.7 Summary 88
CHAPTER 5 COSTING IN ECONOMIC ANALYSIS ALONGSIDE MULTINATIONAL STUDIES 89
5.1 Background 89
5.2 Methods 91
5.2.1 Study design 91

5.2.2 Patients and setting 91
5.2.3 Collection of resource use data 91
5.2.4 Valuation of resource use data 93
5.2.5 Data analysis 99
5.2.6 Sensitivity analysis 99
5.3 Results 100
5.3.1 Patient characteristics 100
5.3.2 Success in methods of obtaining unit cost data for the valuation of resource use 102
5.3.3 Resource use 104
5.3.4 Health care costs 109
5.3.5 Days off work and loss of productivity 112
5.3.6 Sensitivity analysis 112
5.4 Discussion 114
5.4.1 Suggestions and future work 119
5.5 summary 121
CHAPTER 6 HEALTH OUTCOMES IN ECONOMIC ANALYSIS ALONGSIDE MULTINATIONAL STUDIES: AN EXPLORATION OF THE CHOICE OF EQ-5D VALUE SETS 122
6.1 Introduction 122
6.2 Aim 1: Assessing the construct validity of the EQ-5D-3L in patients with acute cough/lower respiratory tract infections 125
6.2.1 Methods 125
6.2.2 Data collection: Case report form 126
6.2.3 Data analysis 126
6.2.4 Results 130
6.3 Aim 2: A comparison of the United Kingdom value set (UKVS) with the European harmonised value set (EVS) and the country-specific value set (CVS) in the context of a multinational study 139
6.3.1 Value sets 139
6.3.2 Methods 144
6.3.3 Results 148
6.4 Discussion and conclusion 162
6.4.1 Summary of chapter 162
6.4.2 Validity of the EQ-5D-3L 162

6.4.3 Comparison of various EQ-5D-3L tariffs in the context of multinational trials ... 163
6.4.4 Comparison with other studies ... 164
6.4.5 Strengths and limitations of the study ... 165
6.4.6 Implications for policy and practice ... 165
6.4.7 Conclusion ... 166
6.5 Summary ... 167
CHAPTER 7 ANALYTICAL APPROACHES TO ECONOMIC EVALUATION ALONGSIDE MULTINATIONAL TRIALS ... 172
7.1 Background ... 172
7.2 Methods: Case study 1 (GRACE WP10A) ... 174
7.2.1 Study design ... 174
7.2.2 Resource use data ... 174
7.2.3 Unit costs ... 175
7.2.4 Health outcomes ... 177
7.2.5 Data analysis ... 177
7.2.6 Analytical approaches explored (Case study 1) ... 178
7.3 Results from Case study 1 ... 179
7.3.1 Baseline characteristics ... 179
7.3.2 Results: fully pooled one country costing (Approach I) ... 181
7.3.3 Results: fully pooled multicountry costing (Approach II) ... 187
7.3.4 Results: fully split one country costing (Approach III) ... 188
7.3.5 Results: fully split multicountry costing (Approach IV) ... 189
7.4 Methods: Case study 2 (WP10 B) ... 192
7.4.1 Study design ... 192
7.4.2 Resource use ... 192
7.4.3 Unit costs ... 192
7.4.4 Health outcomes ... 193
7.4.5 Antibiotic prescribing ... 194
7.4.6 Data analysis ... 194
7.4.7 Analytical approaches explored ... 195
7.5 Results: Case study 2 ... 195
7.5.1 Fully pooled one country costing (FPOC) (Approach I) ... 196
7.5.2 Results: fully pooled multicountry costing (FPMC) (Approach II) ... 201
7.5.3 Results: fully split one country costing (FSOC) (Approach III) ... 205

7.6 Comparison of results from analytical approaches (Case studies 1 and 2) 207
7.6.1 Resource use 208
7.6.2 Costs 209
7.6.3 Health outcomes 210
7.6.4 Cost-effectiveness 211
7.7 Sensitivity analysis 212
7.8 Discussion 221
7.8.1 Summary of main findings 221
7.8.2 Strengths and limitations of the study 225
7.8.3 Comparison with existing literature 226
7.8.4 Implications for policy and practice 226
7.8.5 Future research 228
7.9 Conclusion 229
7.10 Summary 230
CHAPTER 8 GENERAL DISCUSSION AND CONCLUSION 231
8.1 Introduction 231
8.2 Principal findings of the research 232
8.3 Strengths and limitations of the research 238
8.4 Comparison with other studies 240
8.5 Recommendations, implications for policy and future research 241
8.6 Conclusions 263
Appendix 1: Data extraction form 265
A1.1 Data extraction form 265
Appendix 2: Distribution of countries across studies 266
A2.1 Distribution of countries across studies 266
Appendix 3: Summary of countries based on World Bank classifications 278
A3.1 Summary of countries based on World Bank classifications 278
Appendix 4: Transferability checklist 281
A4.1 Summary of the transferability checklist 281
A4.2 Assessment of studies based on the transferability checklist 282
Appendix 5: List of participating institutions and countries in the GRACE network 284
Appendix 6: Study protocols for the economic analysis alongside the GRACE project 285
A6.1 Protocol for observational study: Estimating the resource use and costs of treating LRTI alongside the observational cough study in workpackage 8 (Observational study). 285

A6.2 Protocol for work alsongside Case study 1: Exploring the cost-effectiveness of antibiotics compared with placebo: A randomised placebo controlled double-blind trial. 303
A6.3 Protocol for work alongside Case study 2: Exploring the cost-effectiveness of a web based intervention aimed at reducing antibiotic prescribing 324
Appendix 7: Approaches to obtaining unit cost in participating countries 333
A7.1 Example of letter sent out to UK HESG members 333
A7.2 Example of cost questionnaire sent out to National Network Facilitators and National Network Coordinators 334
Appendix 8: Conversion rates derived from the market basket approach 351
A8.1 Conversion rates used in the market basket approach 351
Appendix 9: Comparison of mean costs obtained from nominal exchange rates and purchasing power parities 352
A9.1 Mean costs from exchange rates and PPPs 352
Appendix 10: Distribution of EQ-5D scores across countries with different value sets 353
Appendix 11: Results from the partially split analysis 357
A11.1 Resource use, cost and health outcomes (Case study 1) 359
A11.2 Cost-effectiveness (Case study 1) 360
A11.3 Resource use (Case study 2) 360
A11.4 Cost (Case study 2) 361
A11.5 Mean EQ-5D scores over 4 weeks (Case study 2) 362
A11.6 Cost-effectiveness analysis (Case study 2) 363
Appendix 12: Winbugs code for estimating country-specific cost-effectiveness 364
Appendix 13: Country-specific cost-effectiveness acceptability curves 366
A13.1 Country-specific cost-effectiveness acceptability curves for antibiotics vs control (sensitivity analysis with one country costing) 366
A13.2 Country-specific cost-effectiveness acceptability curves for antibiotics vs control (sensitivity analysis with multi country costing) 367
LIST OF REFERENCES 368

LIST OF TABLES

Table 2.1: Summary of economic evaluations in European countries....23
Table 2.2 Analytical approaches to the economic evaluation of multinational trials....31
Table 2.3 Generalisability of pooled and split approaches....42
Table 3.1 Example of search terms used (Medline and Embase search)....54
Table 3.2 Summary of studies that met the inclusion criteria....56
Table 3.3 Specific characteristics of studies included in the review....65
Table 5.1 Sources of valuation data....97
Table 5.2 Exchange rates and GDP....98
Table 5.3 Number of practices, general practitioner (GPs) and patients....102
Table 5.4 Resource use (Standard deviation)....107
Table 5.5 Mean costs £ (Standard deviation)....110
Table 5.6 Total mean costs £ (Standard deviation)....113
Table 6.1 Responses to dimensions of the EQ-5D-3L....131
Table 6.2 Relationship between dimension on the EQ-5D-3L and individual items on the Symptoms diary (P-Values) using Chi squared test....132
Table 6.3 Regression results: Relationship between EQ-5D-3L and Symptom severity scores....133
Table 6.4 Relationship between Smoking and EQ-5D-3L [Mean (SD)]....134
Table 6.5 Chronic diseases and EQ-5D-3L scores....134
Table 6.6 EQ-5D-3L scores over time and sensitivity to change (standardized response mean)....136
Table 6.7 Patient reported recovery and EQ-5D-3L scores....137
Table 6.8 Known group validity....138
Table 6.9 A summary of EQ-5D-3L Value sets used in this study....143
Table 6.10 Study population characteristics....149
Table 6.11 Mean (SD) EQ-5D-3L scores, change over time and Standardized response mean....153
Table 6.12 QALY gains over four weeks....157
Table 6.13 Health gains between interventions....159
Table 6.14 Incremental cost-effectiveness ratios with different tariffs....160
Table 6.15 Sensitivity analysis....161
Table 7.1 Analytic approaches to economic evaluation alongside multinational trials....173
Table 7.2 Source of valuation data....176
Table 7.3 Country contribution to sample size....180
Table 7.4 Resource use mean (SD) (All approaches)....182
Table 7.5 Costs (£) mean (SD) (All approaches)....183
Table 7.6 Health outcomes mean (SD) (All approaches)....185
Table 7.7 Cost-effectiveness: Base case analysis (All approaches)....186
Table 7.8 Country contribution to sample size (Case study 2)....196
Table 7.9 Resource use (All approaches)....197

Table 7.10 Costs (£) (All approaches) 199
Table 7.11 Mean EQ-5D scores over 4 weeks (All approaches) 202
Table 7.12 Cost-effectiveness analysis (FPOC) 203
Table 7.13 Cost-effectiveness analysis (FPMC) 206
Table 7.14 Cost-effectiveness analysis (FSOC) 207
Table 7.15 Summary of results obtained using various approaches to economic evaluation alongside multinational trials 208
Table 7.16 Cost-effectiveness UKVS and EVS (Case study 1) 214
Table 7.17Cost-effectiveness UKVS and EVS (Case study 2) 215
Table 7.18 Country-specific cost-effectiveness results obtained from Bayesian hierarchical model 218
Table 7.19 Sensitivity analysis: Participants with complete cases (All approaches) 220
Table 8.1 Checklist to aid the design and conduct of economic evaluations alongside multinational trials 251

LIST OF FIGURES

Figure 2.1 Identifying resource use and unit costs in a multinational trial.....37
Figure 2.2 Examples of multilevel data structures.....43
Figure 3.1 Summary of stage 1, stage 2 and stage 3 criteria 52
Figure 4.1 Summary of WP10A study (Case study 1).....85
Figure 4.2 Summary of WP10B study (Case study 2).....87
Figure 5.1 Patient flow chart.....101
Figure 6.1 Bland-Altman plots CVS, EVS and UKVS 168
Figure 7.1 Cost-effectiveness acceptability curves (Case study 1).....187
Figure 7.2 Net benefit curves (Case study 1).....191
Figure 7.3 Cost-effectiveness acceptability curves for usual care vs CRP (Case study 2)....204
Figure 7.4 Cost-effectiveness acceptability curves for usual care vs comm skills (Case study 2) 204
Figure 7.5 Cost-effectiveness acceptability curves for usual care vs combined intervention (Case study 2) 205

LIST OF ABBREVIATIONS

CBA	Cost-benefit analysis
CCA	Cost-consequences analysis
CEA	Cost-effectiveness analysis
CMA	Cost-minimisation analysis
COPD	Chronic obstructive pulmonary disorder
CUA	Cost-utility analysis
CVS	Country-specific value set
DALY	Disability adjusted life year
EQ-5D	EuroQol 5-dimensions
EU	European Union
EVS	European value set
FPOC	Fully pooled one country costing
FPMC	Fully pooled multicountry costing
FSOC	Fully split one country costing
FSMC	Fully split multicountry costing
GDP	Gross domestic product
GNI	Gross National Income
GRACE	Genomics to combat Resistance against Antibiotics in Community-acquired LRTI in Europe
HESG	Health Economists' Study Group
HRQoL	Health related quality of life
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
IQWIG	Institute for Quality and Efficiency in Health Care
LRTI	Lower respiratory tract infection
NHS	National Health Service

NHS/PSS	National Health Service and Personal Social Services
NICE	National Institute for Health and Care Excellence
NNC	National Network Coordinator
NNF	National Network Facilitator
OECD	Organisation for economic cooperation and development
PPP	Purchasing power parities
PSOC	Partially split one country costing
PSMC	Partially split multicountry costing
QALY	Quality adjusted life years
TTO	Time trade-off
UKVS	United Kingdom value set
VAS	Visual Analogue Scale
WHO	World Health Organisation

DECLARATION

I declare that the work presented in this thesis was carried out in accordance with the requirements of the University of Birmingham. The work is original, except where indicated in the text. No part of this thesis has been submitted for any other academic award.

CHAPTER 1 INTRODUCTION

1.1 Background

In the health care sector, new health technologies are constantly being developed and clinical trials are often used to assess their effectiveness (Yusuf et al. 1984; Black, 1996; Meinert, 2012; Piantadosi, 2013). In addition to obtaining information on clinical effectiveness of new health technologies, decision makers also need to know whether a new health technology offers value for money in order to make informed resource allocation decisions. Economic evaluation which is defined as "the comparative analysis of alternative courses of action in terms of both their costs and consequences" (Drummond et al. 2015 PP: 4) can be used for this purpose, and is conducted alongside a clinical trial or using decision modelling. According to the World Health Organisation (WHO), a clinical trial can be defined as "any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes." *(WHO, accessed on 12/01/16).* Clinical trials offer a vehicle through which economic evaluation of health technologies can be made and many funding bodies, such as the UK National Institute for Health Research's (NIHR) Health Technology Assessment (HTA) programme, are now routinely requesting the assessment of cost-effectiveness as part of randomised controlled trials (Petrou and Gray, 2011). In addition, clinical trials offer a great opportunity to collect information on the effectiveness and resource use of health care interventions simultaneously, thus reducing the marginal cost of collecting this information.

Multinational trials can be defined as trials that take place in more than one country or jurisdiction (Marschner, 2010). Approximately 11% of all trials are multinational and 95% of countries have been involved in this type of trial (Richter, 2014). Most multinational trials

have involved developed countries, however the involvement of South American countries has increased over time (Richter, 2014). The main reason for conducting multinational trials is to increase statistical power and this is particularly important when assessing interventions for rare conditions where obtaining the required sample size in one country is particularly difficult (Marina et al. 2009; Augustine et al. 2013). In addition to the requirement of increasing sample size, further reasons for conducting multinational trials include the need to replicate trial results in a wider setting, increase patient representation to ensure results are generalisable, answer a research question promptly, reduce costs by recruiting patients in countries where costs associated with conducting trials are lower and speed up the development of new drugs and other health technologies. A further reason for conducting multinational trials is to increase the relevance of the research or health technology to a greater number of countries in order to potentially influence clinical practice in countries involved in the trial (Cook et al. 2003; Drummond and Pang, 2001; Mulligan and Fox-Rushby, 2005; Richter, 2014). Multinational trials are most often drug trials carried out in disease areas such as oncology and infectious diseases. A recent study showed that a higher proportion of trials assessing rare diseases/orphan drugs employ a multinational design (Bell and Tudur Smith, 2014). However, multinational trials can also be public health interventions studies such as behavioural change interventions (Hsiehchen et al. 2015).

Research has shown that most multinational trials are industry funded (Hsiehchen et al. 2015). Approximately 30% of industry funded trials are carried out in more than one country compared to just 3% of academic trials (ECRIN accessed 30th November, 2016). The main reason for this is that there are barriers to cross border funding opportunities that are available for academic trials and in general national funding bodies have limits on funding that is available for international research (ECRIN accessed 30th November, 2016). Multinational trials are particularly useful to the pharmaceutical industry, since they are able to assess the

effectiveness of their products in different settings and also collect data from different settings which is useful for drug pricing and marketing as well as strategic planning (Rivero-Arias and Gray, 2010).

The increase in the number of multinational clinical trials (Wild et al. 2009) has led to a corresponding increase in the number of economic evaluations that have been conducted alongside them (Briggs 2010). However, due to the fact that data collection take place in more than one country, this type of analysis is particularly problematic. One common problem relates to dealing with costs from a number of countries, in particular, obtaining unit costs in all participating countries is a difficult task (Torti et al. 2006). Additional major challenges include inconsistent reporting of results (Rivero-Arias and Gray, 2010) and the problem of generalisability and transferability (Drummond et al. 2009). Decision makers are often only interested in results for only their own country and may not be interested in those obtained from other countries unless the results are generalisable (Wilke et al. 1998; Manca et al. 2010). In the same way that country specific trial based economic evaluations are used to inform a model when an intermediate outcome needs translating to explore health and resource use impact for a longer time horizon or just to extend the time horizon, this is also true of data from multinational trials (Vermer and Rutten van Molken, 2013). There is therefore the need to develop methods that would ensure that these inputs can be made more useful for decision making in individual country contexts when the need arises.

A number of approaches have been used by researchers who have conducted economic evaluations alongside multinational trials. These range from fully pooled studies that have considered resource use and effectiveness data from all participating countries and applied unit cost data from just one country, to fully split/partially split studies where resource use and effectiveness data are collected from one or a subset of countries and applied to costs from just one country or from a number of countries (Reed et al 2005). However, it is not

clear which approach is the most appropriate and under which circumstances they should be used. Methods aimed at ensuring that the results can be used and transferred to an individual country have also been developed. These range from very simple approaches such as adjusting resource use to very complex statistical models such as multilevel modelling and decision analytic modelling (Drummond et al., 2009; Reinhold et al., 2010; Manca et al 2010). The extent to which these methods have been used in practice is however unclear.

Presently, there is no consensus with regards to many aspects of conducting economic evaluation alongside multinational trials such as how to obtain country specific cost-effectiveness results and with respect to health outcomes, what value set to apply when using the EQ-5D (Rabin and de Charro, 2001) in a multinational trial setting. Recent research has also concluded that there is a vast variation in methods that are being used for the conduct of economic analysis alongside multinational trials (Torti et al. 2006). Therefore, there is the need to compare and critique various methods in order to come up with appropriate and accepted methods to ensure that scarce resources are allocated effectively.

1.2 Research purpose and objectives

The main objective of this research is to explore the implications of conducting economic evaluation alongside multinational studies/trials and to document the challenges associated with conducting economic evaluation alongside multinational trials. The specific objectives of this research include the following:

1. To review the literature in order to document challenges that have been reported in published economic evaluations alongside multinational trials.

2. To explore various approaches to obtaining and estimating unit costs in multinational trials using case studies.
3. To explore the impact of using different tariffs to value EQ-5D health state descriptions in economic evaluation alongside multinational trials.
4. To compare methods that have been used to conduct economic evaluation alongside multinational trials.
5. To make recommendations to guide the design and conduct of future economic evaluations carried out alongside multinational trials.

1.3 Outline/structure of the thesis

The rest of this thesis is structured as follows:

Chapter 2*:* Chapter 2 provides a background to economic evaluation, multinational trials and the methodological approaches to economic evaluation alongside multinational trials.

Chapter 3*:* *Review of economic evaluation alongside multinational trials*: In Chapter 3, a systematic review of the literature in the area of economic analysis alongside multinational trials is presented in order to determine current practice and also identify challenges that have been reported by researchers. The focus here is on published economic evaluations alongside multinational trials.

Chapter 4*:* *Case studies:* In Chapter 4, an overview of the Genomics to combat resistance to antibiotics in community acquired LRTI in Europe (GRACE) project is presented and an in depth description of the three case studies that would be used for the analysis are introduced.

Chapter 5*:* *Costing in multinational studies:* Chapter 5 explores some of the issues associated with costing alongside multinational trials using case studies.

Chapter 6*:* *Outcome measurement in multinational studies:* Chapter 6 focuses on the EQ-5D as an outcome measure and applies various value sets to the case studies.

Chapter 7*:* *Analytical approaches to economic evaluation alongside multinational trials:* This chapter focuses on comparing the pooled and split approaches to analysis of multinational trials using case studies.

Chapter 8*:* *Discussion:* Chapter 8 presents a discussion of the main findings and conclusions of the research.

CHAPTER 2 BACKGROUND: THEORETICAL FOUNDATIONS OF ECONOMIC EVALUATION AND MULTINATIONAL TRIALS

2.1 Introduction

The aim of this Chapter is to provide some background information about the research work and it consists of two major parts. A general overview of the theoretical foundations of economic evaluation in health care is presented in the first part followed by an overview of the analytical approaches to economic evaluation alongside multinational trials discussed in relation to costs, outcomes and analytical approaches to economic evaluation.

2.2 Theoretical foundations of economic evaluation

Economic evaluation in healthcare seeks to help decision makers with appraising health technologies and allocating resources by comparing competing alternatives in terms of their costs and consequences (Drummond et al. 2015). When a new health technology is introduced, there are two main questions that decision makers seek to address. (I) Whether the health technology is effective i.e. if it improves patient outcomes and (II) whether it provides value for money (Whitehurst and Bryan, 2013). Health economics in general, is concerned with the second question, and economic evaluation serves as a tool which helps to

determine whether the available resources should be spent on the health technology in question or whether it could be better spent elsewhere (Cunningham, 2000).

In many developed countries, specific bodies have been set up to help ensure that healthcare resources are allocated appropriately and a process known as health technology assessment (HTA) is normally used to achieve this (Gray and Wilkinson, 2016). The UK established what is now known as the National Institute for Health and Care Excellence (NICE) in 1999, a body charged with providing the National Health Service (NHS) and those who rely on it with advice on effective, good value health care (NICE accessed 03/10/15). NICE guidance includes technology appraisals in which the clinical and cost-effectiveness of new technologies are assessed to ensure that the technologies that offer value for money are available to all (NICE accessed 03/10/15). Other developed countries have established similar national bodies that have been tasked with the health technology assessment process. Examples of such institutions include: Institute for quality and efficiency in health care (IQWiG) in Germany and The Swedish Dental and Pharmaceutical Benefits Agency in Sweden (www.iqwig.de; www.tlv.se). Countries such as Belgium, Netherlands and Spain also have such bodies however; the extent to which economic evidence is required varies across countries. In the UK, however, economic evaluation has been used as a basis for developing guidance since the inception of NICE in 1999 (NICE accessed 03/10/15).

Economic evaluation has its theoretical foundations in welfarism and extra-welfarism which are now considered in the section below.

2.2.1 Positive and normative economics

Positive economics is considered to be an objective science which aims to provide "a system of generalisations that can be used to make correct predictions about the consequences of any

change in circumstances" (Friedman, 1953 p.4). Positive economics deals with "what is" as opposed to "what ought to be" and is in principle independent of ethical positions and value judgements using statements of fact and refers to empirical testing. In contrast, normative economics deals with value judgements and opinions (Kolm, 2000). Most of the issues that health economists attempt to address involve making value judgements. For example, in a health care decision making context, determining which treatment option is the most appropriate or how many should be treated etc. involves making judgements. Although health economists may argue that their work is mainly concerned with providing information to aid decision making, rather than make these decisions, the approaches that are used involves judgements (Morris et al. 2012). For example, in economic evaluations, health economists make judgements about the costs and benefits they include. Concepts such as market failure which are positive aspects of health economics are more likely to take an objective approach since they aim to predict and explain. However, there are still value judgements about what should be included in an analysis. The next section of this chapter examines two normative approaches (welfarism and extra-welfarism) which form the theoretical basis of economic evaluation in health care.

2.2.2 Welfare economics and Welfarism

Welfare economics or welfarism is the standard/traditional approach to normative economics and is defined as "the systematic analysis of the social desirability of any set of arrangements, for example a state of the world or allocation of resources, solely in terms of the utility obtained by individuals" (Morris, 2012 p 205). Welfarism is founded on the following normative principles: *The utility principle* which states that individuals are rational and seek to maximise their welfare/utility. *Individual sovereignty,* which states that individuals are the best judges of their own utility and *Consequentialism* which is based on consumer choice theory (Brouwer et al. 2008) and "traditionally only takes into account the outcomes for

people, in terms of their consumption of specific types and quantities of goods and services and the consequent utility obtained" (Morris et al. 2012 p 206). With welfarism, the goodness of any given state can be assessed via the utility obtained by the individuals in that state (Hurley, 2000).

This approach therefore takes the view that utilities of individuals should be the basis for resource allocation decisions, where utility means pleasure/satisfaction (Cookson et al. 2012) and therefore considers the views of individuals as the most important in making social choices. Consequently welfarism limits the evaluative space to individuals (Coast, 2004; Morris et al. 2007; Brouwer et al. 2008) and social welfare is considered to be a product of the utilities of members of society (Morris et al. 2007).

In order to make judgements about the desirability of states of the world, welfarism adopts the Pareto principle, a value judgement which stipulates that overall improvement occurs if one person can be made better off without another being made worse off (Pareto improvement) (Morris et al. 2007). A weak Pareto improvement occurs when a change in the state of the world leads to gains in the utility of everyone whilst a strong Pareto improvement occurs when a change or reallocation of resources leads to one person becoming better off whilst no one else is made worse off (Sen, 1979). This gives a framework for assessing whether a change in the state of the world is socially optimal. If a policy satisfies the Pareto principle then that policy ought to be introduced. Pareto optimality is reached when an increase in an individual's utility can only come about as a result of a decrease in another individual's utility (Broadway and Bruce, 1984; Tsuchiya and Williams, 2001).

In theory, the Pareto principle should be ideal for making decisions in health care since an improvement in society would occur if one person is better off without making others worse off. However, the Paretian principle is not considered practically useful in decision making

since, as a result of scarcity of resources, there will be eventual winners and losers when a policy decision is made (Coast, 2004; Coast et al. 2008).

As a result of this limitation, the compensation principle (potential Pareto improvement) was developed by Kaldor and Hicks. This states that global improvement would occur if individuals who are made better off by a change/policy could potentially compensate those individuals who have been adversely affected by the change and still remain better (Kaldor, 1939; Hicks, 1939). Although this principle could be considered as an improvement, the approach considers the distribution of the welfare as irrelevant since the aim is to improve overall welfare. There are therefore questions in terms of equity with this approach (Coast et al. 2008).

In practice, the economic evaluation methodology based on the compensation principle is cost-benefit analysis which measures benefits in terms of individual's willingness to pay (Coast, 2004; McIntosh et al. 2010; Sugden and Williams, 1978) and would advocate an intervention on efficiency and social welfare grounds if the sum of willingness to pay exceeded the cost imposed, i.e. those who would benefit could compensate those who lost out.

When related to health care, welfarism has a number of limitations. The first relates to equity and distributional concerns. With this approach, individuals value welfare in terms of willingness to pay and this could potentially lead to a situation where resource allocation could potentially favour the rich because they have a greater ability to pay (Coast et al. 2008). It has been noted that "extra-welfarists and many decision makers in the real world of health care are more willing to accept an approach that considers outcomes equitably, rather than accept an approach in which choices are heavily influenced by ability to pay (Weinstein and Manning, 1997 p 127). Another limitation associated with the welfarist approach when

applied to the health sector relates to the market failures that arise as a result of the unique nature of health care (Arrow. 1963). Health care is associated with adverse selection, moral hazards and supplier induced demand and as a result of this, relying on welfarism as the basis for resource allocation in health is not appropriate (Culyer, 1989).

Critiques of welfarism also note that it is restrictive, since it relies on individuals utilities obtained from consumption (Mooney and Russell, 2003) and that other factors are equally as important when making resource allocation/social decisions in health care. Thus, the foundations on which welfarism is built does not allow these other important factors to be considered when making resource allocation decisions (Morris et al. 2007). In addition to this, the ability of utility to capture wellbeing has been challenged by researchers such as Sen who argues that individuals assessment of well-being and utility depends on their circumstances and therefore using individuals utility as a measure of social welfare is misleading (Sen, 1979; Morris et al. 2012).

2.2.3 Extra-welfarism

Despite a practical way of operationalising welfraism in the healthcare sector, an alternative form of normative economics, commonly referred to as extra-welfarism is typically used to inform judgements about the value-for-money of new healthcare interventions. This alternative form of normative economics came about partly for historical reasons and partly in response to some of the perceived limitations of welfarism. Just as the name implies, extra-welfarism expands the evaluative space beyond the maximisation of utility to other relevant outcomes such as health (Brouwer et al. 2008). According to Brouwer, extra-welfarism differs from welfarism in the following ways: "(i) it permits the use of outcomes other than utility (ii) it permits the use of sources of valuation other than the affected individuals (iii) it permits the weighting of outcomes (whether utility or other) according to principles that need not be preference based (iv) it permits interpersonal comparisons of wellbeing in a variety of

dimensions, thus enabling movement beyond Paretian economics." (Brouwer et al. 2008: P 330). The extra-welfarist theory builds upon the work of Amartya Sen (Sen 1980) who pointed out that focusing on individual utility was too narrow and identified the need to go beyond this and adopt a perspective that took "account of the quality of utility and of people's capabilities rather than exclusively of the emotional reaction (i.e. utility) of individuals to the possession of goods or capabilities" (Cookson et al. 2012 P 73). Culyer applied this principle to healthcare by expanding the evaluative space to show how non welfare information such as health could be used (Culyer 1989; Cookson et al. 2012).

Cost-utility analysis, which is one of the most popular approaches used for the economic evaluation of health technologies draws upon this extra-welfarist approach. In the UK for example, the requirements by NICE explicitly state that QALYs should be used to measure health outcomes in cost-effectiveness analysis (NICE, 2013). One of the main criticisms of the extra-welfarist approach relates to the evaluative space used. In theory, extra-welfarism is supposed to include other outcomes apart from utility and does not specifically state that health should be the sole maximand. However, in practice, it is interpreted as health maximisation which is quite narrow (Coast, 2009; Birch and Donaldson, 2003). It is noted that this may have arisen as a "pragmatic response to the methodological challenges in economic evaluation" (Morris et al. 2007 P 236). It can also be argued that extra-welfarism as it is currently used in health care is actually narrower than welfarism since by definition; utility may be derived from the consumption of health care services (Coast, 2009).

There is still quite a lot of debate in health care regarding which of the approaches should be used and each approach clearly has its advantages and shortcomings. Welfarism has the advantage of being theoretically superior whilst extra-welfarism is easier to implement in

practice (Gyrd-Hansen, 2005). The choice between whether to adopt a welfarist or extra-welfarist approach is a value judgement.

A number of alternative economic evaluation techniques have arisen from the above mentioned theories (welfarism and extra-welfarism), all of which can be used alongside trials. However, the extra-welfarist approaches (cost-utility analysis and cost-effectiveness analysis) are used more often in the context of trials. These alternative economic evaluation techniques are now considered in the next section.

2.3 Types of economic evaluation techniques

Economic evaluation forms part of normative economics since it involves value judgements about the cost-effectiveness of interventions. There are various types, which differ mainly in terms of the way benefits/outcomes are considered (Morris et al. 2007; Drummond et al. 2015). These include:

Cost-consequences analysis (CCA): With cost-consequences analysis, all the important costs and consequences of competing alternatives are considered and presented separately to decision makers in a disaggregated form. The problem with cost-consequence analysis is that the decision maker is left to interpret the various options that have been presented, and also assumes that they are capable of assigning the weights to the various options that have been presented (Gray et al. 2010). This approach therefore lacks transparency since the decision is based on the opinion/values of the decision maker as opposed to those of society (McCabe et al. 2009). This approach is not as popular as the others and is not used as the main economic

evaluation approach. However, it is sometimes used in combination with the other approaches, often as a first step in the economic analysis of trials.

Cost-minimisation analysis (CMA): This method is used when the effectiveness of the interventions being compared is considered to be equal. It involves comparing two or more interventions in terms of costs and adopting the least costly option. Currently, the literature in health economics suggests that cost-minimisation analysis should only be considered in equivalence/non inferiority trials (Briggs and O'Brien 2001; Dakin et al. 2013). The problems however, is that it is difficult to find a situation where there is equivalence in terms of treatment effects and as a result, this approach is rarely used in practice. With this approach, statistical tests such as the t-test are used to determine differences in costs between interventions. Although this might seem appropriate, health economists tend to focus on estimating cost and effects jointly which further limits the usefulness of this approach (Briggs and O'Brien 2001; Whitehurst and Bryan 2013).

Cost-effectiveness analysis (CEA): This approach has its theoretical foundations in extra-welfarism and with CEA, outcomes are measured in units that can be counted e.g. natural units such as number of cases detected, improvements in physical function and life years gained. This approach is particularly useful when the natural unit is relevant to the interventions under consideration. For example, the number of cases detected would be an appropriate outcome for comparing screening programmes. One of the main problems with this approach is that its applicability is limited to a particular disease area. For example, the outcome cost per case detected can be used for the comparison of various cancer screening strategies and will not be applicable to non-screening strategies and as a result, comparison of interventions across disease areas is limited (Gray et al. 2010). In addition, CEA does not give an indication of the quality of life from the treatment, although it may be assumed that

improvements in these clinical measures most often translate into improvements in quality of life. It is quite common for the CEA approach to be conducted alongside clinical trials.

Cost-utility analysis (CUA): Cost-utility analysis has its theoretical foundations in extra-welfarism. This approach is considered to be broader than cost-effectiveness analysis and overcomes some of the problems associated with CEA. The most common outcome that is used for the CUA is the quality adjusted life year (QALY) which combines both the length of life gained as a result of a treatment and the associated quality of life within the same metric (Drummond et al. 2015). Unlike the cost-effectiveness approach described earlier, CUA permits comparison across disease areas and unlike the cost-benefit approach described below, it facilitates this comparison without having to place a monetary value on outcomes (Gray et al. 2010; McCabe, 2009). For example; with CEA, it is difficult to compare a lifesaving treatment with a vaccination programme since different types of outcomes are more appropriate for each treatment e.g. number of lives saved versus number of cases detected. However, with cost-utility analysis, these interventions can be compared using a common outcome measure such as the QALY.

In addition, it overcomes some of the problems associated with cost-consequences analysis which is described later since multiple outcomes are combined into a single measure. CUA is considered to be the most widely used economic evaluation approach and it is recommended by decision bodies such as NICE (NICE, 2013). However, the problem with this approach is that it is considered to be narrow and it does not capture some process characteristics of interventions which may be important to patients (Drummond et al. 2015; McCabe et al. 2009). CUA does not also capture all the benefits of a healthcare intervention e.g. the effects on carers and children from a treatment are not captured by the QALY (Whitehead and Ali, 2010; Al-Janabi et al. 2011). Lastly, instruments such as the EQ-5D that are used to measure

outcomes may be insensitive to change in certain disease areas (Brazier et al. 1999; Cleemput et al. 2004; Haywood et al. 2008).

Cost-benefit analysis (CBA): Cost-benefit analysis has its theoretical roots in welfarism. With this approach, the consequences of healthcare programmes are valued in monetary terms using stated preference methods such as willingness to pay and contingent valuation (Drummond et al. 2015; McIntosh et al. 2010). The advantage cost-benefit analysis has over the other approaches is that it is possible to compare not only across disease areas but also to compare across sectors of the economy e.g. a comparison between the education and health sectors of the economy (Gray et al. 2010). One of the main problems with this approach is that individuals may find it difficult to value benefits from healthcare in monetary terms (Ryan et al. 2003). Cost-benefit analysis can be conducted alongside trials. However, unlike the extra-welfarist approaches (CEA and CUA) discussed above, this is quite rare.

2.4 Economic evaluation in practice

Economic evaluation is gradually gaining importance as a tool to aid priority setting in both developed and developing countries (Pitt et al. 2016). In many countries, there is a formal requirement for economic evaluation to be included as an important criterion in reimbursement decisions for new drugs and health technologies. However, the extent to which it is used varies across countries (Franken, 2014). In countries such as the UK, the approach which has formed a basis for resource allocation is based on extra-welfarism (Birch and Gafni, 2002). In the UK, the NICE reference case explicitly states that economic evaluation should be conducted from a National Health Service/Personal Social Services (NHS/PSS) perspective and that health effects should be expressed in terms of QALYs that

are generated preferably from the EQ-5D (NICE, 2013). The process of conducting an economic evaluation involves comparing the costs and benefits of competing alternatives and generating an incremental cost-effectiveness ratio (ICER) which is defined as the difference in costs divided by the difference in benefits. From equation 2.1, A and B represent two competing health care technologies where A is the new technology and B the existing technology.

$$\text{ICER} = \frac{\text{Cost A} - \text{Cost B}}{\text{Benefit A} - \text{Benefit B}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2.1)$$

The ICER represents the cost per unit of benefit. Thus, if it is assumed that the QALY is the measure of benefit used, the ICER would represent the cost of an additional QALY gained. Looking at equation 2.1 above, the lower the ICER, the more preferable the new health technology is. This is because it is more desirable to minimise cost (numerator) and maximise benefits (denominator). However, it is not enough to make a judgement about the cost-effectiveness of health technologies based on the ICER alone and there is therefore the need for some form of external criterion in order to judge whether an intervention is cost-effective or not. In the UK, the cost-effectiveness threshold advocated by NICE which tells us how much per unit of benefit is worth paying is used for this purpose. In other words, the threshold tells us the maximum amount the ICER can be for the intervention to be considered cost-effective (Drummond et al. 2015; Morris et al. 2007).

In the UK, NICE uses a cost per QALY threshold of between £20,000 and £30,000 per QALY to judge whether an intervention is cost-effective or not (Appleby, 2007). When a cost-utility analysis is conducted and the ICER falls below the £20,000 per QALY mark, the intervention is considered to be cost-effective. If the threshold is above £20,000 per QALY, then additional judgements are required about whether the technology/intervention represents

a good use of resources, should it be adopted (McCabe et al. 2008). If the ICER falls above £30,000 per QALY, the intervention is clearly not cost-effective since we would have to pay above an acceptable threshold for a 1 QALY gain. In such an instance, the case for adopting the health technology in the NHS has to be very strong (Raftery, 2006). An example is the case of end of life care where a threshold higher than the recommended £20,000 to £30,000 per QALY is applied (Collins, 2013). There has been much debate about the acceptability of this threshold and it has been suggested that this threshold is arbitrary and has no theoretical basis (Raftery, 2014). A recent study suggests that the threshold should be much lower than it is if it should represent opportunity cost and the revised threshold should range between £13,000 and £18,000 per QALY gained (Claxton et al. 2015). However, this research has been criticised as one which relies on several assumptions (Raftery, 2014, Barnsley et al. 2013) and at present, the threshold in the UK still stands at between £20,000 and £30,000 per QALY.

Other countries also employ a cost-effectiveness threshold, although in most cases, it is not as explicit as in the UK. For example, in Netherlands, the threshold is set at €20,000 per QALY (Zwart-van Rijkom et al. 2000; Jit et al. 2009) whilst $50,000 per QALY is often cited in the USA (Bridges et al. 2010). In addition to this, the World Health Organisation (WHO) recommends that GDP per capita should be used as a threshold since this represents the populations fair share of a countries wealth and it recommends that for an intervention to be cost-effective, the threshold should be between per capita GDP (very cost-effective) and three times per capita GDP (cost-effective) (Hutubessy et al. 2003; Tan-Torres, 2003; WHO 2001). There have been debates about the usefulness of the cost per QALY approach and some have suggested that there is the need to go beyond QALYs when making resource use decisions and that maximising health is not the most important thing that decision makers are interested in. Such studies have suggested that the "use of a single outcome for cost-effectiveness

analysis fails to recognise that decision making involves making judgments about a variety of important effects rather than just one" (Coast, 2004: P 1235).

In low and middle income countries (LMICs), estimating the cost-effectiveness of health care interventions has also become increasingly important (Shillcutt et al. 2009; Pitt et al. 2016). It has been used in several prioritization exercises, such as the World Bank Health Sector Priorities Review (World Bank, 1993) as well as in the WHO Choosing Interventions that are Cost-Effective (WHO-CHOICE) initiative which was developed in 1998 with the objective of "providing policy makers with the evidence for deciding on the interventions and programmes which maximize health for the available resources" (WHO, accessed on 24th October, 2015). In LMICs, the cost per DALY (Disability adjusted life years) is often used as a measure of estimating the cost-effectiveness of interventions. DALYs are more common in developing countries because it has been suggested that these countries account for almost 90 percent of the global burden of disease (Zarate, 2007; Griffiths et al. 2016). However, there have been issues with the choice of the cost-effectiveness threshold in low and middle income countries which is most often left to the discretion of the analyst (Shillcutt et al. 2009), although a cost-effectiveness threshold of $150 per disability adjusted life year has been used quite often. This value was initially used by the World Bank and World Health Organisation to recommend services that should be provided and also define research priorities in low and middle income countries (World Bank 1993; WHO 1996; Shillcutt et al. 2009). A list of thresholds for different regions has been published by WHO but it has been recognised that additional work needs to be done in order to develop appropriate threshold values (Shillcutt et al. 2009). Recent research has also shown that most studies conducted in low and middle income countries tend to refer to the world health organisation thresholds (Griffiths et al. 2016).

On the whole, health economics and cost-effectiveness is becoming more important and relevant to decision making globally. However, it is clear that this is not the most important concern when resource allocation decisions are made. Other issues such as equity concerns, need and priorities also form an important part of the decision making process (Raftery, 2001). For example, a review of the decisions that were made by NICE showed that cost-effectiveness data was not used as often as it should (Raftery, 2001). Another study comparing drug reimbursement decisions in three countries: Austria, UK and New Zealand found that different factors determined whether a drug will be reimbursed (Raftery, 2008). A summary of how economic evaluation (both trial and model based) is carried out in selected countries that participated in the GRACE project which was used for the empirical work carried out in this thesis is provided in Table 2.1. From this table, it can be seen that the most common type of economic evaluation technique used across countries are the cost-utility analysis and the cost-effectiveness analysis. However, countries like Finland also accept the use of cost-benefit analysis whilst others like Belgium accept cost-minimisation analysis (KCE, 2012). The guidance in terms of perspectives also varies between countries. For example, in the Netherlands, Norway and Spain, the most preferred approach is to conduct the economic analysis from a societal perspective as opposed to the National Health Service/Personal Social Services perspective which is favoured in the UK and healthcare perspectives favoured by countries such as Belgium, and Italy (NICE, 2013; Weinstein et al. 1996; Torrance et al. 1996; Johannesson, 2009; EUnetHTA, 2016). In terms of presenting the results, most countries require the use of ICERs. The notable exception is the case of Germany where the efficiency frontier is used to present the benefits and costs for each intervention (IQWIG, 2015). With the exception of the UK where there is an established cost per QALY threshold of £20,000 per QALY, most countries do not have a fixed threshold

which can be used to judge the cost-effectiveness of interventions. In some cases like Sweden, informal thresholds are used to determine cost-effectiveness (Bolin et al. 2013).

Table 2.1 Summary of economic evaluations in European countries

Country	Methods used for Economic evaluation	Important perspective	ICERs presented	Cost-effectiveness threshold	Discount rates (costs and outcome)	Type of Health System
Belgium (KCE, 2012)	(1) CUA or CMA (choice should be justified) (2)Budget impact analysis	Health care payers perspective	Yes	Explicit threshold is not used (Reference to threshold values from other countries should be avoided)	Costs=3% Outcomes=1.5%	Healthcare is publicly funded and privately provided
Finland (Laakkeiden, 2013)	Primary method is the Cost-utility analysis. Reasons for adopting a method must be stated	Primarily payer perspective	Yes	No explicit threshold	3%	Predominantly funded by taxes
France (Haute Autorite de Sante, 2012)	Cost-utility analysis and/or cost-effectiveness analysis	Collective perspective that is sufficiently broad to take into account all stakeholders concerned by the treatments studied	Yes	No explicit threshold used. Interventions are qualified as efficient if they are non-dominated, without prejudging their acceptability in terms of the public decision-maker's maximum willingness to pay	4%	Social health insurance

Country	Methods used for Economic evaluation	Important perspective	ICERs presented	Cost-effectiveness threshold	Discount rates (costs and outcome)	Type of Health System
				for health gain. Acceptability curves inform decision-makers about the probability that interventions are cost-effective at various cost-effectiveness thresholds		
Germany (IQWIG, 2015)	CEA (Efficiency frontier method)	1. SHI insurants 2. Social insurance 3.Societal	No Efficiency frontier used to present benefits and costs for each intervention	Not applicable	3%	Social insurance system
Hungary (Szende et al. 2002)	CUA, CEA or CMA	Unclear	Yes	No explicit threshold	3.7%	Funded by both tax and social insurance
Italy (Ferre et al. 2014 ; Mencaci et al. 2013)	CUA or CEA	Health care	Yes	No explicit threshold is used in Italy although some authors have suggested that a threshold of €25,000 per QALY could be used	3%	Healthcare is funded by taxes and co-payments by patients

Country	Methods used for Economic evaluation	Important perspective	ICERs presented	Cost-effectiveness threshold	Discount rates (costs and outcome)	Type of Health System
Netherlands (Tan et al. 2012; CVZ, 2008)	CUA, CEA or CMA	Societal	Yes	€20,000 per QALY gained	Cost= 4% Outcomes=1.5%	Predominantly social health insurance
Norway (NOMA, 2012; Burger et al. 2014)	CUA and CEA	Societal	No	Lack of consensus with respect to a single cost-effectiveness threshold in Norway. However, a value of NOK 500,000 ($83,000) per QALY is often quoted	4%	Predominantly funded by taxes
Poland (Sagan et al. 2010; Skoupa et al. 2014)	CUA	Public health care payer and the patient	Yes	Not clear A threshold of 3 times GDP per capita has been quoted in the past	Costs=5% Outcomes=3.5%	Predominantly social health insurance. However, there is also private funding
Slovakia (EUnetHTA, 2016)	CUA, CEA	Health care payers	No Incremental analysis Costs and outcomes presented separately for each intervention	Unclear	5%	Compulsory social health insurance

Country	Methods used for Economic evaluation	Important perspective	ICERs presented	Cost-effectiveness threshold	Discount rates (costs and outcome)	Type of Health System
Slovenia (Albreht et al. 2016 ; Šmit et al. 2012)	CUA,CEA or CMA	Health insurance and societal perspective	Yes	Not clear, However, the Slovenian Health authority have proposed a value of €30,000 per QALY	Unclear	Both publicly and privately funded
Spain (Lopez-Bastida et al. 2010; Vallejo-Torres et al. 2016)	Any of the main economic evaluation techniques (CUA, CMA, CBA, CEA) can be adopted. However, the choice of method should be relevant to the particular question.	Societal perspective and third party NHS perspective. However they should be presented separately in the analysis	Yes	A figure of €30,000 per QALY is often used. However, a recent study suggests that it should lie between 21,000 and 24,000 per QALY	5%	Decentralised system where autonomous districts are responsible for providing healthcare to the local population
Sweden (Bolin et al. 2013; Anell et al. 2012)	CUA	Societal	Yes	No explicit threshold. But an informal threshold of €55,371 has been quoted	3%	Predominantly funded by tax
UK (NICE, 2013)	CUA	NHS/Personal Social Services	Yes	£20,000 to £30,000 per QALY	3.5%	Predominantly funded by tax

* Information presented in the table applies to both trial and model-based economic evaluation studies

2.5 Economic evaluation alongside multinational trials

Economic evaluation alongside multinational trials can be defined as "studies that, by design, contain data or generate findings that have relevance to two or more countries for the purpose of informing healthcare decision-makers about the cost-effectiveness of healthcare technologies in different national settings" (Pang et al. 2002: P 76). From the above definition, it is clear that it involves data collection in more than one country and as a result of this; certain issues, which are mostly related to the differences between countries need to be addressed (Zarate, 2007). The differences between countries, some of which have been mentioned in Chapter 1 and which would be considered later on in this chapter have led to a situation where there is a lot of debate and uncertainty with respect to whether data from all participating countries (pooling) or a subset of countries (splitting) should be used for the health economic analysis. Current reviews of the literature seems to suggest that similar to the approach that is used in clinical studies, the main approach used for the economic analysis is to simply pool data across countries (Rivero-Arias and Gray, 2010). The main argument against pooling economic data is that doing this assumes that all participating countries are homogenous, and that pooling would make the results less useful for local decision makers. On the other hand, it is suggested that if the analysis is not pooled, sample size would be affected and as a result, the economic analysis would not be adequately powered (Reed et al. 2005).

It has been recognised that there is an increasing demand for economic evaluations to be conducted alongside multinational trials but there is little consensus about how such studies should be conducted and reported. As a result of the need to standardize the approach for the conduct of economic evaluations alongside multinational trials, researchers reviewed the

literature and decided on a framework for the conduct of economic evaluations alongside multinational trials based on the source of resource use data, effectiveness data and costs (Reed et al. 2005). These approaches are outlined below:

2.6 Analytical approaches to economic evaluation alongside multinational trials

2.6.1 Fully pooled analysis

An analysis is described as fully pooled when resource use and effectiveness data from all participating countries and centres in a trial are used in the analysis (Reed et al. 2005) (Table 2.2). The approach therefore assumes that participating countries are similar and that there is likely to be no restrictions to the adaptation of the results of the study to a particular country (Reinhold, 2010). The fully pooled approach comprises alternative approaches for applying unit costs:

Fully pooled one country costing approach (FPOC)

The fully pooled one country costing approach uses resource use and effectiveness data from all participating countries but applies unit costs from just one of the participating countries to the data. The main advantage of this approach is that the sample size of the study is maintained. However, a major disadvantage of the approach is that by applying unit costs from just one country to resource use in *all* countries, the theoretical relationship between resource use and unit cost is affected (Reed et al. 2005; Manca et al. 2010).

Fully pooled multicountry costing approach (FPMC)

The fully pooled multicountry costing approach (FPMC) makes use of resource use and effectiveness data from all participating countries/centres and applies the unit costs associated with each resource item from all participating countries to the data. For example, in a multinational trial that recruits participants from UK and France, unit costs from the UK will be applied to UK participants and French costs will be applied to French participants. Like the FPOC approach, this approach has the advantage of maintaining the sample size and unlike the FPOC approach this approach maintains the theoretical relationship between resource use and unit costs. However, a potential disadvantage of the FPMC approach is that applying different price weights to the data from different countries has the effect of skewing the results towards the countries with the higher costs (Reed et al. 2005).

2.6.2 Fully split analysis

An analysis is described as fully split when it makes use of resource use and effectiveness data from one or a subset of countries involved in the trial (Reed et al. 2005). In other words, a subsample of the data is used for the analysis. For example, in a trial where recruitment takes place in five countries (UK, Canada, Spain, USA and Germany), the data on resource use and effectiveness are derived from some (e.g. UK, Spain and Germany) rather than from all five participating countries. Similar to the fully pooled approach described above, this approach has two sub categories:

Fully split one country costing approach (FSOC)

The fully split one country costing approach (FSOC) relies on resource use and effectiveness data from the same group of patients in one or a subset of countries involved in the trial but relies on unit costs from just one of the participating countries (Reed et al. 2005). So from the above example where the subsample was UK, Spain and Germany, it may be decided that unit costs from the UK will be applied to all three countries that were selected from the five

countries. Similar to the FPOC approach described above, the disadvantage of this approach is that the relationship between resource use and unit costs is distorted (Reed et al. 2005).

Fully split multicountry costing approach (FSMC)

The fully split multicountry costing approach (FSMC) relies on resource use and effectiveness data from the same group of patients in one or a subset of patients in one or a subset of countries but relies on unit costs from individual countries (Reed et al. 2005). From the example that was given above, unit costs from all three countries would be applied to participants from each country i.e. UK costs applied to UK resource use, German costs applied to German resource use and Spanish costs applied to Spanish participants. The main disadvantage of the fully split approach is the loss of statistical power that arises from limiting the analysis to one or a subset of countries.

2.6.3 Partially split analysis

A partially split analysis is one that relies on effectiveness data from all participating countries but resource use data from just one or a subset of countries. Similar to the fully pooled and fully split approaches that have been described above, this approach can also be considered as either partially split one country or partially split multicountry based on whether unit costs are obtained from one country or multiple countries. The main problem with this approach is that costs and clinical effects are derived from different patients/samples and as a result, the relationship between costs and outcomes is compromised (Reed et al. 2005). A potential advantage of this approach is that attributes such as costs and resource use, which are normally considered to be less generalisable, can be obtained from similar countries.

Table 2.2 Analytical approaches to the economic evaluation of multinational trials

Effectiveness data	Resource use data	Type of analysis	Costing methodology	Classification
All countries	All countries	Fully pooled	Multicountry	Fully pooled with multicountry costing [a]
			One country	Fully pooled with one country costing [b]
All countries	One or a subset of countries	Partially split	Multicountry	Partially split with multicountry costing [a]
			One country	Partially split with one country costing [b]
One or a subset of countries	One or a subset of countries	Fully split	Multicountry	Fully split with multicountry costing [a]
			One country	Fully split with one country costing [b]

Table adapted from Reed et al. 2005

[a] Multicountry costing occurs when unit costs from each country is used to value resource use in each country. [b] One country costing occurs when costs from just one participating country is applied to all resource use data irrespective of the country.

2.7 Issues in the analysis of multinational trials

2.7.1 Costing

Costing is an integral part of economic evaluation and the process is normally carried out in 3 stages: (1) choosing a perspective, (2) identifying and measuring resource use, and (3) valuing the resource use data (Morris et al. 2012; Gray et al. 2010). The perspective of a costing exercise sets the context for the study and facilitates comparison between studies (Drummond et al. 2015). There are a number of perspectives which are normally based on national recommendations. For example, the NHS/PSS perspective is used in the UK, whilst the societal perspective is recommended in the Netherlands (NICE, 2013; CVZ, 2008). Once the perspective of the analysis has been determined, the next step is to identify resource use data based on the perspective that has been chosen. This would depend on whether a micro (bottom-up) or macro (top down) costing approach is proposed (Gray et al. 2010; Wordsworth et al. 2005). It has been noted that the choice between the two approaches depends on how easy the data collection process is. In practice, most costing studies use a combination of both approaches (Gray et al. 2010).

Collecting resource use data can be undertaken prospectively or retrospectively. Prospectively, medical records or case notes which are used mainly for collecting medical/clinical information can be adapted to include resource use data (Ridyard and Hughes, 2010). Databases in which research nurses fill in information about resources used by patients could also be used. Patient diaries or self-reported questionnaires are another source of resource use data which could be used in the context of a clinical trial (Kennedy et al. 2002; Ridyard and Hughes, 2010). There have been ongoing debates with respect to whether self-report or medical records should be used for obtaining resource use data. Self-report questionnaires are affected by issues such as recall, questionnaire response rate and

item completion rate, which affect the accuracy of information obtained. Medical records on the other hand do not provide information such as time off work and patient incurred costs such as over the counter medication (Kennedy et al. 2002). Nurse led interviews are another approach that has been used to collect resource use data.

Resource use data could be collected retrospectively through published sources such as journal articles and government reports. The use of administrative databases is another approach that is used in countries such as the USA. These databases normally hold data that have been collected routinely. However, the main problem with these databases is that they may not hold accurate data and might not have required resource use data (Frick, 2009). Expert opinion is another potential method of obtaining resource use data in cases where this data is not available (Ridyard and Hughes, 2010). The problem with this approach is that it may involve a lot of value judgements.

To help with the process of resource use and cost measurement, resources such as the DIRUM database, provide an open access resource which stores a repository of questionnaires and methodological papers which can be used by researchers involved in trial based economic evaluations (Ridyard et al. 2012).

In the context of a multinational study, there is the possibility that care provided may differ across countries, i.e. some countries may be more labour-intensive than others. Thus, when collecting resource use in this context, there is a need to account for the differences in health systems and the differences in the provision of care across countries.

The final step is to value the resource use data by applying unit cost data. In most developed countries such as the UK, unit cost data can be obtained from national sources such as the publications on costs by the Personal Social Services Research Unit (Curtis and Burns, 2015).

Within the context of a multinational trial, there is the possibility that sources of unit cost data are not available in some countries such as low income countries, where good quality data is scarce (Schulman et al. 1998; Knapp et al. 2008). In such cases, a combination of methods ranging from contacting local researchers and making assumptions about the unit costs in the country of interest may be adopted (Schulman et al. 1998). Once all unit cost data has been collected, this cost is multiplied by resource use in order to obtain total costs using the formula.

$$Total\ cost\ for\ patient_i = \sum_j Resource_{ij} \times Unit\ cost_j \ldots\ldots\ldots 2.2$$

Where $Resource_{ij}$ is the amount of resource j used by patient i and Unit $cost_j$ is the unit cost for resource j.

Figure 2.1 provides a summary of the process of costing within a multinational study. Specific issues related to costing exercise alongside multinational studies are now considered in the section below.

Some issues with costing in multinational trials

Costing alongside multinational studies is quite a controversial issue and also one of the greatest challenges faced by analysts (Wordsworth et al. 2005; Schulman et al. 1998). The literature on costing alongside multinational trials has focused on a number of areas including research into methods for ensuring the comparability of cost data in multinational studies, and one study compared the top-down and bottom-up approach for collecting cost data and concluded that the bottom up approach is preferred to the top down approach for multinational costing (Wordsworth et al. 2005).

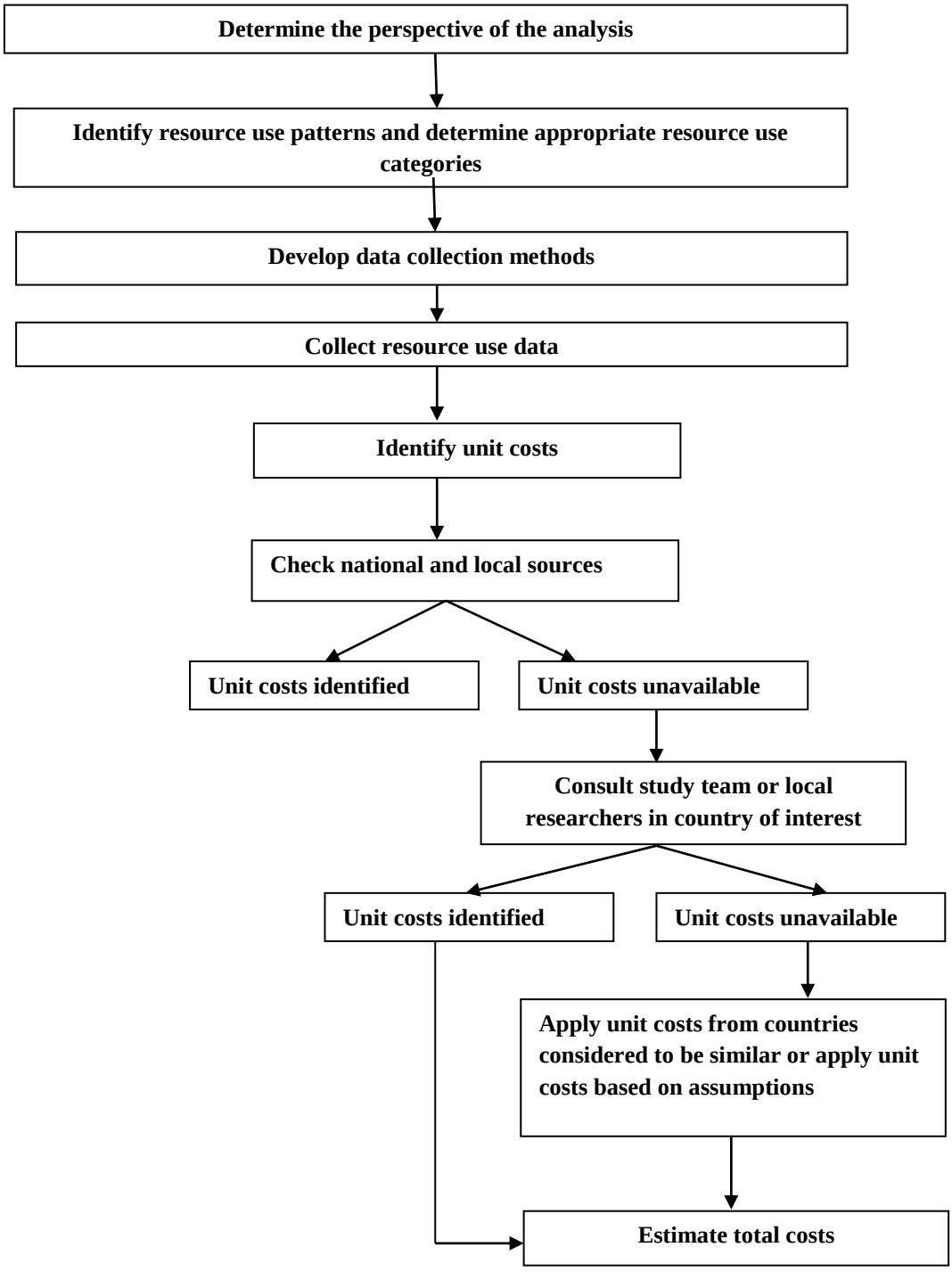
Another issue of importance with costing in multinational trials is the conversion of costs into a single currency particularly when the multicountry costing approach (collecting unit cost data from a number of countries) is used. One approach is to use the official exchange rates to convert costs to a common currency such as the US dollar. However, the problem associated with the use of official exchange rates to convert costs is that they do not necessarily reflect the relative purchasing power of different currencies (Saunders and Marsden, 2014). The acceptable thing to do is to convert the currency using Purchasing power parities (PPPs). The advantage of the PPPs is that they equalize the purchasing power of different currencies rather than provide a reflection of the supply of currencies in the market (Schulman et al 1998). For example, with exchange rates, if a commodity sells for $40 in country X and £150 in country Y, at an exchange rate of 1:10, the good will cost $40 in country X and just £15 in country Y to the consumer in country X. This is because consumers in country X can exchange $15 for £150 (the price of the good in country Y). Consumers in country X will therefore prefer to purchase the good in country Y since it is cheaper. With PPP exchange rates, the exchange rate would be $40 to £150. So it has taken away the disparity. As a result, of this, it is preferable to use PPPs when comparing costs in multinational trial settings (Schulman et al. 1998). One study has suggested that technology specific PPPs which theoretically provide a more robust approach to the comparison of costs in multinational trial settings should be used. However, more work needs to be done in relation to the construction and application of technology specific PPPs (Wordsworth and Ludbrook 2005).

One of the other problematic costing issues in this area is the lack of availability of unit cost data in some countries. This often results in a situation where the one country costing approach is used. However, few studies have explored this problem. One study has however suggested a market basket approach should be used for this purpose. This involves developing indices that reflect the relative cost of a basket of good between countries. For

example, in a trial where unit costs are available for the UK and unavailable for other countries, an index (based on the price of a basket of goods) between the UK and all other countries is developed and used to derive unit costs for all other countries in the trial (Schulman et al. 1998).

Unit costs and resource use data are not always considered to be transferable across countries (Barbieri et al. 2010), and studies have looked at possible ways of ensuring that cost data are more generalisable across countries. Methods that have been proposed in the literature include an analysis of the amount of medical consumption and analysis of costs (Koopmanscap et al. 2001). More recently, literature has focused on establishing whether particular components of costs such as direct and indirect costs are transferable or not (Zhao et al. 2013). Another issue which may affect costing alongside multinational trials is the application of discount rates. There is heterogeneity among countries regarding the discount rate used and as a result, generalisability of study findings may be affected if discount rates used are not relevant to particular settings. For example, in the UK, it is recommended that costs and outcomes are discounted at a rate of 3.5% (1.5% for public health interventions), whilst in France the recommended discount rate is 4% (NICE 2012; NICE, 2013; Haute Autorite de Sante, 2012).

Figure 2.1 Identifying resource use and unit costs in a multinational trial



2.7.2 Outcomes in multinational trials

Within multinational trial settings, outcomes are measured in a similar way to single country trials. With respect to the types of economic evaluations, the literature seems to support the transferability of clinical outcomes such as those which could be used in cost-effectiveness analysis (Barbieri et al. 2010). However, there is some uncertainty with respect to the transferability of health utilities across jurisdictions (Knies et al. 2009; Barbieri et al. 2010; Oddershede and Petersen, 2015). Thus, there are some questions that are raised when a cost-utility analysis is conducted alongside a multinational trial. With respect to outcomes the literature on multinational trials has therefore focused on how utilities should be measured and used. Some studies have focused on comparing utilities obtained from different countries, and their results have shown that differences exist in terms of weights given to various dimensions of the EQ-5D by respondents from different countries. For example, the UK population tend to give greater preference to dimensions such as pain/discomfort and anxiety/depression whilst the Spanish give more preference to dimensions such as mobility and self-care (Badia et al. 2001).

Another key focus of research around utilities in multinational settings have concentrated on the comparison of various value sets/tariffs for obtaining EQ-5D index scores with the results revealing differences in the methods that were used to obtain the various value sets (Norman et al. 2009), however, it is not clear whether the differences between value sets are as a result of cultural or methodological factors (Szende et al. 2007; Knies et al. 2009).

2.7.3 Generalisability and transferability

One question that has remained unanswered in the existing literature on economic evaluation alongside multinational trials is the extent to which the results that are obtained are useful to decision makers in a particular country or jurisdiction. This question becomes pertinent in cases where pooled results from multinational trials are reported. Thus, research in this area

has focused on how best to present results so that they may become more generalisable and transferable (Manca et al. 2010). There are several definitions of generalisability and transferability. In some cases these two terms have been used interchangeably (Mason and Mason 2006). *Generalisability* occurs when the results of a study can be applied to a number of countries without adjustment needed for its interpretation. *Transferability* on the other hand occurs when a study can be adapted to apply to other countries (Drummond et al. 2009). As a result of issues relating to generalisability and transferability, the external validity of results from multinational trials is therefore difficult to determine (Reed, 2012). Some potential threats to generalisability and transferability have been identified in the literature and they include the following:

***Demography and epidemiology of the disease*:** Severity, incidence and prevalence of disease can be important contributing factors to the generalisability of findings from economic evaluations, especially where these factors differ across countries (Koopmanschap et al. 2001, Hughes et al. 2016). In such situations, the overall treatment effect may mask the true treatment effect in a particular country or setting and affect the cost-effectiveness of the intervention. For example, vaccination and screening programs may look more cost-effective in areas where the incidence of a particular disease is high. Therefore, pooling results may lead to misleading conclusions in some settings.

***Clinical practice and conventions*:** Medical practice patterns and health systems differ across countries and this generally affects the extent to which the results of a study are generalisable. Physicians in a particular country may be trained to use adjunct treatments in addition to the health technology under consideration which may cause it to appear more effective and potentially cost-effective. In addition to this, incentives, health systems and regulatory structures in different countries will lead to differences in resource use across countries (Briggs, 2010; Koopmanschap et al. 2001). For example, patients in some parts of Europe

have direct access to secondary care as opposed to going through primary care in countries such as the UK.

Relative and absolute price levels: Absolute prices differ across countries e.g. prices in one country may be above average, whilst those in another country may be below average. This has the effect of making the price of an intervention more expensive in one country and cheaper in another. Differences in relative prices (price of a good in relation to other goods) on the other hand would lead to the substitution of resources from the more expensive to the cheaper resource use which would affect practice patterns across countries and eventually affect cost-effectiveness (Briggs, 2010; Koopmanschap et al. 2001). In addition to this, obtaining unit costs in different currencies affects generalisability and comparability of results from multinational trials (Wordsworth et al. 2005).

In terms of analytical approaches to economic evaluation alongside multinational trials, it has been suggested that the source of unit cost, resource use and effectiveness data will determine how generalisable and transferable the results from the economic analysis is (Reinholdt et al. 2010). Reinholdt and colleagues suggests that if effectiveness, resource use and unit cost data are derived from just one country, the results would be applicable to that particular country. However, as the number of countries from which resource use, effectiveness and cost data are derived from increases, the results from the study become more generalisable to a number of countries but are now less applicable to the country of interest (Table 2.3). A recent study reviewed 27 national pharmacoeconomic guidelines in order to assess the extent to which countries considered clinical and economic data transferable to their countries. The study found that although there was a vast difference in recommendations amongst countries with respect to what data is considered transferable to their local settings, in general, clinical data was considered to be more transferable than economic data (Barbieri et al. 2010). The result from this study therefore suggests that additional research is needed in order to standardize

practice across countries in terms of what data is considered transferable and how to deal with it.

Table 2.3 Generalisability of pooled and split approaches

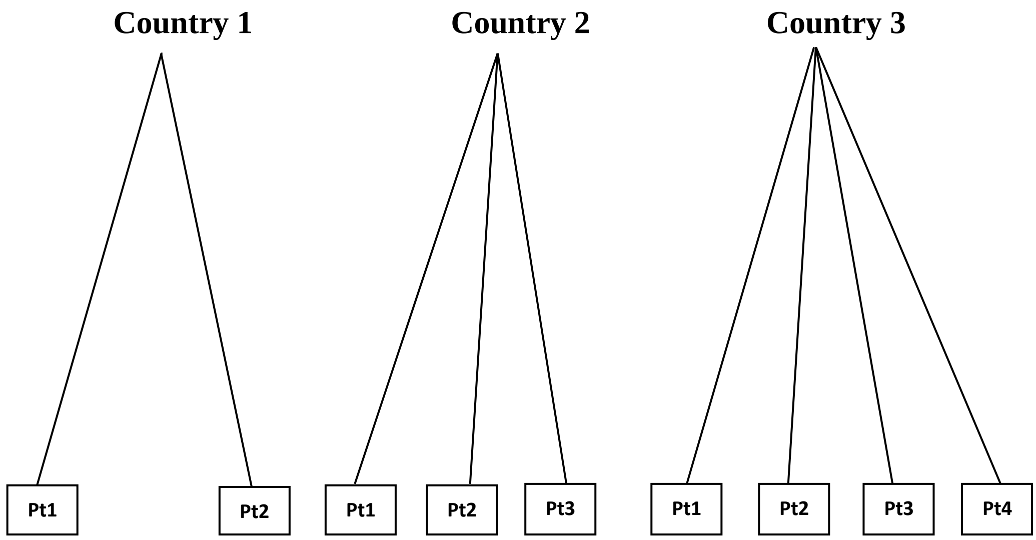
Source of resource use and cost data	Source of effectiveness data: **All countries**	**Subset of countries**	**One country**
All countries	Reduced transferability to one country		
Subset of countries			
One country			Increased transferability to one country

* Table adapted from Reinholdt et al. 2010

2.7.4 Multilevel modelling for the analysis of data from multinational trials

It has been proposed that multilevel models should be used for the analysis of data from multinational trials (Manca et al. 2005; Grieve et al. 2007; Boehler and Lord. 2015). This is mainly due to the fact that data from multinational trials have a hierarchical structure. These hierarchical data structures can take several forms, the most common being the two level structures where lower level units e.g. patients are nested within higher level units e.g. countries (Figure 2.2). Multilevel modelling therefore enables researchers to represent the populations that the subjects of research are located in (Steenbergen and Bradford, 2002).

Figure 2.2 Examples of multilevel data structures



Where Pt means patient

Consider a situation where the cost of treating patients recruited from different countries is being modelled. Using the normal Ordinary least square (OLS) regression equation, we have the following:

$$Y_i = \beta_0 + \beta_1 X_1 + e_i \text{ .... } (2.3)$$

Where Y_i is the cost of patient i, β_0 is the intercept and β_1 is the slope or the change in cost as a result of a 1 unit change in the independent variable X_1 and e_i is error term which represents the departure of the patient i's actual cost from the predicted cost. The error term (e_i) is assumed to have a constant variance and is normally distributed (Rabe-Hesketh and Skrondal, 2008).

Extending equation (2.3) to account for the multilevel structure of the data, results in the following:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + u_{0j} + e_{ij} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.4)$$

An additional subscript j which represents the higher level variable has now been added to the equation. Y_{ij} in equation (2.4) now represents the cost for patient i in country j and the existence of two error terms u_{0j} and e_{ij} is what distinguishes the multilevel model from the ordinary least square model.

Several consequences of ignoring such clustered data have been identified and they include the possibility of generating incorrect inferences about the significance of a variable as a result of overestimated standard errors (Park and Lake, 2005). In addition to this, if analysis is carried out at the patient level, the effect of the variation at the country level would be lost and even if the variation at the country level is modelled, the analyst would have to incorporate separate terms for each country a process which is known to be inefficient (Goldstein, 2011). Furthermore, if the hierarchical data structure is not accounted for, it would not be possible for an analyst to know where most of the variation is occurring e.g. patient level or country level. Multilevel modelling therefore provides researchers with more insights about the data. In the context of economic evaluation alongside multinational trials, it has been recognised that between-location variability in cost-effectiveness between countries

may result due to the correlation in costs/consequences between participants who are located in particular countries and failing to account for the multilevel structure of the data may lead to misleading findings (Manca et al. 2005).

2.7.5 Some unresolved issues

Even though the various analytical approaches to economic evaluation alongside multinational trials have their advantages and disadvantages, as shown in Table 2.3, the literature shows that there is a trade-off between the pooled and split approaches (Reinholdt et al. 2010). It also remains unclear when and how these approaches should be used and also what impact each of the approaches would have on decision making. It has also been suggested that there is the need for research into data collection within multinational trials (Reed et al. 2005). With respect to costs, one of the main issues relates to obtaining unit cost data and it has been suggested that additional work need to be done in order to improve costing exercises alongside multinational studies (Schulman et al. 1998). For outcomes, the main issues arise from the extra-welfarist perspective in particular where instruments such as the EQ-5D are used to obtain QALYs. It has been well established in the literature that there are differences in the various EQ-5D value sets (Pullenayegum et al. 2015). However, the unresolved issue is the impact the choice of tariff has on the results/conclusions from cost-effectiveness analysis and how to choose between value sets in multinational trial settings.

2.8 Summary

The work that has been presented in this chapter focused on outlining the theoretical basis for economic evaluations in general and also looking at specific issues relating to the multinational trial based economic analysis. The characteristics of analytical approaches

based on pooling and splitting the data were explored and an overview of the work related to costing, outcomes, analysis of data and generalisability and transferability were presented. Lastly, some unresolved issues were identified. In the next chapter, this thesis takes a closer look at how economic evaluations alongside multinational trials are conducted in practice by conducting a systematic review of published economic evaluation studies that have been conducted alongside multinational trials.

CHAPTER 3 REVIEW OF THE LITERATURE ON ECONOMIC EVALUATION ALONGSIDE MULTINATIONAL TRIALS

3.1 Introduction

In this Chapter, a review of the published literature around economic evaluation alongside multinational trials is presented. The objective is to review the methods used in practice and the challenges that are typically faced by the researchers who conducted the economic evaluations.

This chapter begins by giving a background to the review, followed by an outline of the methods that were used to search the literature and identify relevant studies. The results of the literature search are presented and lastly a thorough discussion of the main results and findings are presented.

3.2 Background to the review

Several issues relating to the economic analysis of multinational trials such as generalisability and transferability which limit the usefulness of cost-effectiveness results from such studies have already been highlighted. Although efforts appear to have been made to address some of these issues, recent reviews of the literature have revealed wide variation in the way these issues are addressed (Torti et al. 2006, Rivero-Arias and Gray, 2010). National guidelines on

the use of data from multinational trials show vast differences in the data different countries consider to be generalisable or transferable to their own settings (Barbieri et al. 2010). This indicates a lack of proper consensus on how multinational economic evaluations should be carried out. Therefore, there is a need to develop and agree on appropriate methods for conducting and interpreting economic analyses based on multinational trials. This would not only make results more useful to decision makers but also avoid the duplication of work in every country/jurisdiction (Drummond et al. 2005).

The aim of the current systematic review is to assess the methods used by published empirical studies that have conducted economic evaluation alongside multinational trials. This review also explores how the study results have been reported and also outlines the challenges which have been encountered by researchers who conducted the economic evaluation.

The results presented here have been previously published in the following peer reviewed article*: Oppong et al. (2015) Economic evaluation alongside multinational studies: A systematic review of empirical studies PLoS ONE 10(6)*

3.3 Methods

3.3.1 Search strategy

A systematic review was conducted following the guidelines of the Centre for Reviews and Disseminations (CRD, 2009). The following keywords were used in the search: multinational, cost, cost-effectiveness, cost-utility, cost-benefit, multicountry, multicentre, trial, economic evaluation, cross-country. Table 3.1 below gives a summary of the search terms used. The following databases were searched: MEDLINE; EMBASE and the National

Health Service Economic Evaluation Database (NHS EED). The search was limited to the period 2002 to 2012 for pragmatic reasons and in order to capture the most recent studies.

3.3.2 Selection of papers for review

Following an approach used by Roberts and colleagues (Roberts et al. 2002), a three stage process was used to select relevant papers for the review. The stages were as follows: (1) Categorization of studies (II) Further categorization of studies (III) Inclusion and exclusion criteria.

Stage I Categorization of studies

The studies identified through the literature search were classified into the following five groups based on an inspection of the titles and abstracts only.

(A) The study is a multinational/multicountry study and includes a full economic evaluation

(B) The study is a multinational/multicountry study and reports on costs and/or outcomes but is not a full economic evaluation study

(C) The study does not fall clearly into categories (A) or (B) above but could have information which is relevant to the overall objective of this study

(D) The study discusses issues/methodological aspects relating to economic analysis alongside multinational trials

(E) The study is not/does not have any relevance to economic evaluation alongside multinational trials.

The groups were determined through an initial scoping of the literature. Studies that fell into category A were deemed relevant for the review whilst studies that fell into categories B, C, D and E were excluded from the review and no further action was necessary.

Stage II Further categorisation of studies

All studies that fell into category A were further categorized into five sub groups based on a more detailed inspection of the selected articles.

A1. Economic evaluation that reports an Incremental cost-effectiveness ratio (ICER) or net benefits

A2. Economic evaluation in that a comparison of cost and outcomes of two or more interventions is presented thus meeting the definition of an economic evaluation, but for which an ICER or net benefit is not reported.

A3. Methodological study/study protocol

A4. Systematic review

A5. The study is not/does not have any relevance to economic evaluation alongside multinational trials.

Studies that fell into categories A2, A3, A4 and A5 were excluded from the study whilst those that fell into category A1 were taken forward. The above groups were determined in such a way as to choose all the most relevant economic evaluations.

Stage III Inclusion and exclusion criteria

Studies that fell into category A1 were included if the economic evaluation was carried out alongside a randomised or non-randomised multinational clinical trial. Although randomised studies were the main focus, non-randomised studies were also included. Studies were excluded if they were modelling studies or if they did not use patient level data.

The quality of the economic evaluations was not assessed because of the study objectives and the need to include as many studies as possible.

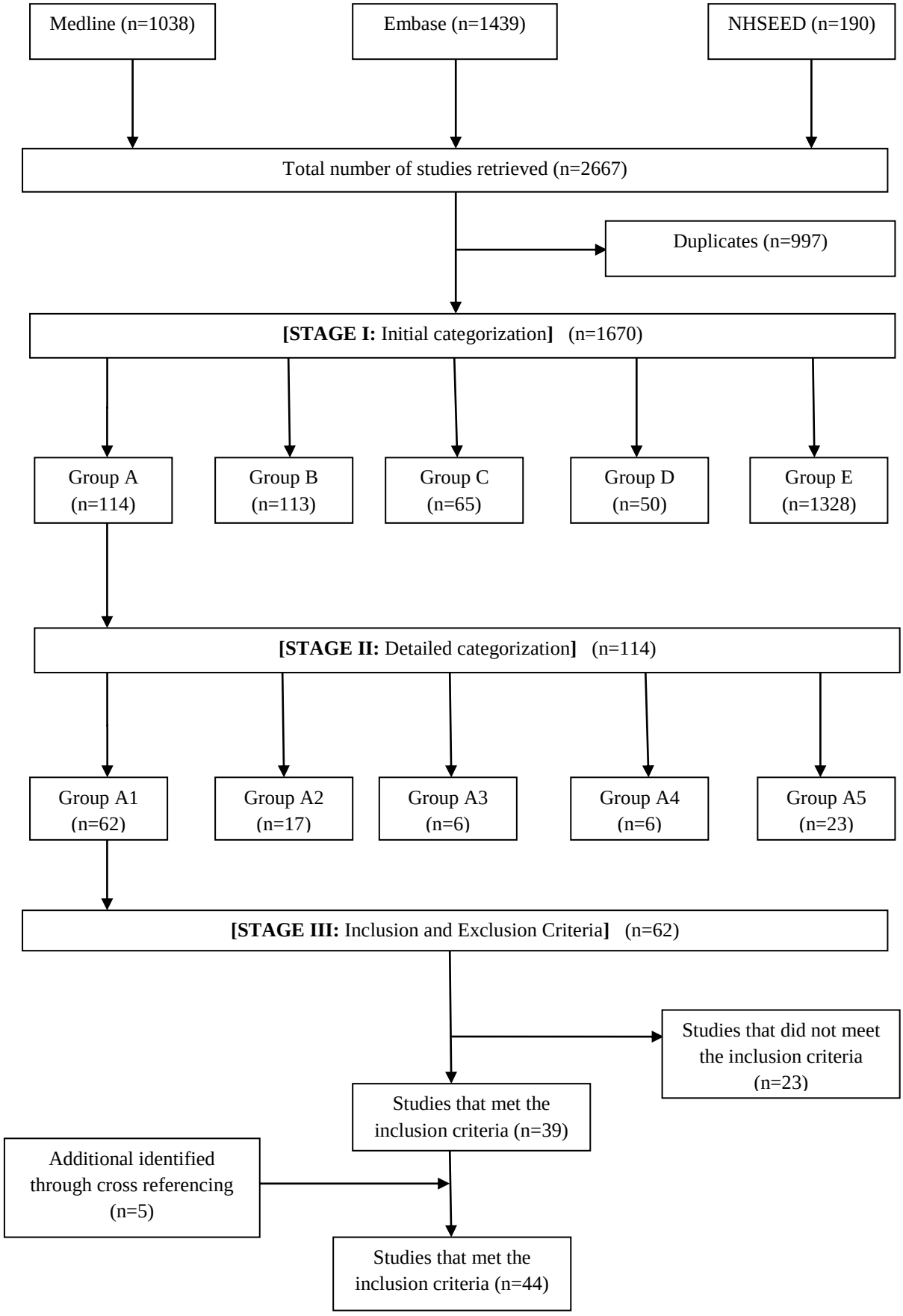
3.3.3 Data extraction

Data were extracted using a predefined data extraction form (Appendix 1), and the following data were extracted from the included studies: Type of economic evaluation, health outcomes considered, study perspective, number of countries included, analytical approach to the economic evaluation used and challenges faced.

3.4 Results

A total of 2667 articles were retrieved through the searches that were carried out with the electronic databases. The number of articles retrieved from each database was 1038, 1439 and 190 for Medline, Embase and NHS EED respectively. The 2667 articles were exported into Reference Manager and after accounting for duplicates, a total of 997 articles were excluded. An inspection of the titles and abstracts of the 1670 remaining articles yielded 114 potentially relevant articles (i.e. articles falling into group A). Out of these studies, a total of 62 studies were classified as economic evaluations that reported an ICER or net benefit (group A1) and the remaining 52 were excluded. Out of the 62 studies that passed through the first and second stages successfully, a total of 39 studies met the inclusion criteria. The remaining 23 studies were excluded mainly because they were model based studies (20 studies) or not relevant (3 studies). In addition to the 39 studies, a further 5 studies were identified through cross referencing. A total of 44 studies passed through the three stages and were included in the final sample of papers to be reviewed (Figure 3.1).

Figure 3.1 Summary of stage 1, stage 2 and stage 3 criteria



3.4.1 Summary of identified studies

The identified studies were published between 2002 and 2011. Approximately 20% were published in 2003. The types of economic evaluation conducted were mainly cost-effectiveness analysis (31 studies) and cost-utility analysis (18 studies) (Table 3.2). Of these, five studies conducted both a cost-effectiveness and a cost-utility analysis (Canoui-Piotrine et al 2009; Glasziou et al 2010; Marcoff et al. 2009; Mittman et al. 2009; Reed et al. 2004). In one study, the cost-utility analysis was performed as a secondary analysis but an incremental cost-effectiveness ratio (ICER) was not estimated (Canoui-Poitrine et al. 2009).

Sixteen studies were related to cardiovascular disease representing a substantial proportion of the included papers. A total of 21 trials that the economic evaluations were based on were placebo controlled trials with a common characteristic being their assessment of drug therapies (Table 3.3). The number of countries included in an individual trial in the analysis ranged from 2 to 48 and approximately 80% of studies included in the review recruited patients from the United Kingdom (UK). The countries that were included in each study are presented in Appendix 2, and a total of 79 countries were identified across all studies.

Table 3.1 Example of search terms used (Medline and Embase search)

1	cost$.mp.
2	cost-utility.mp.
3	cost-benefit.mp. or cost-benefit analysis/
4	cost-effectiveness.mp.
5	economic evaluation$.mp.
6	(multinational or multi-national).mp.
7	(multicentre or multi-centre).mp.
8	cross country.mp.
9	trial$.mp. or exp Clinical Trial/
10	or/1-5
11	or/6-8
12	9 and 10 and 11
13	limit 12 to yr="2002 - 2012"

3.4.2 Classification of studies based on geographical region and gross national income (GNI)

Using World Bank country classifications which are based on the income of the country (The World Bank, 2012); this study identified 38 high income countries, 24 upper middle income countries, 12 lower middle income countries and only 5 low income countries (Appendix 3). A summary of the World Bank classifications is presented in Appendix 3. Only two studies recruited participants from low income countries: one that assessed interventions for preeclampsia (Simon et al. 2006) and another that evaluated interventions for treating malaria (Lubell et al. 2009). Studies that recruited patients from lower middle income countries primarily assessed interventions for chronic obstructive pulmonary disorder (COPD) and asthma (Sullivan et al. 2003, Briggs et al. 2006, Briggs et al. 2010, Lofdahl et al. 2005).

Classification according to geographical region was based on the six populated continents: North America, South America, Europe, Africa, Asia and Australia. Recruitment of participants from European countries was dominant with twelve studies (approximately 29%) classified as European studies (i.e. recruitment of patients took place solely in European countries). Most studies included patients from more than one continent (including Europe). Only one study did not recruit participants from Europe (Mittman et al. 2009).

3.4.3 Health outcomes

The studies that carried out the cost-effectiveness analysis reported general outcomes such as cost per life year gained or used disease specific measures such as cost per cardiovascular event avoided. For the studies that adopted the cost-utility analysis approach, the quality adjusted life year (QALY) was the main outcome that was used; however, various methods were used to estimate QALYs. The main approach was to use responses obtained from the EQ-5D questionnaire and use them to obtain health utilities. Nine studies gave an indication of how they generated EQ-5D index scores (Barchert et al. 2007; Canonica et al. 2007; Fernandez et al. 2005; Manca et al. 2003; Garry et al. 2004; Nasser et al. 2008; Bracco et al 2007; Briggs et al. 2010; Knapp et al. 2008) and in all cases, it was evident that the UK tariff (Dolan, 1997) was used because it was well established (Bracco et al. 2007), recommended (Briggs et al. 2010) and readily available (Knapp et al. 2008). Only one study used the Health Utility Index to obtain QALYs (Mittman et al. 2009). Mapping was another approach used for this purpose; one study (Briggs et al. 2006) used a mapping algorithm to obtain QALYs from the Asthma Quality of Life Questionnaire.

Table 3.2 Summary of studies that met the inclusion criteria

Author/Year	Study aims	Number of countries included	Type of economic analysis	Health outcomes	EQ-5D Value set used	Study perspective	Analytic approach to the economic evaluation used	Country-specific results presented	Adjustments made to account for country variations	Discussed challenges associated with multinational studies
Canoui-Piotrine et al 2009	Assess the cost-effectiveness of sirolimus-eluting stents compared with bare metal stents.	15	Cost-effectiveness analysis and cost-utility analysis	Cost per target vessel revascularization avoided	N/A	Health service perspective	Fully split one-country costing	Yes	No	No
Glasziou et al 2010	Determine the cost-effectiveness of a fixed combination of perindopril and indapamide	20	Cost-effectiveness analysis and cost-utility analysis	Cost per death averted at 4.3 years average follow-up, cost per life year gained and cost per QALY	N/A	Healthcare purchaser perspective	Fully pooled one-country costing	Yes	Yes	Yes
Marcoff et al 2009	Examine the cost-effectiveness of enoxaparin compared with unfractioned heparin as adjunctive therapy for fibrinolysis	48	Cost-effectiveness analysis and cost-utility analysis	Cost per life year gained and cost per QALY gained	N/A	Societal perspective	Fully pooled one-country costing	Yes	Yes Regression approach	Yes
Mittman et al 2009	Assess the cost-effectiveness of cetuximab in metastatic colorectal cancer	2	Cost-effectiveness and cost-utility analysis	Cost per life year gained and cost per QALY gained	N/A	Payer perspective (Canadian government)	Fully pooled one-country costing	No	No	Yes
Reed et al. 2004	Estimate the cost-effectiveness of zoledronic acid versus placebo for dressing skeletal complications in men with prostate cancer	17	Cost-effectiveness analysis and cost-utility analysis	Cost per skeletal complication avoided; cost per patient free of skeletal-related event and cost per QALY	N/A	Societal perspective	Fully pooled multicountry costing	No	Yes through currency conversion	Yes
Simon et al 2006	To assess the cost-effectiveness of using magnesium sulfate to prevent preeclampsia	33	Cost-effectiveness analysis	Cost per case of preeclampsia prevented	N/A	Treatment provider perspective (hospital)	Fully pooled multicountry costing	Yes region-/group-specific cost-effectiveness	Yes through currency conversion and country classification	Yes.
Lubell et al 2009	To explore the cost-effectiveness of artesunate versus quinine for the treatment of severe	4	Cost-effectiveness analysis	Cost per death averted	N/A	Provider perspective	Fully pooled multicountry costing	Yes	Yes	Yes

Author/Year	Study aims	Number of countries included	Type of economic analysis	Health outcomes	EQ-5D Value set used	Study perspective	Analytic approach to the economic evaluation used	Country-specific results presented	Adjustments made to account for country variations	Discussed challenges associated with multinational studies
	malaria									
Sullivan et al. 2003	Estimate the cost-effectiveness analysis of early intervention with budesonide in mild, persistent asthma	32	Cost-effectiveness analysis	Cost per symptom-free day	N/A	Healthcare payer and societal perspective	Fully pooled one-country costing			Yes
Briggs et al 2006	Estimate the cost-effectiveness of a single inhaler versus fluticasone proportionate in aiming for total control in asthma patients	44	Cost-utility analysis	Cost per QALY gained	Mapping	Health service perspective	Fully pooled one-country costing	Yes	Yes Regression approach	Yes
Briggs et al 2010	Inform decision makers about the cost-effectiveness of alternative COPD treatments	42	Cost-utility analysis	Cost per QALY gained	UK tariff	Not clear	Fully split multicountry costing	Yes region-specific	Yes	Yes
Lofdal et al 2005	Compare the healthcare costs and effects of budesonide/formoterol in a single inhaler with those of budesonide and formoterol monotherapies and placebo in patients with COPD	15	Cost-effectiveness analysis	Cost per avoided exacerbation	N/A	Healthcare payer perspective	Fully pooled one-country costing	No	Yes Followed study protocol rigorously in all countries	No
Bachert et al 2007	Assess the cost-effectiveness of grass allergen tablet compared with symptomatic medication for preventing seasonal grass pollen-induced rhinoconjunctivitis	7	Cost-utility analysis	Cost per QALY gained	UK tariff	Societal perspective	Fully split multicountry costing	Yes	No	Yes
Canonica et al 2007	Assess the cost-effectiveness of GRAZAX for preventing grass pollen-induced	8	Cost-utility analysis	Cost per QALY gained	UK tariff	Societal perspective	Fully pooled multicountry costing	Yes	No	No

Author/Year	Study aims	Number of countries included	Type of economic analysis	Health outcomes	EQ-5D Value set used	Study perspective	Analytic approach to the economic evaluation used	Country-specific results presented	Adjustments made to account for country variations	Discussed challenges associated with multinational studies
	rhinoconjunctivitis									
Fernandez et al 2005	Assess the relative cost-effectiveness of escitalopram compared with venlafaxine in patients with major depressive disorder	8	Cost-utility analysis	Cost per QALY gained	UK tariff	Payer perspective	Fully pooled multicountry costing	No	Yes Regression approach	Yes
Manca et al 2003 [40]	Assess the cost-effectiveness of tension-free vaginal tape compared with open burch colposuspension as a primary treatment for urodynamic stress incontinence	2	Cost-utility analysis	Cost per QALY gained	UK tariff	Health service perspective	Fully pooled one-country costing		No	No
Garry et al. 2004	Evaluate the cost-effectiveness of laparoscopic, abdominal and vaginal hysterectomy	2	Cost-utility analysis	Cost per QALY gained	UK	UK NHS perspective	Fully pooled one-country costing	Yes		
Nasser et al. 2008	To assess the cost-effectiveness of GRAZAX in patients with rhinoconjunctivitis and coexisting asthma	8	Cost-utility analysis	Cost per QALY gained	UK tariff	Societal perspective	Fully pooled one-country costing	Yes	No	No
Bracco et al 2007	Assess the cost-effectiveness of tegaserod in treating irritable bowel syndrome	Not stated	Cost-utility analysis	Cost per QALY gained	Appears to be UK tariff	Third-party payer perspective	Fully pooled one-country costing (check)	No	Yes Regression approach	Yes
Knapp et al 2008	Determine the cost-utility of treating schizophrenic patients with olanzapine compared with other antipsychotics	10	Cost-utility analysis	Cost per QALY gained	UK tariff	Health service perspective	Fully pooled one-country costing	No	Yes Regression approach	Yes
Buxton et al 2004	Assess the cost-effectiveness of early intervention with budesonide in mild	32 (Mentioned 8 in paper)	Cost-effectiveness analysis	Cost per symptom free day	N/A	Healthcare payer perspective and societal	Partially split multicountry costing	Yes	Yes Used country-specific costs	Yes

Author/Year	Study aims	Number of countries included	Type of economic analysis	Health outcomes	EQ-5D Value set used	Study perspective	Analytic approach to the economic evaluation used	Country-specific results presented	Adjustments made to account for country variations	Discussed challenges associated with multinational studies
	asthma					perspective				
Rutten Von Molken et al 2007	Assess the cost-effectiveness analysis of roflumilast for treating patients with severe chronic obstructive pulmonary disease	14	Cost-effectiveness analysis	Cost per exacerbation avoided	N/A	Societal and NHS perspectives	Fully pooled one-country costing	No	Yes through currency conversion	Yes
Willan et al 2006	Assess the cost-effectiveness of rivastigmine in patients with Parkinson's disease dementia	12	Cost-utility analysis	Cost per QALY gained	N/A	Societal perspective	Fully pooled multi-country costing		Yes Regression approach	Yes
Radeva et al 2005	Determine the cost-effectiveness of everolimus compared with azathioprine one year after de novo heart transplantation	14	Cost-effectiveness analysis	Cost per additional patient free of efficacy failure	N/A	Societal perspective	Fully pooled multicountry costing	No	Yes Regression approach	No
Edbrooke et al 2011	To assess the implications of intensive care unit triage decisions on patient mortality	7	Cost-effectiveness analysis	Cost per life-year saved and cost per life year	N/A	Not clear	Fully pooled multicountry costing	No	Yes Regression approach	Yes
Lamy et al 2004	Assess the cost-effectiveness of the use of clopidogrel in acute coronary syndromes	28	Cost-effectiveness analysis	Cost per CV death prevented	N/A	Societal perspective	Fully pooled multicountry costing	Yes	Yes Regression approach and event costs	Yes
Drummond et al 2003	Determine the cost-effectiveness of sequential i.v./po moxifloxacin therapy compared with i.v./po co-amoxiclav with or without clarithromycin in treating community-acquired pneumonia	10	Cost-effectiveness analysis	Cost per additional patient cured	N/A	Health service perspective	Fully pooled one country costing	Yes	Yes Regression approach	Yes
Gomes et al. 2010	Assess the cost-effectiveness of general versus local anesthesia for carotid surgery	24	Cost-effectiveness analysis	Cost per event-free day	N/A	Health service and personal social services	Fully pooled one-country costing	Yes	No	Yes

Author/Year	Study aims	Number of countries included	Type of economic analysis	Health outcomes	EQ-5D Value set used	Study perspective	Analytic approach to the economic evaluation used	Country-specific results presented	Adjustments made to account for country variations	Discussed challenges associated with multinational studies
Lorgelly et al 2010	Assess the cost-effectiveness of rosuvastatin treatment in systolic heart failure	21	Cost-effectiveness analysis	Cost per major CV event avoided	N/A	Healthcare perspective	Fully pooled one-country costing	No	Yes Used event cost	Yes
Price et al 2002	Assess the cost-effectiveness of chlorofluorocarbon-free beclomethasone dipropionate in treating chronic asthma	4	Cost-effectiveness analysis	Cost per symptom free day	N/A	Healthcare provider	Fully pooled one-country costing	Yes appeared to be UK	Yes Adjusted resource use	Yes
Weintraub et al 2005	Assess the long-term cost-effectiveness of clopidogrel in patients with acute coronary syndromes	28	Cost-effectiveness analysis	Cost per life year gained	N/A	Societal perspective	Fully polled one-country costing	Yes	No	Yes
Wade et al 2008	Evaluate the cost-effectiveness of escitalopram versus duloxetine in treating major depressive disorder	9	Cost-effectiveness analysis	Change in Sheehan Disability Scale	N/A	Societal perspective	Fully pooled one-country costing	No	Yes Regression approach	Yes
Kolm 2007	Assess the cost-effectiveness of clopidogrel in acute coronary syndromes	28	Cost-effectiveness analysis	Cost per life year gained	N/A	Canadian health system	Fully pooled one-country costing	Yes	Yes	Yes
Jowett et al 2009	Assess the cost-effectiveness of computer-assisted anticoagulant dosage versus manual dosing in patients on long- or short-term oral anticoagulant therapy	13	Cost-effectiveness analysis	Cost per clinical event avoided	N/A	Healthcare perspective	Fully pooled one-country costing	No	No	Yes
Dukhovny et al 2011	Evaluate the cost-effectiveness of caffeine for apnea of prematurity	9	Cost-effectiveness analysis	Survival without bronchopulmonary dysplasia (BPD) or neurodevelopmental impairment (NDI)	N/A	Third-party payer perspective	Fully pooled one-country costing	No	Yes Regression approach	Yes
Annemans et al 2003	Assess the cost-effectiveness of	4	Cost-effectiveness	Cost per life year saved	N/A	Healthcare payer	Fully pooled multicountry	Yes	No	

Author/Year	Study aims	Number of countries included	Type of economic analysis	Health outcomes	EQ-5D Value set used	Study perspective	Analytic approach to the economic evaluation used	Country-specific results presented	Adjustments made to account for country variations	Discussed challenges associated with multinational studies
	recombinant urate oxidase in hematological cancer patients		analysis				costing			
Aspelin et al 2005	Assess the cost-effectiveness of iodixanol in patients at high risk of contrast-induced nephropathy	5	Cost-effectiveness analysis	Cost per adverse drug reaction avoided	N/A	Hospital perspective	Fully pooled one-country costing	Yes		No
Bakhai et al. 2003	Evaluate the cost-effectiveness of coronary stenting and abciximab for patients with acute myocardial infarction	9	Cost-utility analysis	Cost per QALY gained	N/A	Third-party payer perspective	Fully split one-country costing	Yes	No	No
Brown et al. 2003	Establish the cost-effectiveness of eptifibatide treatment for acute coronary syndrome patients	28	Cost-effectiveness analysis	Cost per life year gained	N/A		Fully split one-country costing	Yes	No	No
Janzon et al 2003	Assess the cost-effectiveness of extended treatment with low molecular weight heparin (dalteparin) in unstable coronary artery disease	3	Cost-effectiveness analysis	Cost per avoided death or myocardial infarction	N/A	Healthcare provider perspective	Fully pooled one-country costing	No	Yes Tested the impact of price differences between countries	No
Lamy et al 2003	Assess the cost implication of using ramipril in high-risk patients based on the heart outcomes prevention evaluation (HOPE) study	19	Cost-effectiveness analysis	Cost per primary event saved	N/A	Third-party payer perspective	Fully pooled one-country costing	Yes	No	No
Lindgren et al. 2005	Assess the cost-effectiveness of formoterol and salbutamol in patients with asthma	24	Cost-effectiveness analysis	Cost per avoided severe exacerbation	N/A	Healthcare payer perspective	Fully pooled multicountry costing	Yes	No	No
Martin et al 2003	Determine the cost-effectiveness of epoetin-	15	Cost-utility analysis	Cost per QALY gained	N/A	Health service	Fully pooled one-country	No	No	No

Author/Year	Study aims	Number of countries included	Type of economic analysis	Health outcomes	EQ-5D Value set used	Study perspective	Analytic approach to the economic evaluation used	Country-specific results presented	Adjustments made to account for country variations	Discussed challenges associated with multinational studies
	Alfa versus placebo in stage IV breast cancer.					perspective	costing (Not clear)			
Reed et al 2004	Assess the cost-effectiveness of valsartan in patients with chronic heart failure	16	Cost-effectiveness analysis	Cost per life year saved	N/A	Societal perspective	Fully pooled multicountry costing	No	Yes Used country-specific costing and other approaches	Yes
Welsch et al 2009	Cost-effectiveness of enoxaparin compared with unfractionated heparin in ST elevation myocardial infarction patients	48	Cost-effectiveness analysis	Cost per life year gained	NA		Fully pooled one-country costing	Yes	No	Yes

I A fully pooled analysis is a study that relies on resource use and effectiveness data from all participating countries II A fully split analysis is one that relies on resource use and effectiveness from one or a subset of countries. III Partially split analysis relies on effectiveness data from all participating countries but relies on resource use data from one or a subset of countries. IV One-country costing applies the unit cost from one country V Multicountry costing applies unit costs from two or more participating countries.

3.4.4 Costing and study perspective

Costing methodology varied across studies. Twenty eight studies applied unit costs from just one participating country to the data (one country costing) whilst the others applied unit costs from all participating or a subset of countries to the data (multi country costing) (Table 3.2). The average number (range) of countries per study was 17 (2 to 48) and 16 (4 to 42) for studies that adopted the one-country and multicountry approaches, respectively. One reason for adopting a one-country costing approach was the availability of good-quality data in countries such as the UK (Knapp et al. 2008). Most studies presented results from one perspective (health service/healthcare or societal) (Table 3.2), although three adopted multiple perspectives for the purpose of comparison (Sullivan et al. 2003, Buxton et al. 2004, Rutten-van Molken et al. 2007). The results obtained from the different perspectives were comparable (Sullivan et al. 2003, Rutten-van Molken et al. 2007), although one study had results that were sensitive to the perspective adopted (Buxton et al. 2004). In terms of what was considered societal costs, most studies included productivity losses using human capital (Sullivan et al. 2003, Bachert et al. 2007, Canonica et al. 2007, Nasser et al. 2008, Buxton et al. 2007) or friction costs approaches (Rutten-van Molken et al. 2007). One study included caregiver time (Willan et al. 2006), whereas others were not explicit about what was included in terms of societal costs (Table 3.2).

In terms of the approaches to multi country costing, the level of detail given about the sources of unit costs varied from simply stating that official tariffs and retail prices in each country had been used (Bachert et al. 2007) to providing detailed references of each country's unit costs (Canoui-Poitrine et al. 2009, Buxton et al. 2004, Willan et al. 2006).

In most cases, it was unclear how costs had been obtained (Table 3.2). As noted in Chapter 2, one of the problems associated with costing alongside multinational studies is the lack of unit cost data and one approach to costing when unit costs were unavailable was to assume that countries were similar in terms of geographic proximity and level of development and apply the mean cost from countries that were assumed to be similar to the countries for which costs were not available (Simon et al. 2006, Radeva et al. 2005). In contrast, the market basket approach (described briefly in Chapter 2), which involves developing an index that reflects the relative costs of a basket of resources used in a pair of countries (Schulman et al. 1998), was used in two studies (Reed et al. 2004, Radeva et al. 2005). Other approaches included using recognized international databases such as the WHO-CHOICE database (Lubell et al. 2009), contacting local health economists and researchers through surveys that elicited unit cost information (Reed et al. 2004, Radeva et al. 2005) and the top-down/macro-costing approach, which considers costs at an aggregate level (Lamy et al. 2004). This approach has been shown to be effective in cases when obtaining unit costs is not feasible (Morris et al. 2007). Two studies used a combination of methods to obtain unit costs. One study developed a survey which was sent to local health economists in participating countries and also used the market basket approach in cases where costs were not available (Reed et al. 2004). The second study, contacted local health economists in some participating countries, made assumptions about countries being similar and used the market basket approach. (Radeva et al. 2005).

In terms of presenting costs, the most common currencies used were the US dollar, the Euro and the UK pound, with one study (Drummond et al. 2003) presenting its results using more than one currency (Table 3.3).

Table 3.3 Specific characteristics of studies included in the review

Author	Placebo controlled trial	Provided sources of unit costs in each country	Currency used			
			Pounds	Euro	US dollar	Other
Canoui-Piotrine et al 2009	×	×		✓		
Glasziou et al 2010	✓	×				✓
Marcoff et al 2009	×	×			✓	
Mittman et al 2009	×	×			✓	
Reed et al 2004	✓	×			✓	
Simon et al 2006	✓	×			✓	
Lubell et al 2009	×	×			✓	
Sullivan et al. 2003	×	×			✓	
Briggs et al 2006	×	×	✓			
Briggs et al 2010	✓	×			✓	
Lofdal et al 2005	✓	×		✓		
Bachert et al 2007	✓	✓		✓		
Canonica et al 2007	✓	✓		✓		
Fernandez et al 2005	×	×		✓		
Manca et al 2003	×	×	✓			
Garry et al. 2004	×	×	✓			
Nasser et al. 2008	✓	×	✓			
Bracco et al 2007	✓	×		✓		
Knapp et al 2008	×	×	✓			
Buxton et al 2004	✓	✓			✓	
Rutten Von Molken et al 2007	✓	×		✓		
Willan et al 2006	✓	✓	✓			✓
Radeva et al 2005	×	×			✓	
Edbrooke et al 2011	×	×		✓		
Lamy et al 2004	✓	×	✓	✓	✓	✓
Drummond et al 2003	×	×		✓		
Gomes et al. 2010	×	×	✓			
Lorgelly et al 2010	×	×	✓			

Author	Placebo controlled trial	Provided sources of unit costs in each country	Currency used			
			Pounds	Euro	US dollar	Other
Price et al 2002	×	×	✓			
Weintraub et al 2005	✓	×			✓	
Wade et al 2008	×	×	✓			
Kolm 2007	✓	×			✓	
Jowett et al 2009	×	×		✓		
Dukhovny et al 2011	✓	×			✓	
Annemans et al 2003	×	×		✓		
Aspelin et al 2005	×	×		✓		
Bakhai et al. 2003	×	×			✓	
Brown et al. 2003	×	×				✓
Janzon et al 2003	✓	×	✓			✓
Lamy et al 2003	✓	×			✓	
Lindgren et al. 2005	×	×		✓		
Martin et al 2003	✓	×	✓			
Reed et al. 2004	✓	×			✓	
Welsch et al 2009	✓	×			✓	

3.4.5 Analytical approach to economic evaluation

Based on a well-known classification system for economic evaluation alongside multinational trials which was proposed at a consensus workshop and which has been described in Chapter 2 (Reed et al. 2005), 26 studies representing approximately 59% of the total sample were classified as fully pooled one country costing studies and 13 were fully pooled multi country costing studies. Four studies adopted the fully split approach with 50% of these using the one country costing approach. Just one study was classified as a partially split multi country costing study (Table 3.2). Three studies justified why they pooled data across all participating countries. Two of the studies made mention of the fact that the sample size in some of the participating countries was too small (Canonica et al. 2007; Bachert et al. 2007). The third study went one step further to test whether it is appropriate to pool data across countries (Drummond et al. 2003).

3.4.6 Methods for addressing the multinational nature of the data and ensuring the generalisability and transferability of results

Reporting of country specific results

A total of 25 studies reported country or regional specific results in some form. Of these studies, 15 applied unit costs from just one participating country, whilst the rest used costs from multiple countries. All other studies reported general results that were not applied to any specific country. In most cases, it was not clearly stated that country specific cost-effectiveness were being estimated. In terms of obtaining country specific cost-effectiveness estimates, methods differed. Two studies used subgroup analysis within sensitivity analysis to estimate ICERs using only data from the country of interest (Glasziou et al. 2010, Gomes et al. 2010), and in both cases, the results were similar to

the main (pooled) analysis. A third study ignored data from all other countries and used data from only the country of interest (Canoui-Poitrine et al. 2009). Empirical Bayesian shrinkage, a method that involves borrowing strength from the overall trial to estimate country-specific cost-effectiveness (Manca et al. 2007), was used by only one study; however, the authors did not present the country-specific estimates (Marcoff et al. 2009). The simplest approach was to state that the perspective of the analysis was related to a particular country and to apply unit costs from that country to the trial-wide data (Briggs et al. 2006, Lofdahl et al. 2005, Bachert et al. 2007, Buxton et al. 2004, Drummond et al. 2003). With regard to reporting the country-specific results, one study (Lamy et al. 2004) reported the cost-effectiveness results in the country's own currency, whereas other studies presented their results in currencies such as US dollars or Euros.

3.4.7 Methodological approaches to dealing with the multinational trials

A number of studies made attempts to deal with the multinational nature of the data using various methods, the most common being regression methods.

Regression methods

Multilevel modelling was used in three studies to account for the clustered nature of the data (Marcoff et al. 2009, Radeva et al. 2005, Edbrooke et al. 2011). Other regression approaches such as controlling for country when estimating outcomes such as the QALY (Bracco et al. 2007), adjusting for length of stay and costs within countries (Lamy et al. 2004) and including interaction terms and country dummy variables (Briggs et al. 2006, Fernandez et al. 2005, Knapp et al. 2008, Drummond et al. 2003) were also used. In one of the studies, the authors went further to test whether the country dummy variables were significant (Fernandez et al. 2005).

Other approaches

Event rather than daily costs were used to eliminate effects such as differences in lengths of stay across countries (Lamy et al. 2004, Lorgelly et al. 2010). Close adherence to the study protocol (Grieve et al. 2005) was also used to eliminate differences in practice patterns and resource use in different countries (Lofdahl et al. 2005). Resource use and costs were also adjusted to account for the differences between countries. One study adjusted resource use to reflect UK Department of Health asthma guidelines (Price et al. 2002).

3.4.8 Challenges associated with the economic evaluation of multinational trials

One of the aims of this study was to identify challenges associated with conducting economic analysis alongside multinational trials and potential challenges were discussed in 29 studies (Table 3.2), including:

Differences between countries

It was noted that there are numerous differences between countries but no accepted guidance on how to account for them (Briggs et al. 2010, Buxton et al. 2004, Gomes et al. 2010). These differences include: differences in resource use, prices, health systems and practice patterns (Table 3.2). Estimating country-specific cost-effectiveness was another area in which there are no established guidelines (Briggs et al. 2010, Buxton et al. 2004). One study acknowledged this and outlined the advantages and disadvantages of some of the approaches that had been suggested in the literature (Briggs et al. 2010).

Sample size and lack of data

Sample size problems were mentioned by some researchers, who noted that uneven recruitment across countries could potentially lead to unreliable cost-effectiveness estimates, especially in cases in which pooling data across all countries is not an option

(Simon et al 2006, Lubell et al. 2009, Briggs et al. 2010, Bachert et al. 2007, Buxton et al. 2004). The lack of country-specific price weights/costs and the challenges associated with collecting data in multicountry studies were also highlighted in some studies (Radeva et al. 2005, Wade et al. 2008, Jowett et al. 2009). Most often, the researchers conducting the economic analysis were based in one country and were unlikely to know the sources of unit costs in other countries. In addition, there is also a lack of good-quality data in some participating countries, particularly in developing countries (Reed et al. 2004, Knapp et al. 2008, Dukhovny et al. 2011). One study was aware of the advantage of using country-specific price weights but went on to use price weights from only one country (Dukhovny et al. 2011).

Additional challenges

The cost-effectiveness threshold, which represents society's willingness to pay for an additional unit of benefit, is often used to determine whether an intervention is cost-effective (Appleby et al. 2007). However, with regard to analyzing multinational trials, researchers are faced with the problem of how to determine and choose the appropriate threshold (Mittmann et al. 2009, Briggs et al. 2006, Bracco et al. 2007). One study adopted a threshold of €50,000 per QALY but stated that the decision was based on what other studies had done in the past (Bracco et al. 2007). Another important issue relates to the generalisability of study findings. Two studies noted that owing to the multinational nature of the data, decision makers in various countries might face problems with making judgments about the cost-effectiveness of interventions in their own country/jurisdiction (Reed et al. 2004, Briggs et al. 2010). Finally, only one study mentioned the choice of the EQ-5D tariff as a challenge (Knapp et al. 2008).

3.4.9 Assessment of studies based on the Euroheed transferability checklist

A 16 point transferability check list was used to assess the quality of the economic evaluations that were identified in this study. This checklist was originally developed to assess the potential for studies to be generalisable and transferable (Boulenger 2005; Nixon et al. 2009). The results showed that most studies scored higher than 70% with this checklist. The 16 points are presented in Appendix 4. The highest scores were obtained by 2 studies (Drummond et al. 2003 and Wade et al. 2008). A summary of the score obtained from the checklist is provided in Appendix 4.

3.5 Discussion

3.5.1 Summary of main findings

The systematic review reported in this chapter has assessed published economic evaluations that were conducted alongside multinational trials. The results indicate that most studies applied costs from one country but resource use from all countries, possibly owing to a lack of cost data in some countries or to the fact that researchers sought to inform decisions in a particular country. However, of the studies that reported results from a single country, 50% of them applied one-country costing. The major problem that has been associated with this approach is the possibility of overestimating or underestimating costs (Reed et al. 2005, Manca et al. 2010, Glick et al. 2007).

Most studies did not give reasons for having pooled resource use and effectiveness data; although it can be inferred that increasing sample size is a possible motivation for this. One study did test for heterogeneity and homogeneity before pooling data (Drummond et al. 2003). With regard to pooling resource use, unless the study protocol is followed rigidly, issues related to practice patterns across countries could potentially affect the analysis (Reed et al. 2005). However, it should be noted that although protocols have the potential to reduce differences in treatment patterns across countries, they do not necessarily dictate all care provided.

The UK tariff was used in all studies that used the EQ-5D questionnaire to elicit information on health-related quality of life, and although its widespread use can be attributed to its availability (Briggs et al. 2010, Bracco et al. 2007, Knapp et al. 2008), it is also possible that other tariffs such as the EU tariff, which was derived from 6 countries, were not used because they are based on the visual analogue scale (VAS). Some researchers believe that the VAS should not be used in resource allocation decisions because the values obtained are not considered to be utilities (Torrance, 1986, King et al. 2005, Sakthong et al. 2008). In addition to this, most of the studies that used the UK tariff over the EU tariff were published after 2003, the year the EU tariff was published. This supports the findings from other research papers that the UK tariff is most often used (Sakthong et al. 2008). The choice of the EQ-5D tariff is important because different tariffs could lead to conflicting results (Bernert et al. 2009), and the EuroQol group's current guidance states that the most relevant should be used (Szende et al. 2007). However when the study is multinational it is difficult to determine the most relevant tariff and thus there is a need for further research. A recent study has suggested

that researchers explore the potential for different results using all appropriate tariffs within sensitivity analyses (Oppong et al. 2013b).

A number of studies made some form of adjustment to the data to account for the multinational nature; however, the methods used varied, indicating that methods have not been standardized in this area. With regard to studies that looked at country-specific results, only one study explicitly stated that the reason for doing this was the important role of health economics in policy making (Buxton et al. 2004). The most common method of obtaining country-specific estimates was fully pooled one-country costing. A recent study outlined the various methods that have been used to estimate country-specific cost-effectiveness and concluded that Bayesian methods appeared to be the most appropriate for this purpose (Manca et al. 2010); however, only one study in this review used it, possibly because of the complexity associated with implementing it (Marcoff et al. 2009). This method has been challenged because it assumes that differences between countries are random, whereas in reality, these differences are systematic (Manca et al. 2005).

The multinational nature of the data was acknowledged by most studies, but not all listed the countries that were included, and some merely reported the number of countries in the trial. This may be attributable to word limits imposed by journals. In most cases, it was not clear whether the study was attempting to estimate general or country-specific results, primarily as a result of inadequate reporting. With respect to unit costs, the study found that the sources of the costs were not stated in most cases. This is of great concern because this information would enable researchers and decision makers to judge the validity of the study results and whether it is applicable to their own settings and also

help other researchers identify unit cost sources. It is therefore advisable that future multinational studies include sources of unit costs, and if assumptions about the unit costs were made, this should also be made explicit. This review also found that recruitment is biased towards developed countries, which may reflect the difficulties associated with recruiting patients and the lack of high-quality data in low-income countries (Knapp et al. 2008).

Other reviews have looked at economic evaluations alongside multinational trials. One study found that reporting on economic evaluations of multinational trials is inadequate (Riviero-Arias and Gray 2010), another found that methods of analysis differed between studies (Torti et al. 2006) and a recent review reported that the uptake of the more complicated methods for estimating country-specific cost-effectiveness is slow (Vermer and Rutten van Molken, 2013).

3.5.2 Strengths and limitations of the study

The main limitation of this study is the broad nature of the research question, and it is therefore possible that some articles were missed. However, the best attempt was made to identify all possible studies by developing the search strategy with advice from an information specialist. The study was also limited to the period between 2002 to 2012 and was not updated because some of the results presented here formed a basis for some of the work that was carried out in subsequent chapters. The key strength of this study is that it documented the challenges that have been reported by researchers who have conducted economic evaluations of clinical trials, and no other systematic review of multinational trials has done this.

3.5.3 Implications for future research

The most frequently mentioned challenge was the differences between countries, which could possibly affect the generalisability of study findings. Most clinical results from multinational trials are generalisable to the countries that participated in the study. However, results from economic evaluations are not easily generalisable (Barbieri et al. 2010) because there are differences in economic circumstances and differences in health systems across various countries. Hence, there is the need to consider these issues when countries are being included in trials. However, the requirements for economic evaluation/analysis are not given prominence when countries are being chosen for inclusion in multinational trials, and country selection is based on factors such as convenience (Manca et al. 2010). Recent research has looked at selecting centres for multi-centre clinical trials (Gheorghe et al. 2013), but this research needs to be extended to selecting countries in multinational trials as well because the countries included in a study could potentially determine the extent to which the study results are generalisable. In addition, a very important finding is that different methods were used by different studies for costing and addressing differences between countries. This is an indication that guidance similar to that which has been developed for standard economic evaluations needs to be developed.

A possible solution to the problem of generalisability and transferability is the use of checklists to ensure that the results meet the required standards (Heyland et al. 1996, Spath et al. 1999, Welte et al. 2004, Antonanzas et al. 2009). However, a possible limitation is the fact that individual items on checklists are sometimes equally weighted (Nixon et al. 2009). Another suggestion is for researchers to conduct economic

evaluations using multiple perspectives. For example, the results of a study that considers both a health service and societal perspective may be useful for decision making in both the UK and the Netherlands. There is evidence from this study that most researchers are aware of some of the issues surrounding economic evaluation alongside multinational trials, but they did not offer solutions to these challenges in most cases. Researchers should therefore endeavor to document the challenges they face to guide future research. The main challenge that was identified was how to address the differences between countries, which could be attributed to a lack of guidance and consensus on many aspects such as how to estimate country-specific cost-effectiveness. Future research should therefore focus on reaching a consensus about how to address the challenges associated with multinational trials.

3.6 Conclusion

Despite the difficulties associated with multinational studies, their frequency will increase (Pang, 2002). It is clear that conducting an economic evaluation in every country/jurisdiction is not feasible or efficient, and decision makers are likely to have to resort in some cases to considering results from other countries/jurisdictions, to inform their local decision making, despite the obvious limitations. Conducting economic evaluations alongside multinational trials is not trivial, and there should be a conscious effort by all stakeholders to constantly improve methodology in this area. It is suggested that guidelines be developed to aid in using a consistent approach in this area, and this should be based on understanding the challenges associated with multinational trials and

comparing alternative approaches. The guidelines should also be focused on ensuring that results can be useful to decision makers in individual countries.

3.7 Summary

The work presented in this chapter focused on systematically reviewing published economic evaluations alongside multinational trials. The review yielded a total of 44 relevant studies which were assessed in terms of methodological approach to costing, outcomes, economic evaluation in general, how the multinational nature of the trial was dealt with and also whether the study reported any challenges. The results of the review showed that there was a lack of consistency with respect to many aspects of the analysis such as methods used for obtaining unit costs and dealing with the multinational nature of the data. Most studies applied the fully pooled one country costing approach and the most common challenge was how to deal with the differences between countries. The next chapter presents a summary of the case studies from the Genomics to combat Resistance against Antibiotics in Community-acquired LRTI in Europe (GRACE) project. These are used to explore the various questions posed in this PhD thesis.

CHAPTER 4 SUMMARY OF THE CASE STUDIES

4.1 Introduction

This chapter provides an overview of the three case studies from the *Genomics to combat Resistance against Antibiotics in Community-acquired lower respiratory tract infections (LRTI) in Europe (GRACE) project* that were used to carry out the empirical work. The chapter opens with an examination of the link between acute cough/lower respiratory tract infections and antimicrobial resistance, followed by an overview of the various aspects of the GRACE project. Finally, a description of the specific case studies that were used for this research work are presented.

4.2 Lower respiratory tract infections and antimicrobial resistance

The class of diseases known as acute cough and lower respiratory tract infections (acute cough/LRTI) are known to account for approximately 10% of mortality and morbidity worldwide and costs as much as €10.1 billion annually (Ball et al. 2002; Welte et al. 2010). According to a report by the World Health Organisation published in 2002, the top five respiratory diseases are responsible for 17.4% of all deaths and 13.3% of all Disability Adjusted Life Years (WHO, 2002). Coughs, phlegm production, chest pains and fever are some of the most common symptoms of lower respiratory tract infections

(SIGN, 2014) and some common disease types of LRTIs include pneumonia which is characterised by the inflammation of the lungs and bronchitis which is characterised by the inflammation of the bronchi. The most common management option for lower respiratory tract infection is antibiotic therapy and evidence suggests that up to three quarters of patients who present to their general practitioners with symptoms of acute lower respiratory tract infection are given antibiotic prescriptions (Macfarlene et al. 1997; Verheij et al. 1989). However, this form of therapy may lead to potential problems in the long run. Recent research has shown that there is a link between antibiotic consumption and antibiotic resistance and this is evidenced by the persistent rise in antibiotic resistance among major respiratory pathogens such as pneumonia. For example, there has been a rise in macrolide and penicillin resistance in *S. pneumonia* (Ball et al. 2002; Malhotra et al. 2007; Bruyndonckx et al. 2015). Most cases of lower respiratory tract infections can either be viral or bacterial and in most cases, they are self limiting, therefore antibiotics need to be prescribed appropriately (i.e. given to those with bacterial infections), since antibiotics are not effective against viral infections. If this is not achieved, the continual inappropriate use of antibiotics for lower respiratory tract infections could speed up the development of antibiotic resistance even further.

Antimicrobial resistance is seen as one of the major public health concerns in the world today, with governments and international bodies increasingly focusing their attention on developing strategies to curb the problem. In the UK, the Chief Medical Officer has compared the threat of antimicrobial resistance to the threat of terrorism and has noted that urgent attention is needed to deal with the crisis (Annual Report of the Chief Medical Officer, 2011) and as a result of this, a five year strategy to deal with the problem of

antimicrobial resistance was put in place (Department of Health accessed on 27th of November, 2015). It has also been noted that the problem of resistance to antibiotics should be of concern both nationally and internationally and efforts to curb this public health issue should not only be at the national level but also at the international level. An example of such an international initiative is the *Genomics to combat Resistance against Antibiotics in Community-acquired LRTI in* Europe (GRACE) project (GRACE accessed on 27th November 2015) which will be described in greater detail in the next section. As the problem of antibiotic resistance is becoming increasingly more important, there are a number of international efforts such as Antibiotic Action (http://antibiotic-action.com/), Biotechs from Europe Innovating in anti-microbial resistance (BEAM Alliance) (http://beam-alliance.eu/), Alliance for the prudent use of antibiotics (APUA) (http://www.tufts.edu/med/apua/) and ReAct (http://www.reactgroup.org/) which have been put in place to combat the problem of antibiotic resistance.

4.3 The Genomics to combat Resistance against Antibiotics in Community-acquired LRTI in Europe (GRACE) project

The GRACE project which was funded by the European Union (EU) FP6 programme is a network of excellence which aims to "integrate and coordinate the activities of physicians and scientists from many institutions in 14 European countries to combat antibiotic resistance in community-acquired lower respiratory tract infections" (GRACE accessed on 27th November 2015). The countries that took part in the study were The United

Kingdom (England and Wales), The Netherlands, Germany, Belgium, Italy, Hungary, Slovakia, France, Sweden, Norway, Finland, Spain and Poland. A list of participating institutions has been provided in Appendix 5. The specific aims and objectives of the project include

(1) Studying the major community-acquired lower respiratory tract infections, which are the leading reasons for antibiotic prescribing and development and assessment of practice based interventions which could be used in reducing inappropriate antibiotic prescribing in patients with community-acquired LRTI.

(2) Evaluating the cost-effectiveness of the interventions developed in the observational and intervention studies.

(3) To develop education and training support to disseminate awareness and knowledge relevant to antibiotic resistance.

The full list of the aims and objectives of the GRACE project can be found on the project website (GRACE accessed on 27th November, 2015). The GRACE project was divided into four different platforms, each of which comprised a number of workpackages. The four platforms are:

(1) GRACE-COMIT which was the platform for coordination, management and information technology (workpackages 1 and 2).

(2) GRACE-TECH which is the platform for technological developments (workpackages 3-7).

(3) GRACE-PAT which is the platform for patient studies (workpackages 8-11).

(4) GRACE-EDUT which is the platform for education and training (workpackage 12).

The health economics elements of GRACE formed part of platform 3 (platform for patient studies) and was aimed at (i) studying the economics of molecular diagnostics in community acquired lower respiratory tract infections (CA-LRTI), (ii) modelling the macroeconomic impact of resistance and policies to contain resistance, (iii) modelling the cost-effectiveness of the management strategies developed in the observational studies and (iv) conducting economic evaluations in parallel with the intervention studies (GRACE accessed on 27th November, 2015).

The case studies for this thesis emerged from this platform and the observational and interventional studies that were conducted in workpackage 8 and workpackage 10. These studies were considered appropriate for the aims of the PhD research work because they were multinational studies. A description of both workpackages (8 and 10) is presented below.

4.4 Workpackage 8 (Observational study on determinants of antibiotic use)

GRACE workpackage 8 was an observational study aimed at describing the current presentation, investigation, treatment and outcomes of community acquired-lower respiratory tract infections and analysing the determinants of antibiotic use in 14 primary care networks across 13 European countries, using qualitative and quantitative approaches. Primary care research networks with a good track record in research and potential to contribute to the study were selected from 13 countries: (United Kingdom (England and Wales), the Netherlands, Germany, Belgium, Italy, Hungary, Slovakia,

Sweden, Norway, Finland, Spain and Poland). Each network was led by a National Network Coordinator (NNC) and a National Network Facilitator (NNF) who were responsible for setting up the network, recruitment of patients and the management of data in their respective networks. The trial registration number for this study is Clinicaltrials.gov NCT00353951. Consenting patients were included if they were 18 years and over, consulting with acute cough as one of the main symptoms, consulting for the first time within their current illness episode and who were able to complete study materials (Butler et al. 2009). Study participants were followed up for four weeks and the primary outcomes of the study were prescribing of antibiotics by clinicians and total symptom severity scores over time. Secondary outcome measures included the 3-Level version of the EQ-5D.

4.5 Workpackage 10 (Randomised controlled trials)

Workpackage 10 of the GRACE study was aimed at understanding which subgroups of individuals benefit from antibiotics and how to change doctor-prescribing behaviour. Based on the results that were obtained from two of the previous workpackages (workpackage 8 and workpackage 9), two types of studies were conducted. The first was a multinational, randomised placebo-controlled double-blind trial with patient level randomisation, to study the clinical effectiveness of antibiotic therapy (amoxicillin) in community acquired lower respiratory tract infections (WP10A). This will be known as Case study 1 in this thesis. The second was a multinational, cluster, randomised, factorial controlled trial with primary care practice level randomisation, to study how

improvements in antibiotic prescribing behaviour by general practitioners can be achieved (WP10B or GRACE INTRO TRIAL). This will be known as Case study 2 in this thesis. A summary of both trials in workpackage 10 (WP10A and WP10B) is presented in the next section.

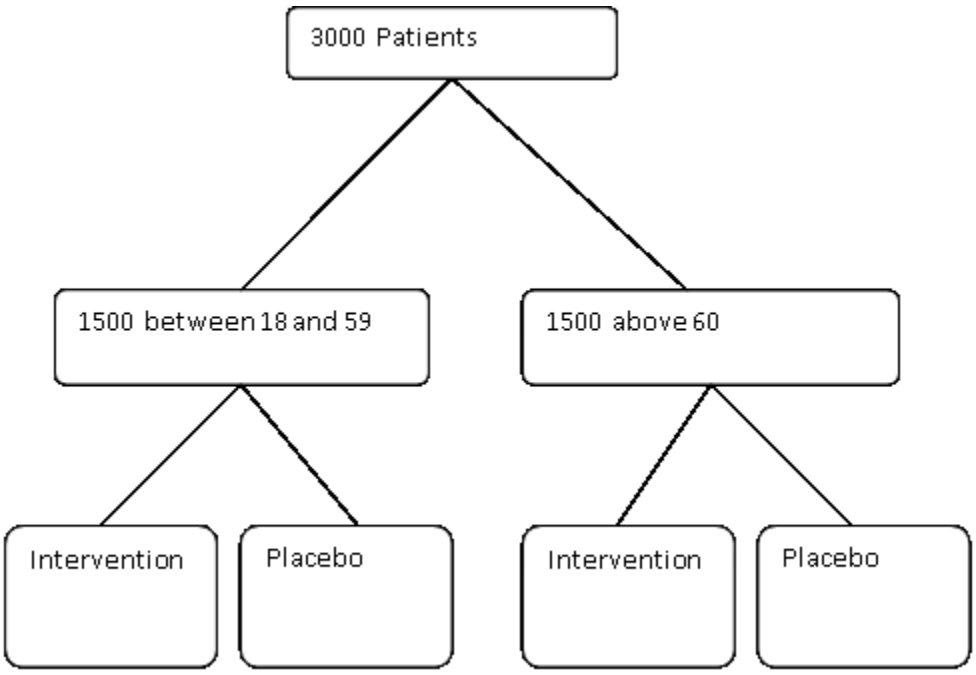
4.5.1 Overview of the WP10A trial design (Case study 1)

The aim of the WP10A trial was to assess the effectiveness of antibiotic therapy for the treatment of community acquired acute cough/lower respiratory tract infections in order to determine which patients actually benefit from antibiotic treatment, and which do not benefit. The study was carried out in 16 primary care networks across 12 countries in Europe: Belgium, France, Germany, Italy, Netherlands, Poland, Slovakia, Slovenia, Spain, Sweden and the United Kingdom (England and Wales). The trial registration number for this study is ISRCTN52261229.

This trial aimed to recruit 3000 patients in two age groups. The first group were those aged between 18 and 59 with the second group aged 60 years and above, thus allowing sub-group analysis to be undertaken. Within each group, patients were randomised to receive either a placebo (control) or the intervention (amoxicillin). Patients were included in the study if they were adults aged 18 and above and consulting with acute cough, or were suspected to have symptoms of acute cough/lower respiratory tract infections by their GP. The primary outcomes of the study included deterioration of illness and symptom severity and duration. A secondary outcome measure was the 3-level version of the EQ-5D (Rabin and de Charro, 2001) which was used as the primary outcome measure for the health economic analysis. Additional details of the trial design and main results

from this workpackage have been presented elsewhere (Little et al. 2013a; Moore et al. 2014). A summary of the trial design is presented in Figure 4.1 below.

Figure 4.1 Summary of WP10A study (Case study 1)



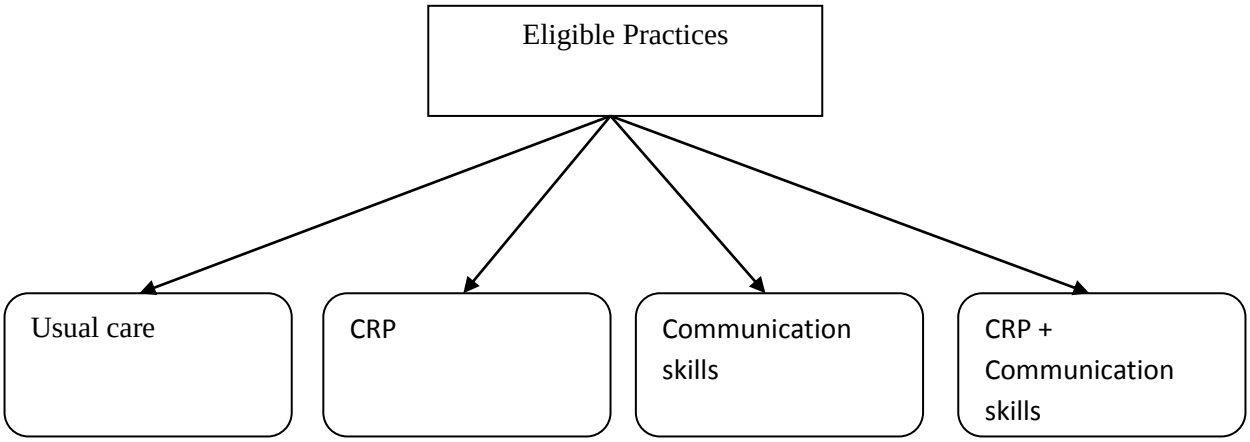
4.5.2 Overview of the WP10B/ GRACE INTRO trial design (Case study 2)

The aim of the Genomics to combat Resistance against Antibiotics in Community-acquired LRTI in Europe/Internet training for reduced antibiotic prescribing for acute LRTI (GRACE/INTRO) or WP10B study was to develop an internet-based training tool for lower respiratory tract infections. To achieve this aim, primary care practices were selected from five European countries Belgium, Netherlands, Poland, Spain, and the United Kingdom (England and Wales). They were randomised into one of the following interventions: (i) usual care; (ii) internet-based training to use a point-of-care CRP test;

(iii) internet-based training in enhanced communication skills; or (iv) combined training in C-reactive protein testing and enhanced communication skills. The trial registration number for this study is ISRCTN99871214.

General practitioners who were randomised to the usual care arm assessed and managed patients according to the practice's usual practices and procedures. The CRP group received internet based training mainly on how to target testing in cases of clinical uncertainty. Those that were randomised to the enhanced communication skills group received training how to gather information on patients' concerns and expectations, exchange of information on symptoms etc., and how to provide guidance on when to re consult. GPs in this group were given an interactive booklet to use during consultations with patients, which included information on symptoms, antibiotic resistance etc. The combined intervention consisted of both training in CRP testing and training in enhanced communication skills. A more detailed description of the study design, interventions and main results have been presented elsewhere (Little et al. 2013b; Yardley et al. 2013; Anthierens et al. 2012; Anthierens et al. 2015). The primary outcome over the four week study period was antibiotic use, as documented on the case-report forms. Secondary outcomes included the 3-level version of the EQ-5D, new or worsening symptoms and symptom severity.

Figure 4.2 Summary of WP10B study (Case study 2)



4.6 Application of the case studies

Each of the studies described above were used to explore specific aspects of the PhD thesis. The observational study (workpackage 8) was used to explore issues relating to costing and outcomes in multinational trials (Chapters 5 and 6) whilst Case studies 1 and 2 were used to explore the various analytical approaches (pooled and split) to economic evaluation alongside multinational trials (Chapter 7).

4.7 Summary

This chapter was primarily aimed at summarizing the case studies that were used to investigate the research questions raised in the thesis. Three main case studies: an observational study (workpackage 8), and two randomised trials (Case study 1) and (Case study 2) were described. All case studies were from the GRACE project, an EU funded project which was aimed at bringing together physicians and scientists from across Europe in order to combat antibiotic resistance in community-acquired lower respiratory tract infections. The next chapter presents the results of the empirical work alongside the observational study (workpackage 8) which focuses on exploring issues relating to costing alongside multinational studies.

CHAPTER 5 COSTING IN ECONOMIC ANALYSIS ALONGSIDE MULTINATIONAL STUDIES

The previous chapter described the case studies that will be used in this PhD thesis. This chapter explores issues relating to costing in multinational trials. The chapter starts by outlining the methods used in the study followed by a presentation of the results. The discussion section focuses on the challenges faced in collecting these data and suggests some solutions and strategies for researchers faced with similar problems in future multinational studies of infection. The results presented in this chapter have been published in a peer reviewed article *"Oppong et al. (2011) Resource use and costs of treating acute cough/lower respiratory tract infections in 13 European countries: results and challenges European Journal of Health Economics 12: 319-329"*

5.1 Background

Considering the results from the systematic review in Chapter 3, it is clear that one of the major unresolved challenges associated with conducting economic evaluation alongside multinational trials relates to costing alongside those studies. Practical difficulties that were identified in the review and which have also been mentioned in previous studies include: problems associated with the acquisition of information about resource use and unit costs in a variety of health systems and countries (Pang et al. 2002, Raikou et al. 2000). Others include problems associated with interpretation as a result of differences in

languages (Virk, 2008). Neither of these issues may be well understood by researchers who undertake costing studies alongside multinational trials. Evidence from the literature also points to the fact that in some settings reliable (or indeed any) information about unit costs may not be available (Schulman et al. 1998, Wordsworth et al. 2005). Chapter 3 also suggested that researchers tended to adopt the one country costing approach suggesting that there are difficulties associated with collecting unit cost data in multinational trial settings. It is therefore important for a study aimed at studying potential problems associated with costing within multinational trial settings to be conducted in order to throw more light on potential problems associated with this type of costing exercise and also help to establish possible solutions to the problems and challenges that have been identified.

Using data from a multinational observational study (workpackage 8 described earlier in Chapter 4) of LRTIs conducted across thirteen European countries: United Kingdom (England and Wales), Netherlands, Spain, Germany, Hungary, Belgium, Poland, Italy, Sweden, Norway, Finland and Slovakia (workpackage 8), this chapter investigates issues associated with costing within a multinational trial setting. Specific objectives of this chapter include the following: First, it aims to throw light on the challenges associated with the collection and use of costing data in multinational trial/study settings and second, it aims to present estimates of resource use, and cost for the treatment of LRTIs overall and in individual countries.

5.2 Methods

5.2.1 Study design

This study was conducted alongside the GRACE workpackage 8 study which was an observational study of the presentation, management and outcomes of patients with acute cough and LRTIs in primary care in Europe (Butler et al. 2009).

5.2.2 Patients and setting

As described in the previous chapter, recruitment into the study took place in 14 primary care networks across 13 European countries (see previous chapter). All consenting patients with acute cough/lower respiratory tract infections who consulted GPs in these primary care networks were recruited sequentially into the study from October – November 2006, and January – March 2007. Additional details of the study and the main clinical results have been presented elsewhere in a published study by the GRACE consortium (Butler et al. 2009).

5.2.3 Collection of resource use data

Resource use data were collected from a societal perspective, including resources used by the health system, by patients and their families and time off work as a result of illness. Resource use data were obtained from a patient diary and a case records form (CRF). The CRF was filled in by the GP and the patient diary was completed by the patient. Both (CRF and patient diary) were completed over a 4 week period. Four different elements of resource use were collected:

Health professionals: number of visits to nurses and/or doctors in various primary care settings; visits to pharmacists; number of referrals to specialists. Where applicable (i.e. the service existed) information was collected about visits to other forms of primary care service such as walk-in-centres.

Investigations: type and number of investigations, e.g. chest x-rays, C-reactive protein tests and procalcitonin

Medication: both prescribed and over-the-counter medication, including dosage and number of days taken.

Lost productivity: including days off work and loss of earnings.

To estimate productivity loses due to absenteeism, data on patients work status, occupation and time off work over the four week period were obtained from patients responses to the patient diary in each country. Data on country-specific average wages by occupation group were also obtained from national databases as well as international databases such as Eurostats and International Labour Organisation. Productivity costs were derived according to the human capital approach which is based on the human capital theory and assumes that the value of lost productivity is equal to the amount of resources an individual would have been paid to do that work. This method values productivity losses by measuring time lost from work and multiplying this by the gross wage (Sculpher 2001, van den Hout, 2010).

5.2.4 Valuation of resource use data

As a result of the study being a multinational costing study and the lack of direct access to unit cost data in all participating countries, a step by step process was used to obtain unit cost data for all aspects of resource using a number of approaches. These comprised four main approaches and other approaches which were developed when unit cost data was not available in some countries. The four main approaches were:

(1) Searching the internet mainly through search engines such as Google to identify international and national sources of cost data such as statistical services websites as well as databases.

(2) An emailed request for information was developed and sent to all members of the UK Health Economists' Study Group (HESG). This asked members to respond if they had previously obtained cost data for any of the countries included in the study or had conducted a multinational study which involved all or some of the countries included in the study. Details of the email sent out to UK HESG members is presented in Appendix 7.

(3) A spreadsheet was developed and circulated to national network co-ordinators (NNCs) and national network facilitators (NNFs) in all participating countries to identify sources of unit cost data and, where possible, actual unit cost data for the various resource use items. Appendix 7 presents details of the spreadsheet that was sent.

(4) A list of health economists and researchers working in the European countries which participated in this study was compiled and each individual contacted directly. These health economists were identified mainly through personal communication.

The four approaches explained above are the main approaches that were used to obtain unit cost data. Where valuation data were still unavailable for some resource use items in some countries, estimates were generated using the following methods developed during the study: First, the thirteen countries included in the study were classified into four groups according to geographical proximity and gross domestic product (GDP). Grouping of countries based on certain similarities is an approach that has been used by other researchers conducting economic analysis alongside a multinational trial (Simon et al. 2006; Radeva et al. 2005). Group one was made up of Germany, Belgium, England, Wales and Netherlands (western European group); group two consisted of Finland, Sweden and Norway (Nordic group); group three included Spain and Italy (southern European group); and Slovakia, Hungary and Poland (eastern European group) made up group four. Where costs were not available in a particular country, it was obtained from countries in the same group. In some cases, there was only unit cost information available for a single country. In such cases, a market basket approach was used to derive a relationship between that country and the other countries in order to generate a cost for each country from the one data point available. This approach is similar to that used in a previous study (Schulman et al. 1998). The market basket approach is a method whereby an index that represents the relationship between prices between countries is generated based on a number of goods in a particular country. This method was applied to resource use items such as walk in centres where costs were only available for the UK. The various indices to show the relationship between the unit costs in different countries have been presented in Appendix 8. The source of valuation data for each item of resource use in all participating country is given in table 5.1.

All costs were initially converted into 2007 prices for the country of origin. Costs were then converted into a value to aid comparison across countries. This was done in two ways. First, values are presented in Euros in 2007 prices with the Euro value obtained using average exchange rates for 2007 (Banque de France). Costs from Germany, Finland, Netherlands, Belgium, Italy and Spain were collected in Euros and did not require any conversions. Second, values are presented in Euros using the PPPs between Germany and all other countries for 2007 (OECD). It would have been preferable to convert into 2007 figures using the EU-25 average purchasing power parity, but the latest date for which this figure is provided by the Organisation for Economic Cooperation and Development (OECD) is 2005. It was anticipated that changes in the purchasing power associated with different currencies would not be so large as to make this figure meaningless. In addition to this and to ensure that costs are generalisable to the UK, costs were also converted into UK pounds sterling using the average exchange rates for 2007. All exchange rates that were used to convert between currencies are provided in Table 5.2.

The following adjustments to costs were made for particular items:

Primary care health professionals: In some networks, the exact consultation time was available. However, in other networks, a range was given. In this case, a mean consultation length was generated from the range. In cases where there was no information available, the consultation period was assumed to be 10 minutes.

Specialist health professionals: Information about consultation length with specialists was not collected. It was therefore assumed that patients spent 10 minutes with a senior specialist health professional.

Investigations: For all investigations, the costs obtained are assumed to include the costs of staff time associated with the investigation. This assumption had to be made as a result of the difficulty associated with collecting unit cost data from all countries.

Over-the-counter medication: As there were several types of over-the-counter drugs taken across the 12 countries, it was not feasible to acquire unit cost data for each, particularly given the limitations of unit cost data experienced. Instead, drugs were first classified into 8 groups: cough medicine, throat medicine, decongestants, pain killers, cold and flu medicine, herbal medicine, antibiotics and 'other'. Over-the-counter antibiotics were used only on two or three occasions across the entire study and these drugs were costed on the basis of their actual unit cost as for all the other antibiotics in the study; the 'other' group mainly comprised various types of vitamins. It was decided that, as these 'other' drugs may have been taken in any case, rather than directly related to the cough symptoms, to exclude them from further costing. For the remaining six groups a single cost was generated by calculating a mean price from a selected list of drugs in each class.

Table 5.1 Sources of valuation data

Country	Wales	England	Netherlands	Spain	Spain	Germany	Hungary	Belgium	Poland	Italy	Sweden	Norway	Finland	Slovakia
Network	Cardiff	Southampton	Utrecht	Barcelona	Mataro	Rotenburg	Balatonfured	Antwerp	Lodz	Milan	Jonkopin	Tromso	Helsinki	Bratislava
GP Visits	4	4	5	1	1	6	1	7	3	2	2	8	9	2
Nurse Visits	4	4	5	1	1	2	1	2	3		2	2	9	2
Specialist Visits			10	1	1	1	1	2	1	2	2		2	2
Out of hours GP	23	23	24	24	24	24	24	24	24	24	24	24	24	24
Walk in centre	23	23	24	24	24	24	24	24	24	24	24	24	24	24
Hospital Admissions	23	23	24	24	24	24	24	24	24	24	24	24	24	24
Investigations	22	22	5	6	6	19	1	21	1	20	2	8	9	2
Medication	12	12	8, 13	11	11	14	2	15	8, 2	8, 2	16	17	18	2

1= Health Basket, 2= Derivation from group, 3= NHF www.nfz.gov.pl, 4= Netten and Curtis (www.pssru.ac.uk), 5= Central Tariff (CTG), 6= Gebührenordnung für Ärzte (GOÄ), 7= http://www.inami.fgov.be, 8= Direct communication, 9= Hujanen (www.stakes.fi), 10= Dutch health service, 11= www.vademecum.es, 12= British National Formulary (www.bnf.org), 13= Dutch healthcare insurance board (www.medicijnkosten.nl), 14= Rote Liste (The red book) www.rote-liste.de, 15= www.bcfi.be, 16= Swedish Pharmaceutical Benefits Board (www.lfn.se), 17= Statens legemiddelverk (Norwegian Medicines Agency), 18= Pharmaca Fennica (www.fimnet.fi/cl_terveysportti/fimnet), 19= Tariffs of CSdM and tariffs of the Barcelona official Physicians (www.comb.cat), 20= Italian NHS costs, 21= www.http://riziv.fgov.be, 22= Southampton University Hospital Trust, Costing Department, 23= NHS Reference costs

24= Market basket approach, 25= ILO labour statistics http://laborsta.ilo.org

Table 5.2 Exchange rates and GDP

Country	**Wales**	**England**	**Netherlands**	**Spain**	**Spain**	**Germany**	**Hungary**	**Belgium**	**Poland**	**Italy**	**Sweden**	**Norway**	**Finland**	**Slovakia**
Network	Cardiff	Southampton	Utrecht	Barcelona	Mataro	Rotenburg	Balatonfured	Antwerp	Lodz	Milan	Jonkopin	Tromso	Helsinki	Bratislava
Average yearly Exchange rates to Euro	0.684	0.684	1	1	1	1	251.32	1	3.783	1	9.250	8.007	1	33.775
PPP Exchange rates	0.754	0.754	1.003	0.841	0.841	1	152.990	1.014	2.186	0.981	10.226	10.091	1.119	19.230
GDP per capita (2007)	35000	35000	39000	33600	33600	34100	19300	36200	16200	30900	37500	53300	36000	20200

Conversion of Euros to British Pounds: Average yearly exchange rate 2007=0.684505 source: http://www.ukforex.co.uk/forex-tools/historical-rate-tools/yearly-average-rates

5.2.5 Data analysis

Descriptive statistics were used to show differences in resource use and costs across countries. Mean cost per patient was estimated and is presented along with standard deviations. No imputation of missing data was carried out, given that it was not clear that missing data would be missing at random and also there was almost no resource use data available for those not completing the diary. To account for clustering at the clinician level, standard deviations of resource use items were inflated for clustering (Donner and Klar, 2000). Data analysis was carried out using STATA version 12 (STATA, 2012) MS Excel and SPSS version 16 (SPSS, 2007). A statistical test (Kruskal-Wallis test) was used to determine whether there were differences between networks. Since costs were likely to be skewed (Glick et al. 2007), non-parametric bootstrapping (1000 replications) was used to obtain confidence intervals around the difference in costs and resource use between the various networks using Cardiff as the reference network.

5.2.6 Sensitivity analysis

The aim of the sensitivity analysis was to explore more complex statistical and regression methods that could be used to estimate unit costs in a multinational trial setting. To achieve this, multilevel modelling, an approach which is able to account for the clustered nature of the data was used to estimate the costs associated with the management of acute cough/LRTIs (Goldstein, 2011). For this particular analysis, patients (level one) were clustered within general practitioners (level two). Is should be pointed out here that this method was only used for the estimation of total costs and not the individual cost items. The results that were obtained from the base case (unadjusted analysis) and the multilevel modelling approach were compared to see if different approaches led to different results.

5.3 Results

5.3.1 Patient characteristics

The study population included a total of 3402 patients across the 14 primary care networks in Europe, for whom data were available for 3368 patients from the CRF and 2714 patients from the patient diary (Figure 5.1). For the purpose of the economic analyses and given the decision not to impute the large quantities of resource use data that would be required where there was no CRF or no diary, patients who filled in both the patient diary and the CRF forms were selected. The final study population used in this study consisted of a total of 2690 patients of whom 63.8% were female and 36.2% were male. The mean age of the study population was 48 years. The number of patients in each network ranged from 90 in the Finish network (Helsinki), to 320 in the Hungarian network (Balatonfured). Table 5.3 includes information about both the number of GP practices and the number of GPs sampled in each network.

Figure 5.1 Patient flow chart

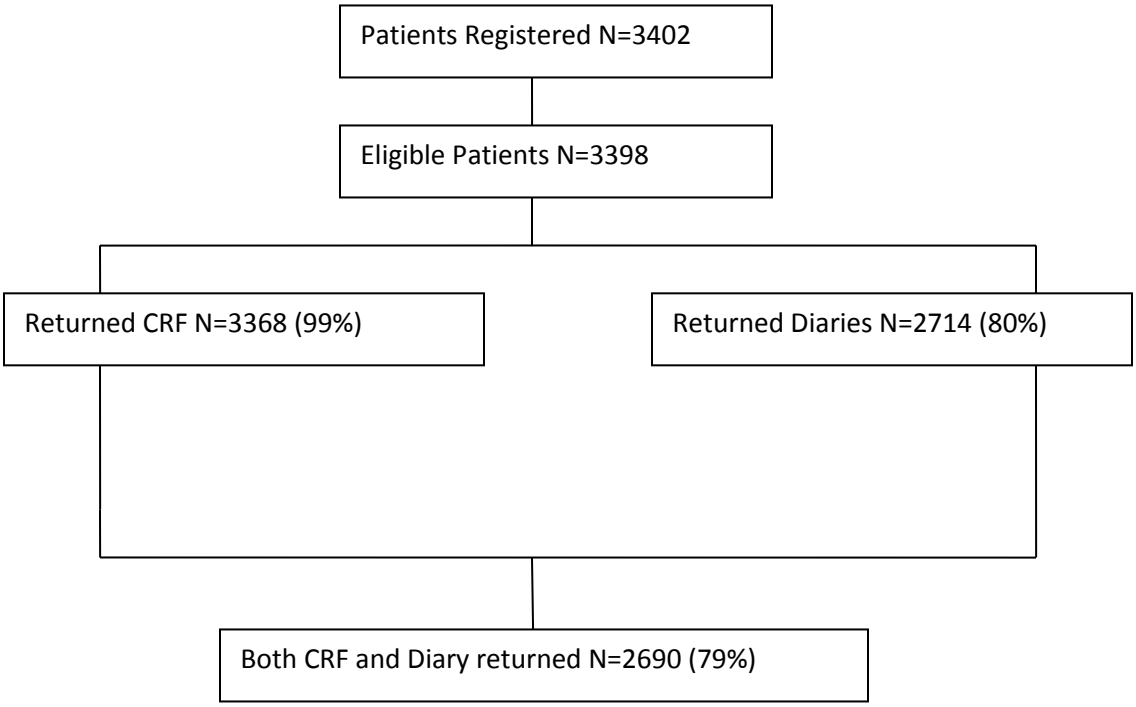


Table 5.3 Number of practices, general practitioner (GPs) and patients

Country	Network	Practices	GPs	Patients
Wales	*Cardiff*	5	26	181
England	*Southampton*	6	23	168
Netherlands	*Utrecht*	11	34	195
Spain	*Barcelona*	3	26	169
	Mataro	3	21	179
Germany	*Rotenburg*	16	16	181
Hungary	*Balatonfured*	11	11	320
Belgium	*Antwerp*	18	26	164
Poland	*Lodz*	9	21	221
Italy	*Milan*	13	12	153
Sweden	*Jonkopin*	12	81	222
Norway	*Tromso*	11	41	148
Finland	*Helsinki*	2	26	90
Slovakia	*Bratislava*	5	23	299

5.3.2 Success in methods of obtaining unit cost data for the valuation of resource use

The four main approaches to obtaining unit cost data collection yielded varying levels of success.

(i) Internet search

The initial internet search tended only to identify websites for national statistical departments, which in general did not contain data at the level of detail required, and pharmaceutical organisations which in some cases (e.g. Sweden and Norway) led to useful links to databases of drug prices. In some cases, e.g. in Hungary even though the search was able to identify national statistical services that could provide cost information; it was only going to be available at an extra cost. Some of the useful databases and websites discovered through this route include the following: International

Labour Office website (ILO) where information on the sources of wages in the different countries was obtained. The Organisation for Economic Cooperation and Development databases where information on exchange rates and PPPs was obtained.

(ii) Email to UK Health Economists Study Group Members

There were two responses to the email sent to members of the UK HESG and no additional sources of cost information were directly identified through this route. However, in some cases, names of people who could potentially help with unit cost information were provided.

(iii) Spreadsheet circulated to National Network Facilitators and National Network Coordinators

Sending out the forms to the networks yielded a number of cost data values and/or sources for nine of the twelve countries but no data from Slovakia, Sweden and Hungary was obtained due to non-response to the questionnaire. For Norway, information was only given about values and no information was provided about sources of data from which these figures had been obtained. Some of the networks found it difficult to engage with the collection of cost data. For example, one of the network facilitators observed that collecting cost information was a very difficult task. Where a response was obtained, information was most often provided for some and not all of the specified items. Around half of the responding networks were able to identify a source for visits such as those by GPs. Only one network, Antwerp, was able to provide a data source for medication costs. Examples of responses obtained from the cost questionnaires sent out to the Belgian and Polish Networks are presented in the appendix. It highlights the differences in the level of

detail that was obtained from the various networks. Clearly, the information obtained from the Polish national network coordinators was not as comprehensive as that which was obtained from the Belgian network (Antwerp) (Appendix 7).

(iv) Direct contact with health economists and researchers

Directly contacting health economists and researchers working in the particular country of interest proved to be the most effective method of identifying unit cost information sources, particularly in relation to eastern European countries where no success had been achieved with the previous approaches. For example, contacting a researcher in Germany led to the discovery of a website where prices of various drugs are presented. It should be noted that some of the researchers that were directly contacted were identified through the email that was sent to HESG members (approach 2 above).

More detailed results in relation to the costs associated with the management of acute cough/lower respiratory tract infections are now presented below:

5.3.3 Resource use

(i) Health Professionals

The mean number of GP visits over the four week period was similar across the various networks (see Table 5.4), with approximately one visit to the GP. Balatonfured was the only network that had a mean of more than two visits to the GP. GPs in Hungary might have influenced patients to re-consult as a result of the €1.2 (£0.82) visit fee that was introduced in 2007 (Rurik, 2007). Visits to nurses were also similar across the networks with the highest number of visits recorded in networks such as Balatonfured (Hungary) and Bratislava (Slovakia). However, there was no recorded visit to the nurse in

Rotenburg or Milan. Most patients across the networks were not referred to a specialist by their GPs, and there was no reported case of visits to specialists in Southampton and Cardiff. Hospital admissions were recorded in about half of the networks but none recorded in Cardiff, Utrecht, Barcelona, Lodz, Tromso and Helsinki. Even though most networks recorded visits to a GP out of hours, the mean number of visits was very low, only 0.011 visits per network. Walk in centres were one of the least used services across the networks, with only 3 (Milan, Balatonfured and Cardiff) of the 14 networks recording any visits to a walk in centre. It should be noted however, that services such as walk in centres were not available in all networks.

(ii) Medical Investigations

All networks reported the use of chest x-rays and full blood count. With the exception of Barcelona and Lodz, all the other networks reported the use of CRP either near-patient or lab based. Procalcitonin was the least used investigation (Cardiff, Balatonfured and Bratislava only). Other investigations with low usage (used in fewer than 6 networks) were blood for serology and sputum for culture. Whilst investigations with high usage (used in more than 8 networks) included Full blood count, chest x-ray, C-reactive protein, electrolyte and spirometry. Mean values of less than one were recorded for the total use of investigations in all networks. Table 5.4 gives a breakdown of medical investigations used across all networks.

(iii) Medication

Antibiotics were prescribed in 53% of the 2690 patients. The eastern European networks Balatonfured (74%), Lodz (72%) and Bratislava (87%) recorded the highest proportions

of patients receiving a prescription for antibiotics. On the other hand, Mataro (34%), Jonkopin (37%) and Tromso (30%) recorded low prescription rates. Advice on delayed prescriptions was higher in Southampton where delayed use was advised in over 50% of patients. This is probably due to campaigns aimed at reducing the prescribing of antibiotics and guidance that advocate delayed prescribing in UK (Little 2005, NICE 2008). Over the counter drugs were taken by 97% of patients in all networks before they consulted their GP. The most common class of over the counter drugs taken were cough medicines and painkillers. Herbal medicines were least used with Rotenburg and Bratislava recording the highest intake. Other studies have reported high sales values for over the counter herbal medicine in Germany (De Smet, 2005). Patients indicated that they purchased antibiotics over the counter in some networks: Barcelona, Mataro, Balatonfured, Lodz, Milan, and Bratislava, although whether this was actually the case is questionable. With the exception of Spain (Goossens et al. 2005) most countries do not allow over-the-counter purchase of antibiotics.

Table 5.4 Mean resource use (Standard deviation)

Country	Wales (n=181)	England (n=168)	Netherlands (n=195)	Spain (n=169)	Spain (n=179)	Germany (n=181)	Hungary (n=320)	Belgium (n=164)	Poland (n=221)	Italy (n=153)	Sweden (n=222)	Norway (n=148)	Finland (n=90)	Slovakia (n=299)
Network	Cardiff	Southampton	Utrecht	Barcelona	Mataro	Rotenburg	Balatonfured	Antwerp	Lodz	Milan	Jonkopin	Tromso	Helsinki	Bratislava
						HEALTH PERSONNEL/SERVICES								
GP Visits	1.37 (0.75)	1.25 (0.60)	1.49 (0.86)	1.32 (0.64)	1.29 (0.48)	1.89 (1.83)	2.08 (2.97)	1.53 (1.12)	1.39 (0.73)	1.38 (0.83)	1.30 (0.67)	1.21 (0.71)	1.20 (0.59)	1.96 (1.90)
Nurse Visits	0.022 (0.24)	0.042 (0.33)	0.041 (0.20)	0.006 (0.77)	0.011 (0.18)	0	0.813 (3.92)	0.024 (0.15)	0.032 (0.24)	0	0.072 (0.32)	0.020 (0.14)	0.056 (0.27)	0.224 (0.96)
Specialist Visits	0	0	0.036 (0.24)	0.053 (0.31)	0.011 (0.11)	0.044 (0.33)	0.025 (0.44)	0.061 (0.34)	0.054 (0.44)	0.085 (0.47)	0.009 (0.12)	0	0.033 (0.23)	0.167 (0.77)
Out of hours GP	0	0	0.010 (0.10)	0.030 (0.17)	0.017 (0.13)	0.006 (0.11)	0.050 (0.40)	0.006 (0.08)	0.045 (0.55)	0.020 (0.18)	0.023 (0.16)	0.020 (0.14)	0	0
Walk in centre	0.011 (0.16)	0	0	0	0	0	0.031 (0.39)	0	0	0.007 (0.08)	0.014 (0.18)	0	0	0
Hospital Admissions	0	0.510 (0.50)	0	0	0.011 (0.11)	0.533 (0.65)	0.041 (0.73)	0.360 (0.42)	0	0.132 (1.63)	0.010 (0.10)	0	0	0.146 (1.18)
Pharmacist Visits	0.204 (0.71)	0.268 (0.83)	0.456 (0.84)	0.142 (0.49)	0.162 (0.37)	0.646 (1.26)	0.671 (2.46)	0.671 (0.86)	0.285 (0.78)	0.216 (0.62)	0.550 (0.86)	0.088 (0.37)	0.278 (0.61)	0.794 (1.82)
						INVESTIGATIONS								
Full Blood Count	0.0055 (0.07)	0.0060 (0.08)	0.0103 (0.10)	0.0059 (0.08)	0.0335 (0.18)	0.1602 (0.37)	0.0281 (0.17)	0.0122 (0.11)	0.0223 (0.15)	0.0261 (0.16)	0.0059 (0.29)	0.1622 (0.37)	0.1444 (0.35)	0.2542 (0.44)
Chest X-ray	0.0110 (0.10)	0.0238 (0.15)	0.0205 (0.14)	0.0414 (0.20)	0.1564 (0.36)	0.0331 (0.18)	0.0469 (0.21)	0.0305 (0.17)	0.0181 (0.13)	0.0392 (0.19)	0.0315 (0.18)	0.0473 (0.21)	0.0778 (0.27)	0.1271 (0.33)
C-RP	0.0055 (0.07)	0.0060 (0.08)	0.0051 (0.07)	0	0.0056 (0.07)	0.1326 (0.34)	0.0125 (0.11)	0.0122 (0.11)	0	0.0196 (0.14)	0.6441 (0.48)	0.9324 (0.25)	0.2111 (0.41)	0.1672 (0.37)

Country	Wales (n=181)	England (n=168)	Netherlands (n=195)	Spain (n=169)	Spain (n=179)	Germany (n=181)	Hungary (n=320)	Belgium (n=164)	Poland (n=221)	Italy (n=153)	Sweden (n=222)	Norway (n=148)	Finland (n=90)	Slovakia (n=299)
Electrolyte Sediment rate	0.0055 (0.74)	0.0060 (0.08)	0.0051 (0.07)	0	0	0.1492 (0.36)	0.0281 (0.17)	0.0122 (0.11)	0.0090 (0.09)	0.0131 (0.11)	0.0495 (0.22)	0.0946 (0.29)	0.0222 (0.15)	0.3076 (0.46)
Procalcitonin	0.0055 (0.07)	0	0	0	0	0	0.0063 (0.08)	0	0	0	0	0	0	0.067 (0.08)
Urea of Creatinine	0.0055 (0.07)	0.0060 (0.08)	0.0051 (0.07)	0.0059 (0.08)	0.0335 (0.18)	0.0331 (0.18)	0.0188 (0.14)	0.0061 (0.08)	0.0045 (0.07)	0.0065 (0.08)	0.0180 (0.13)	0.0068 (0.08)	0	0.0769 (0.27)
Electrolytes	0.0055 (0.07)	0.0060 (0.08)	0	0	0.0355 (0.18)	0.0221 (0.15)	0.0156 (0.12)	0	0.0090 (0.09)	0	0.0090 (0.95)	0	0	0.0736 (0.26)
Sputum for culture	0.0055 (0.07)	0.0179 (0.03)	0	0	0	0	0.0063 (0.08)	0	0	0	0	0.0203 (0.14)	0	0.1271 (0.33)
Blood for Serology	0.0055 (0.07)	0	0	0.0059 (0.08)	0	0	0.0063 (0.08)	0	0	0	0.0450 (0.21)	0.0068 (0.08)	0	0.1003 (0.30)
Nose swab	0.0055 (0,07)	0	0	0.0059 (0.08)	0	0	0.0063 (0.08)	0	0.0045 (0.07)	0	0.0315 (0.18)	0.0338 (0.18)	0.0111 (0.11)	0.1204 (0.64)
Spirometry	0.0055 (0.01)	0.0060 (0.08)	0.0205 (0.14)	0.0118 (0.11)	0	0.0773 (0.27)	0.0094 (0.10)	0.0122 (0.11)	0.0045 (0.07)	0.0131 (0.11)	0.0045 (0.07)	0.0608 (0.24)	0	0.0502 (0.22)

5.3.4 Health care costs

The total mean costs per patient of treating LRTI in each network are presented in Table 5.5. The results show that most items did not contribute significantly to cost because of the low resource use associated with these items. The main cost drivers in most networks were GP visits, medication costs, as well as costs of medical investigations.

Clearly, the use of different methods of conversion leads to different estimates of the mean cost per patient in different settings (see Table 5.5) and in most cases the costs derived from PPP's were smaller than the costs that were derived from the exchange rates. The only cases where the opposite was true was the Spanish networks (Mataro and Barcelona), Belgium and Italy (Table 5.5).

With the normal exchange rates (base case analysis) the highest mean cost per patient was recorded in Jonkopin £79.72 and the lowest was recorded in Balatonfured £16.35. The results of the Kruskal-Wallis test showed that there was a significant difference between the mean costs across networks ($p<0.01$). However, when bootstrapping was used to test the difference in costs between the Cardiff network and the others, there were some cases where the differences were found to be statistically significant (Table 5.6). When PPP exchange rates were used for the conversion, the highest cost was recorded in Antwerp (Belgium) and the lowest cost was recorded in Lodz (Poland).

Table 5.5 Mean costs £ (Standard deviation)

	Cardiff (n=181)	Southampton (n=168)	Utrecht (n=195)	Barcelona (n=169)	Mataro (n=179)	Rotenburg (n=181)	Balatonfured (n=320)	Antwerp (n=164)	Lodz (n=221)	Milan (n=153)	Jonkopin (n=222)	Tromso (n=148)	Helsinki (n=90)	Bratislava (n=299)
	HEALTH PERSONNEL/SERVICES													
GP Visits	20.83 (10.79)	30.18 (14.38)	9.19 (5.28)	7.17 (3.46)	6.96 (2.59)	9.02 (6.59)	7.41 (0.46)	18.94 (11.75)	6.55 (3.43)	9.93 (5.42)	33.68 (17.27)	38.34 (17.67)	38.45 (13.75)	4.63 (2.50)
Nurse Visits	0.12 (0.83)	0.23 (1.88)	0.25 (1.23)	0.01 (0.17)	0.03 (0.34)	0	0.39 (0.51)	0.15 (0.95)	0.04 (0.33)	0	0.51 (2.25)	0.16 (1.11)	0.86 (4.23)	0.21 (0.80)
Specialist Visits	0	0	2.20 (14.44)	1.21 (5.39)	0.24 (2.27)	3.77 (22.25)	0.29 (2.29)	4.47 (25.29)	0.66 (5.44)	1.93 (10.73)	0.71 (7.51)	0	2.25 (15.80)	1.99 (7.79)
Walk in centre	0.31 (4.15)	0	0	0	0	0	0.06 (0.43)	0	0	0.08 (1.02)	0.38 (4.15)	0	0	0
Hospital Admissions	0	2.27 (22.47)	0	0	0.21 (2.02)	1.33 (16.32)	0.13 (2.27)	1.29 (15.18)	0	2.68 (32.96)	0.42 (4.31)	0	0	1.03 (8.34)
	INVESTIGATIONS													
Full Blood Count	0.02 (0.24)	0.01 (0.19)	0.002 (0.02)	0.03 (0.42)	0.18 (0.98)	0.51 (1.18)	0.04 (0.23)	0.06 (0.58)	0.04 (0.26)	0.07 (0.44)	0.40 (1.24)	0.84 (1.90)	0.42 (1.03)	0.29 (0.50)
Chest X-ray	0.42 (3.93)	0.42 (2.68)	0.11 (0.67)	0.44 (2.17)	1.70 (3.95)	0.60 (3.22)	0.03 (0.12)	0.29 (1.63)	0.05 (0.41)	0.46 (2.27)	0.70 (3.88)	0.67 (3.02)	1.06 (3.69)	0.23 (0.61)
C-RP	0.02 (0.28)	0.02 (0.29)	0.005 (0.08)	0	0.05 (0.61)	1.06 (2.71)	0.03 (0.24)	0.03 (0.27)	0	0.08 (0.54)	2.73 (2.04)	2.62 (0.71)	0.85 (1.65)	0.38 (0.84)
Electrolytes Sediment rate	0.01 (0.15)	0.01 (0.015)	0.003 (0.05)	0	0	0.17 (0.40)	0.004 (0.02)	0.01 (0.09)	0.004 (0.04)	0.02 (0.18)	0.09 (0.41)	0.18 (0.54)	0.06 (0.37)	0.09 (0.14)
Procalcitonin	0.09 (1.19)	0	0	0	0	0	0.005 (0.07)	0	0	0	0	0	0	0.12 (0.15)
Urea of Creatinine	0.03 (0.37)	0.03 (0.39)	0.01 (0.08)	0.010 (0.13)	0.36 (0.29)	0.05 (0.29)	0.006 (0.04)	0.02 (0.22)	0.01 (0.06)	0.01 (0.09)	0.06 (0.47)	0.02 (0.29)	0	0.04 (0.15)
Electrolytes	0.04 (0.56)	0.04 (0.58)	0	0	0.36 (1.96)	0.07 (0.46)	0.006 (0.05)	0	0.01 (0.12)	0	0.05 (0.50)	0	0	0.06 (0.22)
Sputum for culture	0.05 (0.68)	0.16 (1.22)	0	0	0	0	0.003 (0.04)	0	0	0	0	0.13 (0.21)	0	0.13 (0.34)

	Cardiff (n=181)	**Southampton (n=168)**	**Utrecht (n=195)**	**Barcelona (n=169)**	**Mataro (n=179)**	**Rotenburg (n=181)**	**Balatonfured (n=320)**	**Antwerp (n=164)**	**Lodz (n=221)**	**Milan (n=153)**	**Jonkopin (n=222)**	**Tromso (n=148)**	**Helsinki (n=90)**	**Bratislava (n=299)**
Blood for Serology	0.11 (1.51)	0	0	0.04 (0.50)	0	0	0.007 (0.08)	0	0	0	0.64 (2.95)	0.09 (1.15)	0	0.23 (0.68)
Nose swab	0.15 (2.05)	0	0	0.06 (0.83)	0	0	0.01 (0.15)	0	0.03 (0.41)	0	0.82 (4.55)	0.87 (4.66)	0.38 (3.65)	0.62 (1.48)
Spirometry	0.12 (1.67)	0.13 (1.73)	0.72 (5.01)	0.18 (1.61)	0	1.17 (4.05)	0.01 (0.11)	0.18 (1.60)	0.02 (0.25)	0.10 (0.89)	0.07 (1.05)	0.76 (3.00)	0	0.13 (0.55)
Total mean costs	30.79 (17.27)	42.05 (40.61)	23.93 (21.37)	26.77 (11.38)	31.83 (61.06)	38.36 (48.63)	16.35 (23.74)	69.27 (82.56)	16.49 (8.61)	25.92 (52.63)	79.72 (23.47)	74.48 (50.73)	57.84 (22.80)	17.60 (17.77)
Total mean costs (PPP)	29.62 (17.27)	41.03 (39.43)	23.78 (21.24)	31.85 (13.55)	37.89 (72.66)	38.36 (48.63)	14.14 (9.49)	73.46 (84.56)	11.34 (6.93)	27.47 (53.60)	50.48 (16.65)	55.47 (39.37)	49.65 (19.86)	16.41 (25.20)
Loss of productivity	152.76 (254.01)	146.11 (178.19)	213.41 (399.18)	121.72 (151.57)	102.02 (114.12)	240.79 (276.01)	76.13 (83.80)	296.28 (435.43)	79.82 (113.62)	162.57 (168.88)	274.19 (294.21)	304.80 (310.28)	290.51 (374.96)	93.80 (112.27)

**Length of GP Visits (mins): Cardiff (7.25); Southampton (11.5); Utrecht (10); Barcelona/Mataro (7.5); Rotenburg (7.5); Balatonfured (6); Antwerp (10); Lodz (13); Milan (10); Jonkopin (17) Tromso (20); Helsinki (20); Bratislava (10)

5.3.5 Days off work and loss of productivity

The mean days lost from work was 3.08 days for all networks. However, this figure varied considerably across networks. In Lodz, Bratislava and Southampton, this figure was 5.12, 4.82 and 2 days respectively. Whilst the mean number of days lost in Mataro, Utrecht and Barcelona was 0.89, 1.73 and 1.54 respectively. In monetary terms, the loss in productivity as a result of days off work for the networks also varied. It ranged from £76.13 and £79.82 in Balatonfured and Lodz to £304.80 and £290.51 in Tromso and Helsinki (Table 5.5).

5.3.6 Sensitivity analysis

The total costs obtained from the base case analysis and the multilevel modelling approach yielded similar results. In most cases with the exception of the following networks: Mataro, Lodz, Tromso, Milan and Helsinki, the costs associated with the multilevel modelling approach were higher than the costs obtained from the base case analysis (Table 5.6). Similar to the results that were obtained with the base case analysis, the multilevel modelling approach also indicated that with normal exchange rates the highest cost network was the Swedish network (Jonkoping) however, the lowest cost network with this approach was the Polish network (Lodz). On the other hand, and similar to the result obtained with the base case analysis, when PPP exchange rates were used, the highest cost was recorded in the Belgian network (Antwerp) whilst the lowest cost network was the Polish network (Lodz).

Table 5.6 Total mean costs £ (Standard deviation)

Country	**Wales n=181**	**England n=168**	**Netherlands n=195**	**Spain n=169**	**Spain n=179**	**Germany n=181**	**Hungary n=320**	**Belgium n=164**	**Poland n=221**	**Italy n=153**	**Sweden n=222**	**Norway n=148**	**Finland n=90**	**Slovakia n=299**
Network	Cardiff	Southampton	Utrecht	Barcelona	Mataro	Rotenburg	Balatonfured	Antwerp	Lodz	Milan	Jonkopin	Tromso	Helsinki	Bratislava
Total mean costs (SD)	30.79 (17.27)	42.06 (40.61)	23.93 (21.37)	26.77 (11.38)	31.83 (61.06)	38.36 (48.63)	16.35 (23.74)	69.27 (82.56)	16.49 (8.61)	25.92 (52.63)	79.74 (23.47)	74.48 (50.73)	57.84 (22.80)	17.60 (17.77)
Difference (Bootstrapped CI)[a]		11.27 (3.82, 20.91)	-6.86 (-12.34, -0.36)	-4.02 (-9.43, 1.10)	1.05 (-8.91, 21.76)	7.58 (-2.12, 26.44)	-14.44 (-18.88, -9.87)	38.48 (16.73, 71.18)	-14.29 (-18.06, -11.15)	-4.87 (-11.86, 6.92)	48.95 (41.77, 55.95)	43.70 (29.11, 60.67)	27.05 (20.01, 36.24)	-13.19 (-17.30, -9.38)
Total mean costs (SD) [c]	29.62 (17.27)	41.03 (39.43)	23.78 (21.23)	31.85 (13.55)	37.89 (72.66)	38.36 (48.63)	14.14 (9.49)	73.46 (84.56)	11.34 (6.93)	27.47 (53.60)	50.48 (16.65)	55.47 (39.37)	49.65 (19.86)	16.41 (25.20)
Difference (Bootstrapped CI)[a]		11.41 (4.16, 21.06)	-5.85 (-11.32, 0.39)	2.23 (-3.64, 8.07)	8.26 (-3.38, 32.61)	8.74 (-1.06, 27.32)	-15.49 (-19.29, -12.06)	43.84 (21.76, 76.77)	-18.29 (-22.05, -15.25)	-2.15 (-9.19, 10.08)	20.86 (15.64, 26.13)	25.85 (14.08, 39.41)	20.03 (13.67, 27.65)	-13.21 (-17.57, 8.67)
TOTAL COSTS OBTAINED FROM MULTILEVEL MODELLING														
Total mean costs [b]	31.38	43.62	24.26	26.98	31.59	38.70	16.45	71.97	16.29	24.73	80.31	72.99	57.62	18.92
Difference (CI) [a]		12.24 (0.95, 23.52)	-7.12 (-18.21, 3.96)	-4.40 (-19.69, 10.88)	0.21 (-11.68, 12.08)	7.32 (-5.56, 20.19)	-14.93 (-24.94, -4.92)	59.28 (26.90, 54.27)	-15.05 (-24.89, -5.30)	-6.65 (-17.41, 4.11)	48.93 (36.86, 60.98)	41.61 (28.51, 54.70)	26.24 (12.80, 39.68)	-12.47 (-21.73, -3.21)
Total mean costs (PPP)[b]	30.77	43.14	24.35	32.23	37.35	38.84	14.63	77.65	11.02	25.82	51.39	54.04	49.21	18.00
Difference (CI)[a]		12.38 (0.49, 24.27)	-6.41 (-18.49, 5.66)	1.47 (-14.84, 17.77)	6.59 (-6.31, 19.48)	8.07 (-5.89, 22.04)	-16.14 (-26.80, -5.46)	68.49 (32.38, 61.39)	-19.74 (-30.72, -8.76)	-4.94 (-16.89, 7.00)	20.63 (7.69, 33.57)	23.27 (9.32, 37.22)	18.44 (4.08, 32.81)	-12.76 (-22.82, -2.70)

[a] Difference with reference to Cardiff [b] Estimates derived from hierarchical model [c] Total cost with PPP exchange rates

5.4 Discussion

General summary of results

In this chapter issues relating to costing alongside multinational studies has been considered. The costs associated with treating acute cough and LRTI in a European primary care setting have been explored and the potential methodological issues and challenges associated with such multinational costing exercises have been outlined. In general, this study found that in practice, conducting multinational costing studies of this kind is extremely problematic for a number of reasons including the lack of available unit costs and resource use data, language difficulties and barriers and a host of other problems such as lack of co-operation from study partners. A more detailed discussion specifically related to the methodological issues and the costing exercise are presented below.

Discussion on methodological issues relating to multinational costing exercises

This study revealed that there are a number of methodological challenges faced by researchers conducting these multinational studies, many of which are normally underestimated at the proposal as well as the design stage of a project. The most problematic challenge that was encountered and which has been discussed by other studies (Raikou et al. 2000, Pang, 2002) relates to the identification of information for use in costing care. Even though a huge effort was made in this study to identify relevant costs (e.g. the four stage approach plus other approaches which were adopted in cases where unit cost data were still unavailable), there were still a number of aspects of

resource use for some countries where no actual unit cost data were found and as a result, assumptions had to be made in order to come up with the unit costs. However, it remained unclear as to whether this reflected a lack of data in the countries concerned or a lack of access to such data although, it is very likely that a combination of both contributed to the unavailability of data.

Out of all the four approaches that were used for the identification of costs, it was shown that the most effective method of acquiring unit cost information across countries appeared to be by contacting researchers working in the various countries. Clearly, this requires patience and forbearance on the part of the researcher being contacted as well as their willingness to help out. With an increase in multinational studies, it is possible that such attempts at contact will be less welcomed, especially in countries where the number of relevant researchers is relatively limited. Even the experience of contacting the national network coordinators and national network facilitators in each site did not yield the desired response although it was anticipated that they would respond positively given they are members of the wider GRACE project team. It should be acknowledged that the national network coordinators and national network facilitators are not economic specialist researchers and would typically not have had any previous experience of conducting or undertaking costing exercises or any form of economic costing and they might have found the task difficult. This was confirmed by a national network coordinator who stated that the costing exercise was very demanding.

There was also the problem of interpreting different languages which was a major hurdle. Most databases were, unsurprisingly, in the local language. For example, the German and Belgian drugs databases were all in the local language and so full and effective

interpretation of the data within as well as navigation of the databases was often problematic. Bilingual dictionaries, translation software such as Google translate and the use of contacts who speak the language fluently were all methods used to try and overcome this problem. Help was obtained for the interpretation of the German and Belgian drug databases in order to ensure that the correct values were used. However, errors may have remained in the interpretation of these data. Resource use data collection and interpretation was not, however, problematic as it was all collected and translated within study.

It should be acknowledged that more advanced statistical approaches for the prediction of costs could be developed. Approaches such as multiple imputation have been suggested as alternative approaches to estimating centre specific unit costs (Grieve et al. 2010). However, this approach was considered within the context of multi-centre studies as opposed to multinational studies and it is therefore difficult to comment on the appropriateness of this approach for multinational trials until it can be empirically investigated. Furthermore, complex statistical approaches are normally based on a number of assumptions and may not yield accurate results. Another issue which is often ignored in the literature is the choice of conversion rates. The results obtained clearly showed that the methods used for the conversion of costs to a common currency is quite important and as we have clearly seen from this study, different conclusions with respect to which is the most expensive and which is the least expensive network were obtained from the exchange rates and the PPP exchange rates. Although exchange rates have their advantages, there are a number of issues that arise when they are used for the conversion of costs. The most important being the fact that they do not necessarily reflect the relative

purchasing power of different currencies. PPPs on the other hand have the advantage of equalizing the purchasing power of different currencies and thus they may be more appropriate to use for conversion of unit costs. In addition to this, PPP exchange rates cannot be manipulated by governments who tend to enforce/fix exchange rates. PPP exchange rates are therefore more suited for the purpose of comparison (Buxton et al. 2004).

Discussion on findings related to costs associated with treating acute cough/lower respiratory tract infections

This chapter also examined the costs and resource use associated with the treatment of lower respiratory tract infections in Europe and found that the mean costs of treating lower respiratory infections differed across geographical regions in Europe. The Nordic countries (Sweden, Norway and Finland) recorded the highest mean total cost whilst the eastern European countries recorded lower costs. The difference is probably due to the lower GDP in eastern European countries as well as the high cost of health care in the western European countries. Other factors that could explain this variation in mean cost include health system features and regional factors. In the Nordic countries, investigations such as point of care C-reactive protein are carried out more often than in other parts of Europe, possibly because, GPs in Norway and Finland are paid extra when they carry out additional investigations (Grytten and Sorensen, 2001) and thus, have a financial incentive to perform these tests. The length of time associated with a visit to the GP was a major contributory factor to mean costs and this cost varied across networks. In the Nordic countries, the mean length of consultation ranged between 17 minutes in Sweden and 20 minutes in Finland and Norway. The higher consultation times in the

Nordic countries are associated with incentives that have been put in place. In Norway for example, to ensure chronically ill patients are treated properly, GPs are paid extra if the consultation with the patient lasts for more than 20 minutes (Grytten and Sorensen, 2001). On the other hand the consultation times in the Eastern European countries were much lower, with the mean consultation period ranging from 6 minutes in Balatonfured to 13 minutes in Lodz.

The results of the costing exercise showed that approximately 53% of patients included in this study received an antibiotic prescription after the first visit to the GP. Given that GPs in all networks stated that only around 10% of patients in all networks asked for an antibiotic prescription and their perception was that approximately 30% of patients wanted them to prescribe an antibiotic, this does not suggest (as has been proposed previously) (Scott et al. 2001) that the main cause of high prescription levels for this largely self-limiting condition is a perception on the part of GPs that patients require them to prescribe. It may be that interventions to reduce prescription of antibiotics should be targeted at GPs rather than at patients.

There are some potential limitations in relation to the estimation of costs associated with acute cough/LRTIs. First, the study period for this project was only 4 weeks and thus some effects, for example prolonged hospitalisation, may not have been captured in this analysis. It is likely, however, that any such effect would be minimal, due to the relatively small number of high cost long-term events and the short duration of most cases of lower respiratory tract infections. Second, these data were collected in an observational study. Thus, although different patients followed different care pathways, these are largely confounded by primary care network and country (for example, with the

Scandinavian networks having a greater tendency towards use of CRP investigations and longer consultation times, and Southampton having a greater tendency towards delayed prescription). Third, the costing study was limited to those who had both complete CRF and patient diaries. This could be a potential source of bias since non responders may have deteriorated more than responders. There is therefore the possibility that patients with severe respiratory illness may have been excluded from the analysis. Also, approximately 19% who had information from the CRF had to be excluded as a result of the non-completion of the patient diary. Fourth, standard questionnaires (CRF and patient diary) were used to collect data across the various countries and as a result, it may be possible that important resource use or work-related data in some countries may have been missed.

5.4.1 Suggestions and future work

Even though some of the problems associated with multinational trial costing exercises have been discussed by other studies (Pang, 2002, Raikou et al. 2000, Virk 2008, Schulman et al. 1998, Wordsworth et al. 2005) there seems to be little consensus with respect to the solutions to these problems which could be an indication that the problems might not have been fully understood by researchers. The easiest/simplest solution to the problem associated with obtaining unit cost would be to apply a one country costing approach which would potentially avoid most of the problems that are associated with searching for unit cost data. However, it is quite clear that this is not quite an acceptable approach because it would affect the relationship between resource use and cost and the approach which has been recommended is to apply country-specific unit costs when undertaking costing exercises alongside multinational studies (Reed et al. 2005).

As mentioned earlier, most of the costing issues that were identified are normally underestimated and should have been considered at the study design and proposal stages of the project. Two suggestions are proposed:

The first is that economists should endeavour to be involved in the selection of recruitment centres and countries. This would enable them to determine whether costing exercises would prove difficult or not. However, it should be recognised that economic questions are not the main motivation for setting up and conducting clinical trials and this might limit the influence economists have on the choice of any study centre. Also, some countries/centres which may be preferred by health economists (i.e. countries where unit costs are available) may not be willing to participate. Clinicians, statisticians and other specialists are more concerned about achieving the required sample size as opposed to selecting centres and countries with information on costs and might push to recruit patients from centres/countries even though health economists might not prefer these centres.

The second suggestion from this research is for the setting up of a website, not dissimilar to those focusing on particular outcome measures such as EuroQol (EQ-5D), but that provides instead information about cost sources for the different countries in the EU and across the world. The problem with this approach is that it might be difficult to set this up due to funding constraints. From the experience of this study, it is suggested that even if the above approaches (economists involvement in the start up of projects and setting up websites) do not yield the desired results, economic researchers should endeavour to hold discussions with project representatives in the various countries on how best unit cost information can be obtained in the respective countries. In this case, discussions could

have been held with national network coordinators and national network facilitators in the various countries before the start of the project to ensure that adequate unit costs are collected.

5.5 summary

Chapter 5 explored costing alongside multinational studies with a focus on exploring methods for obtaining unit costs alongside multinational studies. Four different approaches: (i) internet search (ii) email to HESG members (iii) contact with national network coordinators and facilitators in participating countries and (iv) direct contact with health economists and researchers in participating countries were assessed in terms of how successful they were in terms of obtaining unit cost data. The results of the study showed that direct contact with health economists and researchers was the most effective way of obtaining unit cost data. The next chapter (Chapter 6) now considers issues related to using the EQ-5D to value health stated in a multinational trial setting.

CHAPTER 6 HEALTH OUTCOMES IN ECONOMIC ANALYSIS ALONGSIDE MULTINATIONAL STUDIES: AN EXPLORATION OF THE CHOICE OF EQ-5D VALUE SETS

6.1 Introduction

Economic evaluation was defined as the comparative analysis of interventions in terms of costs and outcomes in Chapter 1 and in the previous chapter, issues relating to costing alongside multinational trials were explored. In this chapter, the outcome side of economic evaluation is considered with a focus on issues relating to the 3-level version of the EQ-5D (EQ-5D-3L). From the review presented in Chapter 3, it was established that one of the unresolved issues associated with the conduct of economic evaluations alongside multinational studies relates to the choice of EQ-5D-3L value sets/tariffs (Knapp et al. 2008).

The work presented in this Chapter has two main aims. The first is to assess the construct validity of the EQ-5D-3L in patients presenting with acute cough/LRTI and the second aim is to compare EQ-5D-3L scores obtained with the European value set (EVS) with those obtained from country specific value sets (CVS) and the United Kingdom's value set (UKVS) and to explore the impact of between-value-set discrepancies on the estimation of cost-effectiveness in multinational studies. This analysis will draw upon data from the observational study conducted alongside the

GRACE project (workpackage 8). The UKVS was chosen as a separate comparator because of its frequent use in non-UK studies and as was established in Chapter 3, the UKVS is the most frequently used value set in multinational studies. The results presented in this chapter have been published previously in the following peer reviewed papers: (1) *Oppong et al. (2013) The Impact of Using Different Tariffs to Value EQ-5D health state descriptions: An example from a study of acute cough/LRTI in seven countries European Journal of Health Economics 14(2) pp 197-209 and (2) Oppong et al. (2011) Assessment of the Construct Validity of EQ-5D in patients with acute cough/LRTI Applied Research in Quality of life 6(4) pp 411-423.*

Under an extra-welfarist perspective, the economic evaluation of interventions ideally requires the use of a generic preference based measure that can provide a single index for overall health to allow calculation of QALYs and the EuroQol's EQ-5D measure provides such an index (EuroQol, 1990). This measure comprises a visual analogue scale (VAS) with which individuals can rate their own health state and a descriptive system with five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), each of which is sub-divided into 3 levels: no problems, some problems and severe problems, resulting in the definition of 243 (3^5) separate states (Morris et al. 2007). States are defined by a number representing the level on each of the five dimensions. For example state 22213 would be a state with some problems in terms of mobility, self-care and usual activities, no pain and severe anxiety/depression. State 11111 is equivalent to full health (value equals 1), and dead is given a value of 0. Individual self-rated assessment of their health on this measure can then be used to generate a single health related quality of life value (Kind, 2003). To generate this index for all 243 states, value sets for individual countries have been developed. Currently, there are EQ-5D-3L value sets for 13 different countries

(Belgium, Denmark, Finland, France, Germany, Japan, New Zealand, Netherlands, Slovenia, Spain, UK, USA and Zimbabwe) as well as an European value set (EuroQol accessed on 15th August, 2015). However, when conducting economic evaluations alongside multinational trials, there is no consensus with respect to the choice of value set. It is therefore important to study the implications of the choice of EQ-5D value sets when conducting economic evaluations alongside multinational trials.

In some cases, generic measures such as the EQ-5D-3L have been shown to be inappropriate or to lack relevance to specific disease areas (Barton et al. 2004; Haywood et al. 2008) and may not fully capture all the relevant changes in health related quality of life in specific disease groups (Cleemput et al 2004). It is therefore important for the EQ-5D to be validated in the context of any disease area or population in which it is to be used. Since the case study is limited to a particular disease area (acute cough/lower respiratory tract infections), it is important to initially test the validity of the instrument in this patient population. This study therefore starts by conducting a validation exercise to determine whether the EQ-5D-3L is appropriate for valuing quality of life in patients with acute cough/LRTI.

6.2 Aim 1: Assessing the construct validity of the EQ-5D-3L in patients with acute cough/lower respiratory tract infections

6.2.1 Methods

Study Population

Validation was conducted alongside the GRACE observational study of the management of patients with cough and LRTI in primary care i.e. Workpackage 8 described in Chapter four (Butler et al. 2009) and was limited to patients from the UK (England and Wales) to avoid problems regarding choice of EQ-5D-3L tariff in multinational trials context. Two main sources of data were used: a patient completed diary and a case report form (CRF) completed by a clinician.

Data collection: Patient Diary

Baseline and other data were collected over four weeks using the patient diary which was given to patients on entry to the study. Data were collected for EQ-5D-3L, symptom severity scores and socio-demographic characteristics.

EQ-5D-3L

Patients were asked to complete the EQ-5D-3L questionnaire at baseline and then weekly for four weeks or until they felt better. In this study, the UK tariff (Dolan, 1997) was obtained from the EuroQol group and was applied to data from the UK (England and Wales) that were included in the study to generate EQ-5D-3L index scores. EQ-5D index scores range from -0.594 (the worst health state) to 1.000 (full health). Zero represents dead. Since data were not missing at random and patients were specifically told to stop the completion of the questionnaire when they got well,

multiple imputation was not considered to be appropriate for dealing with the missing data. The last value carried forward method was therefore used to estimate EQ-5D-3L scores for patients who stopped completing the EQ-5D-3L questionnaire before week 4.

Symptom severity scores

Patients completed a symptoms diary on a daily basis for 28 days or until they felt well. Using this diary, they were asked to score the severity of symptoms associated with LRTI (cough, phlegm, shortness of breath, wheezing, blocked nose, chest pain, fever, muscle ache, headache, disturbed sleep, feeling unwell, interference with normal activities, and interference with social activities). The scores on this instrument ranged from 0 to 6 for each item, with higher values indicating worse symptoms. To develop a symptom severity score for each day, the 13 symptoms were summed and then the resultant score was scaled to range between 0 and 100 (Butler et al. 2009). Similar to the approach used with the EQ-5D, the last value carried forward approach was used to estimate symptom severity scores for individuals who got well before day 28, i.e. if severity scores fell to zero.

6.2.2 Data collection: Case report form

Primary care physicians collected a variety of data at baseline using the CRF. Physicians indicated whether a patient had co-morbidities as well as the existence of chronic conditions such as asthma, COPD and diabetes. Physicians also recorded the use of medical investigations and their choice about whether to prescribe antibiotics.

6.2.3 Data analysis

Descriptive statistics were employed to summarise the main variables. The Chi-square test was used to examine the relationship between individual dimensions of the EQ-

5D (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and items in the patient diary such as chest pain, disturbed sleep, interference with normal activities and interference with social activities. To ascertain the construct validity of EQ-5D-3L in patients with acute cough/LRTI, a number of *a priori* hypotheses were developed with the aid of literature in the area of LRTI and tested. The following six hypotheses were tested:

Hypothesis 1: EQ-5D-3L scores should increase if symptom severity scores decrease

It was expected that there would be an inverse relationship between symptom severity scores and EQ-5D-3L. As noted earlier, increasing symptom severity scores indicate worsening health whilst increasing EQ-5D-3L scores indicate improvements in health. Tobit regression analysis was used to explore the relationship between symptom severity scores and EQ-5D-3L at baseline and weeks one to four. EQ-5D-3L was the dependent variable and symptom severity scores the independent variable. Comorbidities were also included as dependent variables. Since EQ-5D-3L was collected weekly and symptom severity scores were collected on a daily basis, EQ-5D-3L scores were compared to the corresponding symptom severity scores. Thus, EQ-5D-3L scores at baseline was compared with symptom severity scores at day 1 and EQ-5D-3L scores at week 4 was compared with symptom severity scores at day 28.

Hypothesis 2: Smoking is associated with worse EQ-5D-3L scores

It was anticipated that smokers would have worse health outcomes (lower HRQol) than non-smokers for this respiratory condition as reflected in their EQ-5D-3L scores. A study conducted in the UK population revealed that smokers reported worse outcomes on all dimensions of the EQ-5D-3L than non-smokers (Kind et al. 1998). It has also been shown that that there is a positive relationship between smoking and

antibiotic prescribing in patients with acute cough/LRTI (Stanton et al. 2010).This indicates that smokers are likely to have worse symptoms than non-smokers. Ex-smokers are also more likely to have worse outcomes than people who have never smoked. This is because in some cases, ex-smokers might have given up their smoking habits as a direct result of ill health (Mulder et al. 2001). This difference was tested using Kruskal-Wallis tests, to allow for separate groupings for smokers, ex-smokers and non-smokers.

Hypothesis 3: Chronic diseases are associated with worse EQ-5D-3L scores

It was expected that chronic diseases such as asthma and diabetes would be associated with worse outcomes in patients with acute cough/LRTI. Research has shown that patients with chronic diseases like diabetes have an increased risk of developing severe LRTI (Venmans et al. 2008). Patients who reported with at least one underlying chronic disease such as asthma, COPD and diabetes were compared with those that did not report any chronic disease. The difference in EQ-5D-3L scores between the groups was tested using the Mann Whitney test.

Hypothesis 4: EQ-5D-3L scores increase over time

It was anticipated that EQ-5D-3L scores would increase over time in patients presenting with acute cough/LRTI. This is for two main reasons. First, quite a number of lower respiratory tract infections are self-limiting (Van Duijn et al. 2005) and so patients would be expected to improve in health status over the four weeks from their inclusion in the study irrespective of intervention. Second, for those patients prescribed some sort of treatment, this would be anticipated to improve their quality of life (symptom severity score and EQ-5D-3L) although recent results suggest that there is only a very small difference in recovery between patients who receive treatment and those who do not receive any treatment (Butler et al. 2009). The

Wilcoxon sign rank test was used to test the difference between EQ-5D-3L scores at baseline and week four. The responsiveness of EQ-5D-3L was also tested using the standardized response mean (SRM) which is the ratio of the mean change to the standard deviation of the change score (Streiner and Norman 2003). It has been suggested that SRM values of 0.20, 0.50 to 0.80 and 0.80 and above indicate small, moderate and large responsiveness respectively (Husted et al. 2000).

Hypothesis 5: EQ-5D-3L scores and patient reported recovery

The final a priori hypothesis that was tested was that EQ-5D-3L scores should increase if the patient self assessed themselves as being well. In order to test this relationship, mean EQ-5D-3L scores of patients who reported that they felt well on days 7, 14, 21 and 28 were compared to their EQ-5D-3L scores at baseline. The difference was tested using the Wilcoxon sign rank test.

Hypothesis 6: Known group validity

The ability of EQ-5D-3L to discriminate between patients with particular symptoms was tested. The characteristics tested were three specific symptoms (headache, interference with normal activities and disturbed sleep). The particular symptoms were chosen from those available within the symptom diary as being particularly likely to be related to some dimensions of the EQ-5D-3L (headache to pain; interference with normal activities to usual activities; and disturbed sleep to anxiety/depression and pain/discomfort). In considering the ability of the EQ-5D-3L to discriminate, it is helpful to choose those symptoms that are most likely to be related to EQ-5D-3L scores. The ability of EQ-5D-3L to discriminate between patients with signs and symptoms that were related to acute cough/LRTI was also tested. The signs and symptoms that were chosen include fever, chest pain and shortness of breath (Macfarlane 1993; Williams et al. 2004). The most obvious

symptom of LRTI (cough) was excluded from the symptoms that were studied because most patients (approximately 95%) presented this symptom. Differences in the EQ-5D-3L score between the groups with and without the symptom were tested using the Mann-Whitney test.

Data analysis in this study was carried out using STATA version 10 (STATA 2010).

6.2.4 Results

Patient characteristics

A total of 349 patients who completed both the CRF and patient diary in the UK (Wales and England) were included in the analysis. Of these, 201 (58%) were female and the mean (standard deviation) age of the population was approximately 51 (16.44) years. Overall, 23% of the study population were smokers.

Relationship between dimensions of the EQ-5D-3L and the symptoms diary at baseline

With regards to individual items on the EQ-5D-3L index, most patients reported no problems with self care. Around a quarter of patients reported some problems with mobility and anxiety/depression and more than half of patients reported problems with usual activities and, particularly, pain/discomfort, with 64% reporting some problems and 9% reporting severe problems with the latter (Table 6.1).

Table 6.1 Responses to dimensions of the EQ-5D-3L

Dimensions	No problems (%)	Some problems (%)	Extreme problems (%)
Mobility	240 (73.4)	85 (26.0)	2 (0.6)
Self-care	303 (93.2)	22 (6.8)	0 (0)
Usual activities	156 (47.7)	145 (44.3)	26 (8.0)
Pain/discomfort	90 (27.3)	211 (63.9)	29 (9.0)
Anxiety/depression	239 (72.6)	78 (23.7)	12 (3.8)

**This table presents the data for the various levels of each of the EQ-5D dimensions. This analysis was limited to complete cases because imputation was not carried out at the various levels of the EQ-5D dimensions.*

With regard to the relationship between dimensions of the EQ-5D-3L and individual items of the symptoms diary, there was a statistically significant relationship between usual activities and all the individual items of the patient diary with the exception of cough, wheeze and blocked nose. Pain/discomfort was also significantly related to all items of the diary except headache, cough, wheeze and blocked/runny nose. There was also a statistically significant relationship between mobility on the EQ-5D-3L and shortness of breath, wheeze, blocked nose, feeling generally unwell and chest pain. Anxiety and depression was also significantly related to most items on the symptoms diary. On the other hand, self-care did not show a strong relationship with most individual items of the symptoms diary. For example, self-care was only significantly related to shortness of breath (Table 6.2).

Table 6.2 Relationship between dimension on the EQ-5D-3L and individual items on the Symptoms diary (P-Values) using Chi squared test

	EQ-5D-3L Dimensions				
Symptoms	Mobility	Self care	Usual activities	Pain/discomfort	Anxiety/depression
Cough	0.480	0.544	0.746	0.684	0.000
Phlegm	0.142	0.395	0.007	0.010	0.278
Shortness of breath	0.002	0.016	0.000	0.001	0.002
Wheeze	0.011	0.295	0.125	0.047	0.019
Blocked/runny nose	0.043	0.561	0.649	0.303	0.621
Chest pain	0.021	0.486	0.000	0.000	0.100
Fever	0.978	0.107	0.000	0.030	0.880
Muscle aching	0.053	0.197	0.000	0.000	0.749
Headache	0.849	0.832	0.025	0.179	0.895
Disturbed sleep	0.147	0.058	0.001	0.003	0.020
Feeling generally unwell	0.032	0.128	0.000	0.000	0.000
Interference with normal activities	0.119	0.125	0.000	0.000	0.040
Interference with social activities	0.072	0.105	0.000	0.000	0.030

Hypothesis 1: EQ-5D-3L scores and Symptom severity scores

The regression analysis supported the hypothesis that there was a negative statistically significant relationship between symptom severity scores and EQ-5D-3L scores at baseline and over the four week period. All p-values obtained from the analysis were less than 0.01 (Table 6.3).

Table 6.3 Regression results: Relationship between EQ-5D-3L and Symptom severity scores

	N	Coefficient	95% Confidence Interval		P-value
Baseline	307	-0.009	-0.010	-0.007	<0.01
Week 1	306	-0.0011	-0.013	-0.008	<0.01
Week 2	291	-0.014	-0.017	-0.010	<0.01
Week 3	274	-0.019	-0.024	-0.015	<0.01
Week 4	265	-0.026	-0.033	-0.018	<0.01

**It should be noted that the difference in the sample sizes used across the four week period is a result of missing symptom severity scores. Unlike the EQ-5D data, symptom severity scores were not imputed.*

Hypothesis 2: EQ-5D-3L scores and smoking

EQ-5D-3L scores obtained from patients who did not smoke were consistently higher than the scores that were obtained in patients who were ex-smokers or who smoked. In addition, ex-smokers also had higher EQ-5D-3L scores than those who smoked (Table 6.4).

Hypothesis 3: EQ-5D-3L scores and chronic diseases

As was expected, individuals who had underlying chronic conditions (asthma, COPD and diabetes) had worse EQ-5D-3L scores than those who did not have any of the chronic diseases. The difference was statistically significant (Table 6.5).

Table 6.4 Relationship between Smoking and EQ-5D-3L [Mean (SD)]

	N	Baseline (N=318)	Week 1 (N=318)	Week 2 (N=318)	Week 3 (N=318)	Week 4 (N=318)
Smoker	79	0.613 (0.324)	0.656 (0.322)	0.722 (0.321)	0.756 (0.327)	0.788 (0.314)
Non-smoker	169	0.742 (0.241)	0.787 (0.242)	0.845 (0.240)	0.862 (0.246)	0.881 (0.247)
Ex-smoker	101	0.711 (0.241)	0.758 (0.234)	0.833 (0.218)	0.852 (0.215)	0.862 (0.210)
P-values (Kruskal-Wallis test)		<0.05	<0.01	<0.01	<0.05	<0.05

** A total of 318 patients recruited from the UK had complete EQ-5D data at baseline and the last value carried forward method was used to impute missing EQ-5D scores over the 4 week period. As a result the total sample used for the assessment of the relationship between EQ-5D and smoking status in the study was 318*

Table 6.5 Chronic diseases and EQ-5D-3L scores

Chronic Disease present	N	EQ-5D-3L
No	250	0.717
Yes	99	0.669
Difference		0.048*

** significantly different at 5% level*

***significant at 1% level Mann-Whitney test*

Hypothesis 4: EQ-5D-3L scores over time

There was an increase in EQ-5D-3L scores over the four week period. The standardized mean (SRM) for the change score was 0.552. This indicates that the EQ-5D-3L is moderately responsive (Table 6.6).

Hypothesis 5: Patient reported recovery and EQ-5D-3L scores

Patients who reported that they recovered on days 7, 14, 21 and 28 had higher EQ-5D-3L scores compared to their baseline EQ-5D-3L score. These results are presented in Table 6.7.

Hypothesis 6: Discriminating between known groups

When the factors 'headache', 'interference with normal activities' and 'disturbed sleep' were considered, higher EQ-5D-3L scores were recorded for the groups without the condition. With the exception of headache, the resulting differences were statistically significant (Table 6.8). Individuals who did not report headache, interference with normal activities and disturbed sleep had mean EQ-5D-3L scores of 0.737, 0.825 and 0.811 respectively whilst those who did report these symptoms had scores of 0.678, 0.658 and 0.689 respectively. With regards to symptoms that were associated with acute cough/LRTI, patients who had chest pains, fever and shortness of breath had lower EQ-5D-3L scores than patients who did not report these signs and symptoms. The difference between individuals with and without the condition was significant in all cases (Table 6.8).

Table 6.6 EQ-5D-3L scores over time and sensitivity to change (standardized response mean)

Baseline (N=318)	**Week 1 (N=318)**	**Week 2 (N=318)**	**Week 3 (N=318)**	**Week 4 (N=318)**	**Difference from baseline to week 4**	**Standardized response mean**
0.704	0.749	0.814	0.836	0.855	0.151**	0.552

**** Significantly different at 5% level***

***** Significantly different at 1% level Wilcoxon signed rank test***

Table 6.7 Patient reported recovery and EQ-5D-3L scores

Recovery on Day 7		
	N	**EQ-5D-3L**
Baseline	14	0.703
Day 7	14	0.880
Difference		-0.177
Recovery on Day 14		
	N	**EQ-5D-3L**
Baseline	17	0.714
Day 14	17	0.867
Difference		-0.153*
Recovery on Day 21		
	N	**EQ-5D-3L**
Baseline	11	0.682
Day 21	11	0.876
Difference		-0.194*
Recovery on Day 28		
	N	**EQ-5D-3L**
Baseline	12	0.731
Day 28	12	0.885
Difference		-0.154**

** Significantly different at 5% level*

***significant at 1% level Wilcoxon signrank test*

Table 6.8 Known group validity

	Headache		*Interference*		Disturbed sleep	
	N	**EQ-5D-3L**	**N**	**EQ-5D-3L**	**N**	**EQ-5D-3L**
No	150	0.737	95	0.825	61	0.811
Yes	199	0.678	253	0.658	288	0.689
Difference		0.059		0.167**		0.122**
	Chest pain		**Fever**		**Shortness of breath**	
	N	**EQ-5D-3L**	**N**	**EQ-5D-3L**	**N**	**EQ-5D-3L**
No	151	0.761	148	0.735	61	0.815
Yes	167	0.652	170	0.676	257	0.677
Difference		0.109**		0.059*		0.138**

** significantly different at 5% level **significant at 1% level Mann-Whitney test*

6.3 Aim 2: A comparison of the United Kingdom value set (UKVS) with the European harmonised value set (EVS) and the country-specific value set (CVS) in the context of a multinational study

As stated earlier, the second aim of this Chapter is to compare the results of using different EQ-5D-3L value sets within the context of multinational trials/studies. Most studies looking at the impact of using different tariffs have not considered the impact that using different value sets would have on the results of cost-utility analyses and the potential impact on decision making. To achieve this aim, the section starts by explaining the various value sets under consideration in this study and compares them using data obtained from the GRACE workpackage 8 data.

6.3.1 Value sets

Estimating a value set involves the following steps: (1) A description of the health states by dimensions and levels (2) A selection of a subset of health states from the total number of available health states (3) Eliciting the preferences of the general public about the subset of health states (4) Modelling the preference data obtained from the general public in order to predict the remaining health states (Tsuchiya et al. 2002).

THE UK value set (UKVS)

To obtain the UKVS, data were collected from non-institutionalised adults in England, Scotland and Wales between August and December 1993. Respondents were drawn from the national postcode address file. In a face-to-face time trade-off

exercise, each respondent valued 2 very mild states, 3 mild states, 3 moderate states, and 3 severe states drawn from a subset of 41 health states. All respondents also valued full health (11111), the most severe state (33333), as well as unconscious and dead (Dolan, 1997). After applying the exclusion criteria where respondents without complete valuation data were excluded, a total of 2997 respondents were included in the analysis. Regression analysis was used to develop a scoring algorithm for the UK. The UK model consists of variables which represent each dimension of the EQ-5D-3L with two dummy variables representing the levels of each dimension. Another dummy variable (N3), which indicates if any dimensions of the EQ-5D-3L is at level 3, was added to the model. This is commonly known as the N3 model.

Most countries that have developed country specific value sets (CVS) have followed the general form of the UK protocol (Devlin et al. 2011). These other countries include Finland, Germany, Netherlands and Spain (Ohinmaa et al. 1995; Badia et al. 1997; Greiner et al. 2005; Lamers et al. 2006). There are, however, differences in the procedures used to derive each of the value sets (Oppe et al. 2007a; Oppe et al. 2007b; Szende et al. 2007). For example, none of the studies were as large as that conducted in the UK, with numbers included in the development of value sets ranging from 294 in Spain to 1634 in Finland. Other differences from the UK valuation study include the total number of scenarios valued (ranging from 11 to 43), the method of data collection, which in some cases was by postal survey rather than face-to-face interview, the use of different 'props' to aid the collection of time trade-off data (Lamers et al. 2006), and the application of different exclusion criteria in the analysis of data (Greiner et al. 2005). Specific details of the country-specific value sets are now presented below:

Finland

Data were collected in November 1992 with questionnaires randomly distributed to members of the public by post. After data exclusion, a total of 928 persons were included in the study of whom 53.8% were female. Respondents valued a subset of 46 health states, and the health states 11111, 33333 and dead were valued twice by the respondents. The Finish model was similar to the UK model however, no value for N3 (which indicated whether at least one dimension is at level 3) was included in the Finish value set (Ohinmaa et al. 1996).

Belgium

Data for the Belgian study were collected in the summer of 2001 by mail. After exclusions, a total of 548 patients were included in the final analysis. The model used was similar to the UK model and in this case, an N3 dummy which indicates whether 1 or more dimensions is at level 3 (Cleemput, 2003; Cleemput, 2010).

Spain

Data were collected between October and December 1996, from a random sample of individuals attending a primary care setting. After exclusions, 294 respondents were included in the final analysis. The model used in the Spanish study was similar to the UK N3 model and rescaling was also carried out on the VAS scores (Badia et al., 1997).

Germany

Data were collected between October 1997 and March 1998. After exclusions, 339 respondents were included in the final analysis. This study also applied the MVH study protocol (Greiner et al., 2005), however, the exclusion criteria was slightly different. The UK N3 model was used, but the regression coefficients of usual

activities (UA), anxiety and depression (AD), self-care (S) and usual activities at level 2 (U2) were excluded from the model because they were not significant.

Netherlands

In the Netherlands, data were collected using quota sampling from patients between the ages of 18 and 75 and after applying the exclusion criteria, a total of 298 respondents were included in the analysis. The Dutch study was a replica of the MVH study and applied the MVH study protocol. Respondents filled in the EQ-5D-3L questionnaire and described their own health states, then ranked 17 health states including 11111 and immediate death. The 19 states were put on a visual analogue scale. Afterwards, Time trade-off (TTO) was used to value the set of 17 health states. For the Dutch study, a computer program replaced the TTO boards used in the UK study. The study applied the UK N3 model (Lamers et al., 2006).

European value set (EVS)

A further combined value set has been generated to aid the comparison of health state values in multinational studies, and is known as the European value set (EVS). In generating this tariff, visual analogue scale (VAS) data were obtained from 11 studies across the six European countries with more than one study used from some countries such as Germany (Greiner et al. 2003). Once respondents with incomplete data and respondents with inconsistent responses were excluded, 6870 respondents were included in the analysis. To generate the EVS, a multi-level random effects model with valuations of health states (level 1) nested within respondents (level 2), and respondents nested within studies (level 3) was used (Greiner et al. 2003). A full description of the value sets is now presented in table 6.9 below.

Table 6.9 A summary of EQ-5D-3L Value sets used in this study

	UK	Finland	Belgium	Spain	Germany	Netherlands	European
Full Health	1	1	1	1	1	1	1
At least one 2 or 3 (Constant)	-0.081	-0.158	-0.152	-0.024	-0.001	-0.071	-0.128
At least on 3 (N3)	-0.269	-	-0.256	-0.291	-0.323	-0.234	-0.229
Mobility= 2	-0.069	-0.058	-0.074	-0.106	-0.099	-0.036	-0.066
Mobility= 3	-0.314	-0.230	-0.148	-0.430	-0.327	-0.161	-0.183
Self care= 2	-0.104	-0.098	-0.083	-0.134	-0.087	-0.082	-0.117
Self care= 3	-0.214	-0.143	-0.166	-0.309	-0.174	-0.152	-0.156
Usual activities= 2	-0.036	-0.047	-0.031	-0.071	-	-0.032	-0.026
Usual activities= 3	-0.094	-0.131	-0.062	-0.195	-	-0.057	-0.086
Pain/discomfort= 2	-0.123	-0.111	-0.084	-0.089	-0.112	-0.089	-0.093
Pain/discomfort= 3	-0.386	-0.153	-0.168	-0.261	-0.315	-0.261	-0.164
Anxiety/depression= 2	-0.071	-0.160	-0.103	-0.062	-	-0.062	-0.089
Anxiety/depression= 3	-0.236	-0.196	-0.206	-0.144	-0.065	-0.144	-0.129

6.3.2 Methods

Study Population

This patient population was from the same study described in the validation work presented earlier in this chapter in section 6.2, but was expanded to seven countries (Belgium, Finland, Germany, Netherlands, Spain and UK (England and Wales). These countries were selected because they had country specific value sets (CVS) available.

Data collection

The data collection process was similar to that of the validation study which has been described previously i.e. the main sources of data were the patient diary which was completed weekly over a four week period and a clinician completed case report form (CRF). The EQ-5D-3L value sets (CVS, UKVS and EVS) were obtained from the EuroQol group and applied to data from the seven countries that were included in the study to generate EQ-5D-3L index scores. The value sets in UK, Netherlands and Germany were based on TTO whilst the value sets in Belgium, Finland, Spain and the EVS was based on VAS. The patient diary included information on the patient's age, gender, and employment status as well as years spent in education. Primary care physicians indicated whether antibiotics were prescribed in the CRF.

Data analysis

Descriptive summary statistics were generated for all the main variables. Plots of the mean EQ-5D-3L scores as measured by the different value sets were constructed to examine trends at baseline. As a result of the non-normality of the EQ-5D-3L scores, the Kruskal-Wallis test (Ott and Longnecker, 2008) was applied to measure the

difference in the EQ-5D-3L scores across the seven countries. The difference between EQ-5D-3L scores generated from the different tariffs was tested using the Wilcoxon signed rank sum test. The responsiveness of EQ-5D-3L scores obtained from the CVS, EVS and UKVS was tested using the standardized response mean (SRM) which is the ratio of the mean change to the standard deviation of the change score (Streiner and Norman, 2003). Spearman correlation coefficients was used to assess association and Bland-Altman plots were used to determine whether there was agreement between the EQ-5D-3L scores obtained from the CVS, EVS and the UKVS at baseline.

To illustrate the potential impact of using different tariffs on cost-utility analyses, the study population was divided into two groups based on whether an antibiotic was prescribed (group A) or not (group B). Using the area under the curve approach (Matthews et al. 1990), quality adjusted life years (QALYs) were generated over the four week period for both groups. The difference between the QALYs that were generated using the different tariffs was tested using paired t-tests.

In order to establish the effect of using different tariffs on the results of cost-utility analysis, an assumption was made that the difference in costs between intervention A and B was £50 (in all countries). Alternative scenarios where the difference in costs between interventions A and B was raised to £100, £300 and £500 were also explored. These differences in cost were then used to generate incremental cost-effectiveness ratios (ICERs) based on the differences in QALYs that were obtained from the different tariffs.

The resulting ICERs were used to make a judgement about the cost-effectiveness of the intervention based on the existing thresholds in the various countries.

In the UK (England and Wales), the existing cost-effectiveness threshold is £20,000 to £30,000 per QALY (Appleby et al. 2007). In Netherlands, the threshold is set at €20,000 (£17,832) per QALY (Zwart-van Rijkom et al. 2000). In Finland, Belgium, Germany and Spain, and many other countries, there is no explicit threshold (Strom et al. 2007), although some thresholds have been used in different countries. For example, In Belgium a threshold of €35,000 (£31,206) per QALY has been used to inform some decision making (Jit et al. 2009). According to the World Health Organisation (WHO), for an intervention to be cost-effective, the threshold should be between per capita GDP (very cost-effective) and three times per capita GDP (cost-effective) (Hutubessy et al. 2003; Edejer, 2003; WHO, 2001). For the countries without an explicit threshold, GDP at current prices for the year 2009 using purchasing power parities were: €22,936 (£20,450) for Spain; €29,374 (£26,190) for Germany; €31,944 (£28,481) for Finland and €32,061(£28,585) for Belgium. For the purpose of this study, these values were used as cost-effectiveness thresholds in the various countries. In order to examine the effects of varying thresholds, a common threshold of £30,000 per QALY in each country (with the exception of UK and Netherlands where there is an explicit threshold) was assumed.

Sensitivity Analysis

Comparison of the UKVS and EVS in countries without country-specific value sets

Since the analysis conducted above was limited to the countries that had a country specific value set, sensitivity analysis focused on exploring the impact of using the EVS and UKVS in cases where countries do not have their own specific tariffs. This analysis was carried out using the same data from workpackage 8 (observational study described in Chapter 4) however was limited to a sub study which was restricted to

patients that were recruited from Norway and Sweden as these countries do not have country specific value sets. The overall aim of the sub study was three fold: first, to evaluate the impact of point of care CRP testing in two European countries (Sweden and Norway) on antibiotic prescribing; second, to confirm that point of care CRP did not compromise patient recovery; and third to assess the cost-effectiveness of point of care CRP for diagnosing acute cough/lower respiratory tract infections from a health service perspective (Oppong et al. 2013a). The focus of the sensitivity analysis is on the third objective, to estimate the cost-effectiveness of the test. Using the area under the curve approach, quality adjusted life years (QALYs) were estimated from EQ-5D-3L scores obtained from responses to the EQ-5D-3L questionnaire at baseline, week1, week 2, week 3 and week 4. The European value set (EVS) was used for the base case analysis. Costs were obtained by multiplying resource use items by unit costs. A net monetary benefit value (Drummond et al. 2015) was estimated for each patient using the formula *(threshold x QALY) – total cost.* Hierarchical modelling was used to obtain differences in costs, QALYs and net benefits between the intervention and control group i.e. those who received a point of care CRP test and those that did not receive the test. The primary explanatory variable was whether or not point of care CRP testing was administered. Explanatory variables/confounders in the model included: baseline EQ-5D-3L scores, baseline composite severity score, patient co-morbidities, symptoms at baseline, findings from chest auscultations and country. Model estimates of the difference in health care costs and QALYs between patients receiving and not receiving point of care CRP testing were compared. This was used to estimate an incremental cost per QALY gained (ICER) using the formula: *difference in cost/difference in QALYs* and an Incremental net benefit. In addition, EQ-5D-3L scores were also generated using the United Kingdom value set (UKVS)

and used to obtain incremental cost per QALY gained and an incremental net benefit using the approach described above.

The results obtained from the base case (EVS) were then compared to the results that were obtained with the UKVS to see if there were any differences in the cost-effectiveness of the interventions. This base case methods and results of the analysis presented here have been published previously: "*Oppong et al. (2013) Cost-Effectiveness of Point of Care CRP Testing to Inform Antibiotic Prescribing Decisions. British Journal of General Practice 63(612) pp e465- e471"*

6.3.3 Results

Patient Characteristics

This study included a total of 1327 patients who had information in both the patient diary and CRF, and were based in one of the seven countries with its own CVS. As presented in Table 6.10 below, the number of patients in each country ranged from 90 (6.8%) in Finland to 348 (26.2%) in Spain. In the study, 61.4% of the patients were female with mean ages ranging from 45.1 in Germany to 52.9 years in Wales. A very high proportion of the study population were employed while the range of the mean number of years spent in education was from 9.4 in Wales to 12.8 years in Finland.

Table 6.10 Study population characteristics

	Wales	England	Netherlands	Spain	Germany	Belgium	Finland
Total no of patients n (%)	181 (13.6)	168 (12.7)	195 (14.7)	348 (26.2)	181 (13.6)	164 (12.4)	90 (6.8)
Age mean (SD)	52.85 (16.31)	49.87 (16.49)	52.26 (16.18)	49.28 (16.49)	45.07 (15.92)	51.56 (15.44)	46.73 (14.14)
Female (%)	55.3	60.1	55.9	63.8	69.6	52.4	78.9
Male (%)	44.8	39.9	44.1	36.2	30.4	47.6	21.1
Employed (%)	90.1	88.1	84.6	95.4	93.9	90.2	95.6
Years in Education mean (SD)	9.40 (3.02)	9.66 (4.17)	9.50 (5.50)	10.28 (4.85)	9.40 (3.38)	11.30 (5.05)	12.78 (3.84)

Mean EQ-5D-3L Scores

Appendix 10 show the distribution of the EQ-5D-3L scores at baseline for each of the countries based on the three value sets used in this study. All of the EQ-5D-3L scores exhibit left skewness. As expected, the EQ-5D-3L scores obtained from value sets in all countries also showed that some patients had a score of 1 which implies full health. The percentage of respondents with EQ-5D-3L scores at full health (ceiling effect) at baseline ranged from 12.1% in Germany to 24.2% in Netherlands."

All mean EQ-5D-3L scores are shown in Table 6.11. The mean EQ-5D scores obtained using the EVS were higher than those obtained using the CVS for all countries except the Netherlands and Germany. With the exception of Spain, where the difference was not significant at baseline, week 3 or week 4, the mean EQ-5D-3L scores obtained from the CVS and EVS were significantly different in all other countries over the four week period.

At baseline, the differences between the EQ-5D-3L scores obtained from the UKVS and EVS were statistically significant for all countries except for Finland. Using both value sets, England had the highest scores whilst Finland had the lowest scores (Table 6.11). For weeks one to four, the differences between the mean EQ-5D-3L scores obtained from each of the two value sets were in all cases statistically significant. However, the direction of the differences were not the same in all cases, with the mean EVS scores being higher than UKVS in Netherlands and Germany at week one and the mean EVS score being lower than the UKVS for Finland, Belgium and Spain at weeks one to four. The highest scores (for both value sets) were recorded in Spain (weeks one to three) and Finland (for week four). The lowest scores were recorded in Wales over the four weeks. There were also statistically significant differences

between the mean scores obtained from the UKVS and CVS in Netherlands, Spain, Germany, Belgium and Finland at baseline and weeks one to four.

Association and agreement between value sets

The correlation coefficients between the EQ-5D-3L scores obtained from the CVS, EVS and the UKVS were very high in all countries, with all values above 0.87. Between the EVS and CVS, the highest correlation coefficient was 0.9 for Belgium at week 4 whilst the lowest correlation of 0.87 for Germany, also at week 4. Between the UKVS and EVS, the highest correlation coefficient had a value of 1.0 for German data at weeks 2 to 4 and Finland at weeks 1, 3 and 4, with the lowest in the Netherlands. For the comparison between the UKVS and CVS, the highest correlation of 0.99 was recorded in Belgium at week 4 whilst the lowest value of 0.91 was recorded in Germany at baseline. However, the results from the Bland-Altman plots (Figures 6.1 to 6.7c) did not suggest that there was as much agreement between the EQ-5D-3L scores obtained from the different tariffs as was suggested by the correlation coefficients. From the plots, it can be observed that there is a positive difference at lower ends of the scale however, at higher ends of the scale, the difference becomes negative. This was true when the EVS was compared with the CVS and UKVS. With the exception of Spain, when the UKVS was compared to CVS, the difference seemed to be lower at lower ends of the scale and higher at the higher ends of the scale.

EQ-5D-3L scores over time

EQ-5D-3L scores increased over the four week period which indicates that there was an improvement in health over the four week period in all countries and with all value sets (CVS, EVS and UKVS) (Table 6.11). Finland recorded the largest difference

with the UKVS whilst the smallest difference was recorded in Germany with the CVS. The difference over the four week period was significant in all countries and with all value sets. When the difference between EQ-5D-3L scores at baseline and week four was tested, the values obtained with the EVS and UKVS were significantly different in Netherlands, Spain and Germany but not in Belgium and Finland (Table 6.11). Comparing the EVS and CVS, there was a significant difference in Netherlands, Germany and Belgium but not in Finland and Spain. With the exception of Belgium and Finland, there was a significant difference between the UKVS and CVS with the largest difference being for Germany. Values from the standardized response mean (SRM) indicate that the EQ-5D-3L showed moderate to large responsiveness with all tariffs (Table 6.11). With the exception of Finland, the EQ-5D-3L scores obtained with EVS were more responsive than EQ-5D-3L scores obtained with the UKVS and CVS. It has been suggested that SRM values of 0.20, 0.50 to 0.80 and 0.80 and above indicate small, moderate and large responsiveness respectively (Husted et al. 2000).

Table 6.11 Mean (SD) EQ-5D-3L scores, change over time and Standardized response mean

Country		Baseline	Week 1	Week 2	Week 3	Week 4	Difference Week 4-Baseline	Standardized response mean (SRM)
Wales	**EVS**	0.697 (0.235)	0.719 (0.233)	0.777 (0.241)	0.809 (0.238)	0.827 (0.234)	0.130**	0.62
	CVS	0.675 (0.296)	0.711 (0.279)	0.770 (0.284)	0.803 (0.282)	0.821 (0.282)	0.146**	0.54
	EVS- CVS	0.022**	0.008**	0.007**	0.006**	0.006**	-0.015	
England	**EVS**	0.739 (0.188)	0.794 (0.206)	0.866 (0.184)	0.876 (0.194)	0.896 (0.185)	0.157**	0.68
	CVS	0.737 (0.221)	0.792 (0.239)	0.864 (0.215)	0.873 (0.229)	0.893 (0.216)	0.156**	0.57
	EVS- CVS	0.002 **	0.002**	0.002**	0.003**	0.003**	-0.004**	
Netherlands	**EVS**	0.725 (0.210)	0.765 (0.200)	0.854 (0.163)	0.892 (0.157)	0.902 (0.154)	0.177**	0.80
	CVS	0.768 (0.209)	0.807 (0.193)	0.883 (0.154)	0.915 (0.140)	0.922 (0.139)	0.154**	0.74
	UKVS	0.719 (0.257)	0.764 (0.237)	0.856 (0.184)	0.897 (0.167)	0.905 (0.166)	0.186**	0.72
	EVS- CVS	-0.043**	-0.042**	-0.038**	-0.023**	-0.020**	-0.023**	
	EVS-UKVS	0.006**	0.001**	-0.002**	-0.005**	-0.003**	-0.010*	
	UKVS- CVS	-0.049**	-0.043**	-0.027**	-0.018**	-0.017**	0.032**	
Spain	**EVS**	0.730 (0.184)	0.812 (0.146)	0.905 (0.136)	0.930 (0.127)	0.932 (0.126)	0.202**	1.01
	CVS	0.730 (0.179)	0.799 (0.145)	0.901 (0.138)	0.927 (0.128)	0.930 (0.128)	0.200**	0.74
	UKVS	0.721 (0.236)	0.823 (0.163)	0.912 (0.140)	0.935 (0.133)	0.938 (0.132)	0.217**	0.88
	EVS- CVS	0.000	0.013**	0.004*	0.003	0.002	0.002**	
	EVS-UKVS	0.009**	-0.011**	-0.007**	-0.005**	-0.006**	-0.014**	
	UKVS- CVS	-0.009**	0.024**	0.011**	0.008**	0.008**	0.016**	
Germany	**EVS**	0.717 (0.179)	0.795 (0.205)	0.881 (0.166)	0.917 (0.138)	0.928 (0.133)	0.211**	1.0
	CVS	0.825 (0.195)	0.870 (0.204)	0.934 (0.135)	0.958 (0.983)	0.962 (0.991)	0.110**	0.65

Country		Baseline	Week 1	Week 2	Week 3	Week 4	Difference Week 4-Baseline	Standardized response mean (SRM)
	UKVS	0.713 (0.226)	0.788 (0.248)	0.882 (0.182)	0.920 (0.144)	0.930 (0.143)	0.217**	0.86
	EVS-CVS	-0.108**	-0.075**	-0.053**	-0.041**	-0.034**	0.073**	
	EVS-UKVS	0.004**	0.007**	-0.001**	-0.003**	-0.002**	-0.006**	
	UKVS- CVS	-0.112**	-0.082**	-0.052**	-0.038**	-0.032**	0.080**	
Belgium	**EVS**	0.727 (0.222)	0.772 (0.157)	0.873 (0.152)	0.891 (0.142)	0.892 (0.152)	0.165**	0.73
	CVS	0.705 (0.235)	0.748 (0.162)	0.859 (0.168)	0.878 (0.156)	0.880 (0.155)	0.175**	0.73
	UKVS	0.715 (0.284)	0.782 (0.184)	0.881 (0.161)	0.900 (0.144)	0.900 (0.156	0.185**	0.66
	EVS- CVS	0.022**	0.024**	0.014**	0.013**	0.012**	-0.010**	
	EVS-UKVS	0.012**	-0.010**	-0.008**	-0.009**	-0.071**	-0.021**	
	UKVS- CVS	0.010**	0.034**	0.022**	0.022**	0.020**	0.011	
Finland	**EVS**	0.669 (0.230)	0.798 (0.174)	0.878 (0.165)	0.924 (0.139)	0.933 (0.136)	0.264**	1.04
	CVS	0.665 (0.193)	0.766 (0.182)	0.858 (0.185)	0.911 (0.158)	0.922 (0.155)	0.257**	1.20
	UKVS	0.652 (0.281)	0.801 (0.187)	0.881 (0.171)	0.927 (0.135)	0.937 (0.132)	0.285**	0.95
	EVS- CVS	0.004**	0.032**	0.020**	0.013**	0.011**	0.007	
	EVS-UKVS	0.017	-0.003**	-0.003**	-0.003**	-0.004**	-0.020**	
	UKVS-CVS	-0.013	0.035**	0.023**	0.016**	0.015**	0.028	

**** Significantly different at 5% level***

***** Significantly different at 1% level Wilcoxon signed rank test***

QALY scores and cost-effectiveness

There were gains in QALYs over the four week period in all countries and with all value sets. From the analysis, the difference between the QALYs gained generated from the different tariffs was very small in most cases (Table 6.12). The difference obtained between the EVS and the UKVS was very small in all countries and was zero in Netherlands, Spain and Germany. The difference between the CVS and the EVS was larger than the difference between the EVS and UKVS in most countries. The size of the difference between the UKVS and the CVS in all countries was similar to the difference that was recorded between the EVS and the CVS. Overall the greatest difference in QALYs between the EVS and CVS and UKVS and CVS was in Germany.

When the difference in QALYs gained between intervention A and intervention B was considered, there was no difference in the QALYs gained obtained from different tariffs in England, Spain and Germany. However, there was a difference in QALY gained in Wales, Netherlands, Belgium and Finland. However, these differences were very small (Table 6.13).

Based on the thresholds in Netherlands and UK (England and Wales), and Spain, the results of this study indicate that based on cost-effectiveness alone, a similar decision would be taken regardless of which tariff was used in the study. In Belgium and Finland however, the results suggest that different results might be reached if different tariffs were used to value EQ-5D-3L. Similar results were achieved when a threshold of £30,000 per QALY was assumed (Table 6.14).

It should be noted however, that this depends on the difference in the cost between the interventions. For example, in Belgium, if the difference in cost between the

interventions is greater than £50, a similar decision would be reached regardless of the tariff that is used. When the difference between the interventions was £50 the intervention was cost-effective when either the EVS or CVS was used but was not cost-effective when the UKVS was used. In Finland if the difference in cost between the interventions is greater than £100 a similar decision would be reached regardless of the tariff that is used. If the difference between the interventions is £100, the intervention would be cost-effective with the CVS but not cost-effective when the EVS or the UKVS are used. If the difference between the interventions is £50 the intervention would appear cost-effective regardless of the tariff that is used.

Results obtained from sensitivity analysis

The population for this sub study included 370 patients (222 from Sweden and 148 from Norway) of whom 32.4% were male. QALYs generated with the EVS and UKVS were 0.0012 and 0.0018 respectively and resulted in ICERs of £8,342 per QALY and £5,561 per QALY for the EVS and UKVS. At a threshold of £30,000 per QALY, the intervention (CRP) was cost-effective irrespective of the tariff that was used.

Table 6.12 QALY gains over four weeks

Country		QALYs gain over four weeks
Wales	EVS	0.059
	CVS	0.058
	EVS- CVS	0.001
England	EVS	0.064
	CVS	0.064
	EVS- CVS	0.000
Netherlands	EVS	0.064
	CVS	0.066
	UKVS	0.064
	EVS- CVS	-0.002**
	EVS-UKVS	0.000
	UKVS- CVS	-0.002
Spain	EVS	0.067
	CVS	0.066
	UKVS	0.067
	EVS- CVS	0.001**
	EVS-UKVS	0.000
	UKVS- CVS	0.001
Germany	EVS	0.066
	CVS	0.070
	UKVS	0.066
	EVS-CVS	-0.004
	EVS-UKVS	0.000
	UKVS- CVS	-0.004
Belgium	EVS	0.064
	CVS	0.063

Country		QALYs gain over four weeks
	UKVS	0.065
	EVS- CVS	0.001
	EVS-UKVS	-0.001*
	UKVS- CVS	0.002
Finland	EVS	0.065
	CVS	0.064
	UKVS	0.065
	EVS- CVS	0.001
	EVS-UKVS	0.001
	UKVS-CVS	0.000

**** Significantly different at 5% level ** Significantly different at 1% level paired t-test***

Table 6.13 Health gains between interventions

		Intervention A	*Intervention B*	**Absolute Difference**
Wales	EVS	0.058	0.061	0.003
	CVS	0.057	0.061	0.004
England	EVS	0.063	0.066	0.003
	CVS	0.063	0.066	0.003
Netherlands	EVS	0.064	0.064	0
	CVS	0.066	0.067	0.001
	UKVS	0.063	0.065	0.002
Spain	EVS	0.068	0.066	0.002
	CVS	0.068	0.066	0.002
	UKVS	0.069	0.067	0.002
Germany	EVS	0.065	0.066	0.001
	CVS	0.070	0.071	0.001
	UKVS	0.065	0.066	0.001
Belgium	EVS	0.063	0.065	0.002
	CVS	0.061	0.064	0.003
	UKVS	0.064	0.065	0.001
Finland	EVS	0.064	0.067	0.003
	CVS	0.062	0.066	0.004
	UKVS	0.064	0.067	0.003

Table 6.14 Incremental cost-effectiveness ratios with different tariffs

Wales	**EVS**				**CVS**							
Difference in cost (£)	50	100	300	500	50	100	300	500				
Difference in QALYs	0.003	0.003	0.003	0.003	0.004	0.004	0.004	0.004				
ICER (£) [a]	16,666.7	33,333.3	100,000	166,666.7	12,500	25,000	75,000	125,000				
Netherlands[b]	**EVS**				**CVS**				**UKVS**			
Difference in cost (£)					50	100	300	500	50	100	300	500
Difference in QALYs					0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.002
ICER (£)[a]					50,000	100,000	300,000	500,000	25,000	50,000	150,000	250,000
Belgium	**EVS**				**CVS**				**UKVS**			
Difference in cost (£)	50	100	300	500	50	100	300	500	50	100	300	500
Difference in QALYs	0.002	0.002	0.002	0.002	0.003	0.003	0.003	0.003	0.001	0.001	0.001	0.001
ICER (£)[a]	25,000	50,000	150,000	250,000	16,666.7	33,333.3	100,000	166,666.7	50,000	100,000	300,000	500,000
Finland	**EVS**				**CVS**				**UKVS**			
Difference in cost (£)	50	100	300	500	50	100	300	500	50	100	300	500
Difference in QALYs	0.003	0.003	0.003	0.003	0.004	0.004	0.004	0.004	0.003	0.003	0.003	0.003
ICER (£)[a]	16,666.7	33,333.3	100,000	166,666.7	12,500	25,000	75,000	125,000	16,666.7	33,333.3	100,000	166,666.7

[a] Cost per QALY gained [b] Excluded because the difference between interventions A and B with the EVS was zero.

Table 6.15 Sensitivity analysis

	EVS			UKVS		
	Coefficient	**CI**	**P-Value**	**Coefficient**	**CI**	**P-Value**
Cost difference	£10.01	-1.86, 0.004	0.09	£10.01	-1.86, 0.004	0.09
QALY difference	0.0012	-0.001, 0.004	0.35	0.0018	-0.001, 0.005	0.23
ICER	£8,342 per QALY gained			£5,561 per QALY gained		

6.4 Discussion and conclusion

6.4.1 Summary of chapter

This chapter has considered issues around outcomes when conducting economic evaluations alongside multinational studies/trials. The first aspect of the chapter validated the EQ-5D-3L in patients with acute cough/lower respiratory tract infections and the second section focused on a comparison of different EQ-5D-3L tariffs in a multinational study context.

6.4.2 Validity of the EQ-5D-3L

This study found that patients reported more problems with usual activities and pain/discomfort than with other dimensions of the EQ-5D-3L, a result which can be anticipated as most LRTIs are not expected to affect mobility and self-care but may significantly affect individual's daily activities and cause considerable pain and discomfort. Many individual items on the EQ-5D-3L were related to individual items on the symptoms diary. However, the relationship between headache on the symptoms diary and pain/discomfort on the EQ-5D-3L is a cause for concern. One might expect this relationship to be significant, but the non-significant relationship that was found in this study could be due to the sample size as the validity study was limited to UK participants only. Overall, results obtained were in line with the a priori expectations and the EQ-5D-3L showed evidence of being able to capture severity in this patient population and also discriminate between patients who had chronic diseases.

The trend in EQ-5D-3L shows an improvement in the health of patients over the four week period and the EQ-5D-3L was found to be moderately responsive indicating that the EQ-5D-3L is quite sensitive to change in health state over time in this patient population. This is an important result because it shows that the EQ-5D-3L is useful for acute conditions with

relatively shorter durations, as long as assessments are made in a timely manner. In addition it would be an adequate measure for economic evaluation studies and for policy makers seeking to establish the effect of an intervention in the context of acute cough/LRTI. From this validation study, it can be inferred that the EQ-5D-3L appears to have reasonable construct validity in patients with acute cough/LRTI. As a result, it is an appropriate measure of quality of life in this patient group.

6.4.3 Comparison of various EQ-5D-3L tariffs in the context of multinational trials

EQ-5D-3L scores were left skewed irrespective of the value set that was used. At baseline, a number of individuals already had a score of 1, despite having symptoms of acute cough/LRTI, and suggesting the possibility of ceiling effects, a problem which has been associated with the EQ-5D-3L (Brazier et al. 2004; Goldsmith et al. 2009). The EVS yielded higher EQ-5D-3L scores than the CVS in some countries (e.g. Wales, England) and lower scores than the CVS in other countries (e.g. Netherlands, Germany). The higher EQ-5D-3L scores that were obtained with the CVS in Germany could have been as a result of the exclusion of the non-significant factors in the model that was used in the development of Germany's CVS (Greiner et al. 2005) i.e. the German value set has no disutility for levels 2 and 3 of usual activities and level 2 of anxiety/depression and as a result of this, nothing is subtracted from the model if an individual reports a value of 2 or 3 for usual activities or a value of 2 for anxiety/depression.

There was considerable association between all value sets as shown by the correlation coefficients. The Bland-Altman plots also suggest that there is considerable agreement between EQ-5D-3L scores based on the different value sets although at an individual level the limits of agreement would be wide. However, there was greater agreement between the EVS and UKVS in all countries. This is not surprising, as the EVS includes a much higher

number of respondents (approximately 47%) and valuations from the UK than from any other country in the EVS valuation study.

The difference in the EQ-5D-3L scores for the different value sets between baseline and week four were statistically significant in most cases and this seemed to imply that using different value sets may result in differences in cost-effectiveness results. However, when QALYs were generated with the various tariffs, the resulting differences in QALYs obtained from the CVS, EVS and UKVS were very small. From this study, it was observed that in most countries (with the exception of Belgium and Finland) the choice of value set did not change the decision that would be taken based on the results of the cost-utility analysis. It should however be noted that difference in costs between intervention and the cost-effectiveness thresholds used also play a part in determining whether an intervention is cost-effective or not.

6.4.4 Comparison with other studies

The finding that the EQ-5D-3L scores obtained with the CVS was higher than the EVS in Netherlands and Germany was similar to that found in another study (Bernert et al. 2009). However, this study only considered three countries and did not assess the impact on cost-utility analysis. The work presented in this chapter considered seven countries and also considered the implication of using different value sets for cost-utility analysis in an illustrative example. Some studies have concluded that the choice of value set for cost-utility analysis does not matter (Sakthong et al. 2008) whilst others have suggested that the use of different value sets matters and have suggested that different weights may result in differences in QALYs which would yield different cost-effectiveness estimates (Huang et al. 2007; Johnson et al. 2005).

6.4.5 Strengths and limitations of the study

This study is associated with the following strengths: First, due to the observational nature of this study, it includes all patients presenting with cough and not just a sub-group of patients and covers the whole range of severities, which is important in validation studies. Second, the study also considered the impact of the tariffs in cases where a particular country does not have their own tariff. Some limitations associated with this study include the following: First, the validation study included only the UK participants (England and Wales) and the comparison study limited to patients with acute cough/LRTI therefore the generalisability of the results to other countries and patient populations might be limited. Second, patients were asked to complete the diary up until the point that they felt well and as a result of this, it is not always clear whether the patient left the study when they actually felt well or for some other reason. The analysis comparing different value sets was therefore conducted excluding missing data. Third, the study was conducted with the 3-level version of the EQ-5D questionnaire and not the latest version of the EQ-5D which is the 5-level version of the questionnaire (Herdman et al. 2011). However, it should be noted that the 5-level of the questionnaire was not collected as part of the GRACE study.

6.4.6 Implications for policy and practice

Studies have previously concluded that cultural and other factors affect valuation of health states (Guillemin et al. 1993; knies et al. 2009). In addition, guidance produced by the EuroQol group suggests that value sets that are relevant to individual countries should be used (Devlin and Parkin, 2007). Therefore, for studies in individual countries, it is preferable that country specific value sets should be used if they exist. However, it has been recommend that there is a need for much larger studies which are more representative of the individual country to be used in the development of country specific value sets (Bernert et al. 2009). If a

tariff does not exist for that particular country, then it is possible to use either the UKVS or EVS in such circumstances.

In the case of multinational studies and international comparisons, it is difficult to determine whether the EVS or UKVS is a better measure for cost-utility analysis. Previous research has concluded that either the EVS or UKVS is appropriate for multinational studies (Bernert et al. 2009). The results presented here show that the difference in QALYs between the CVS and EVS as well as the difference between the CVS and UKVS was larger than the difference between the EVS and UKVS. This seems to suggest that either the EVS or UKVS is suitable. Considering the fact that the EVS was developed from studies drawn from different countries, it would seem as if it would be appropriate to use the EVS for multinational studies in Europe. However, researchers (Greiner et al. 2003) have observed that some of the studies that were used in the development of the EVS were not representative of the individual countries and also the countries that were selected only represent a small portion of Europe (Western and Northern Europe). None of these studies inform the question of what value set is most appropriate for Eastern and Southern European countries. In addition, the EVS is based on VAS and there are those who argue that VAS should not be used for cost-utility analysis and resource allocation decisions because it lacks a theoretical foundation and violates the axioms of utility theory (King et al. 2005). However, there is research currently exploring the predictive value of the EQ-VAS for EQ-5D utilities (Feng et al. 2014).

6.4.7 Conclusion

This study has considered the use of different EQ-5D-3L tariffs in the context of multinational trials and although differences in cost-effectiveness resulting from use of these tariffs were generally small, it is recommended here that analysts ensure that the potential for different findings is checked in all multinational studies, by the use of all appropriate tariffs (particularly the UKVS and EVS) within sensitivity analysis to ensure that results are

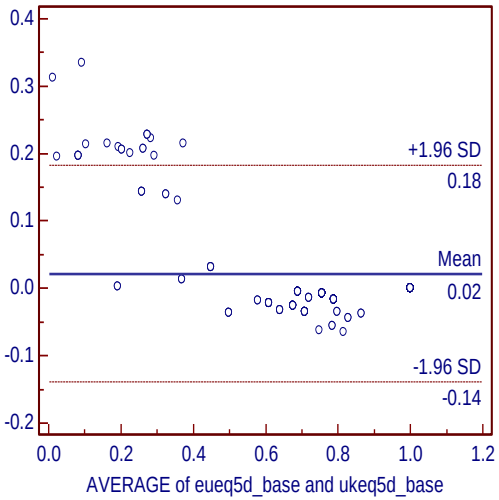
generalisable. It is also recommended that the EuroQol group should produce additional guidance on the use of EQ-5D for valuing health states in the context of multinational trials.

6.5 Summary

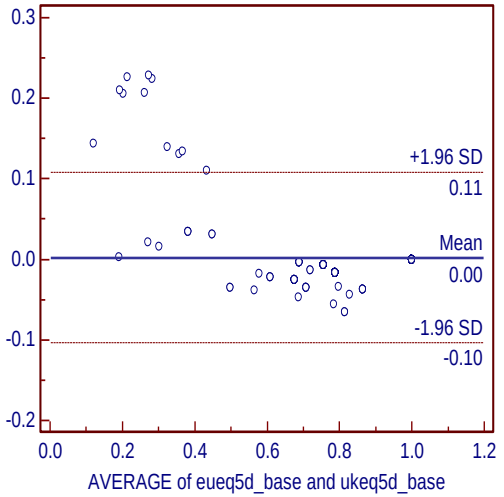
This chapter was aimed at assessing outcomes in multinational trials with a focus on the EQ-5D-3L. The 3-level version of the EQ-5D questionnaire was initially validated in patients with acute cough/LRTI and the impact of using different tariffs to value EQ-5D in multinational studies was assessed. The results showed the EQ-5D-3L to be a valid measure for use in patients with acute cough/LRTI and that in most cases, the choice of tariff did not make a difference to the conclusions from cost-utility analysis. This thesis has now covered issues relating to costs and outcomes in multinational trials. The next chapter now focuses attention on issues relating to conducting economic evaluation as a whole with a focus on comparing pooled and split approaches to economic evaluation alongside multinational trials.

Figure 6.1 Bland-Altman plots CVS, EVS and UKVS

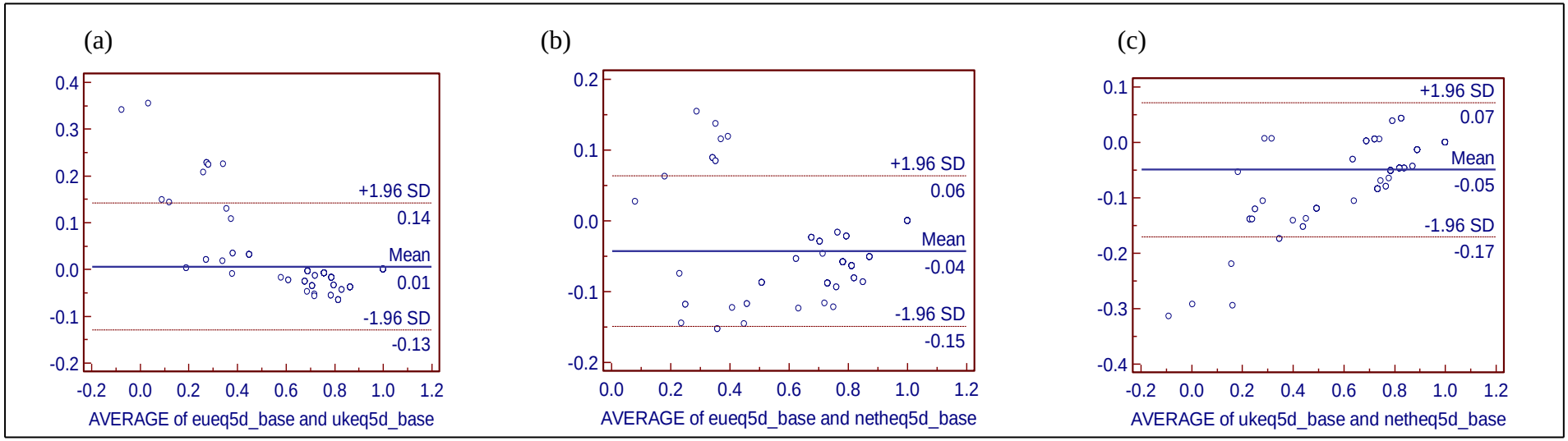
(A) PLOTS FROM CVS AND EVS (WALES)



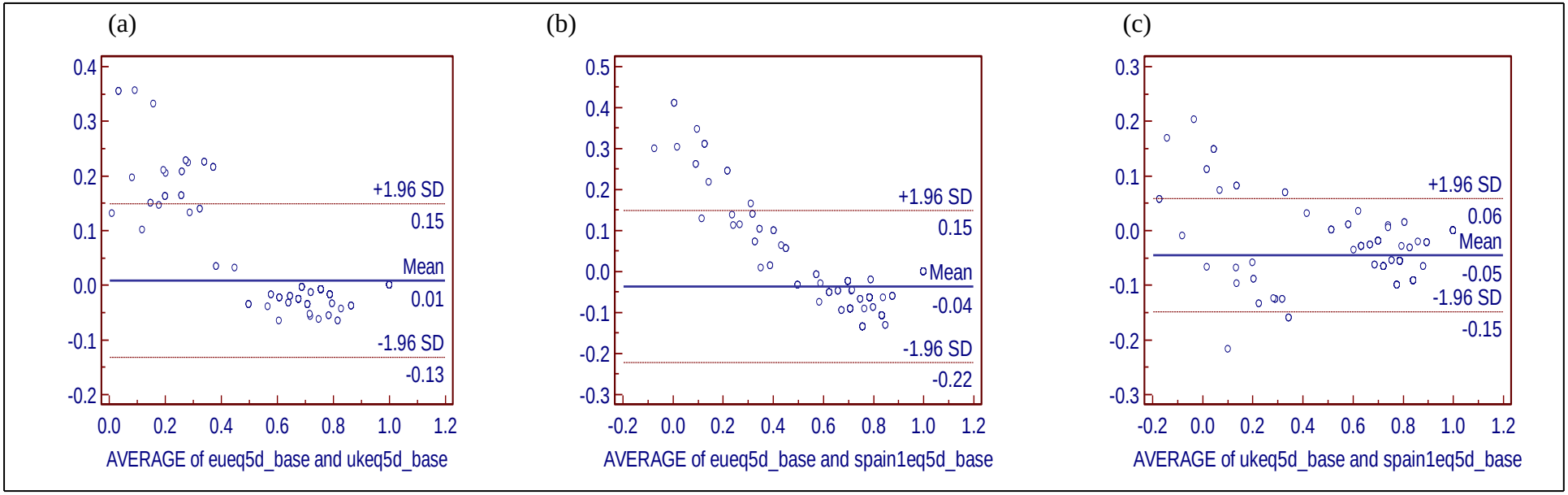
(B) PLOTS FROM CVS AND EVS (ENGLAND)



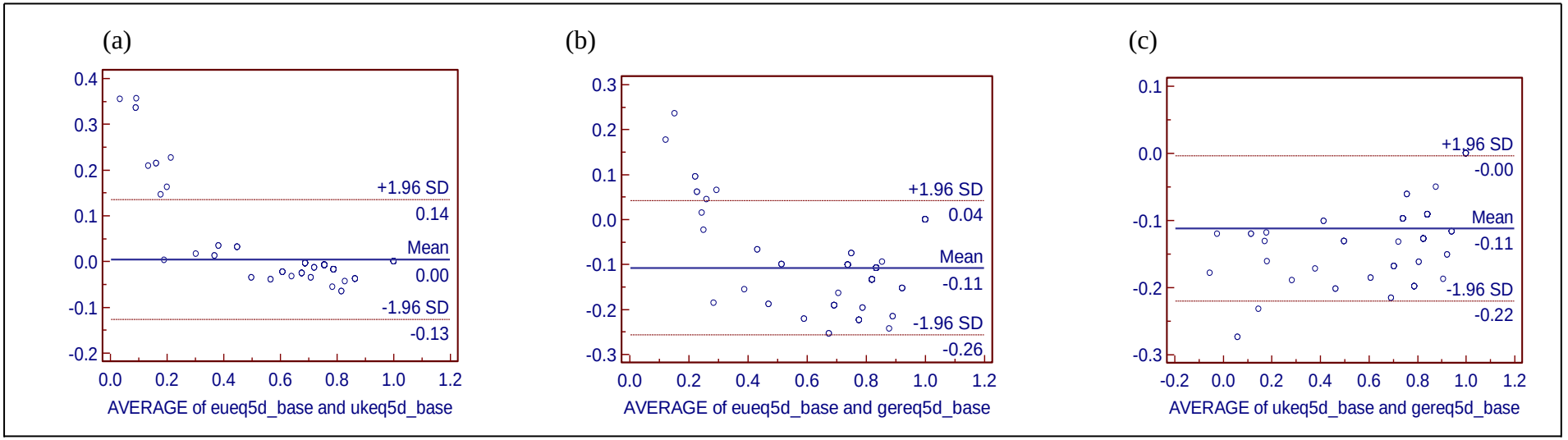
BLAND-ALTMAN PLOTS OF EQ-5D SCORES FROM CVS EVS and UKVS (NETHERLANDS)



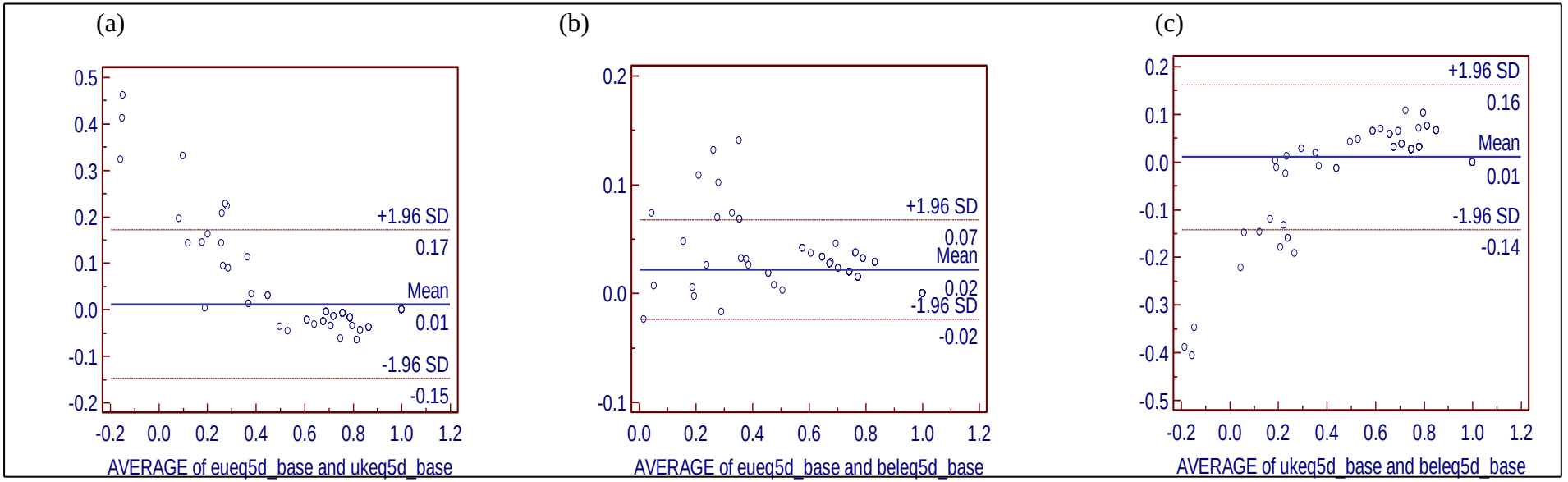
BLAND-ALTMAN PLOTS OF EQ-5D SCORES FROM CVS EVS and UKVS (SPAIN)



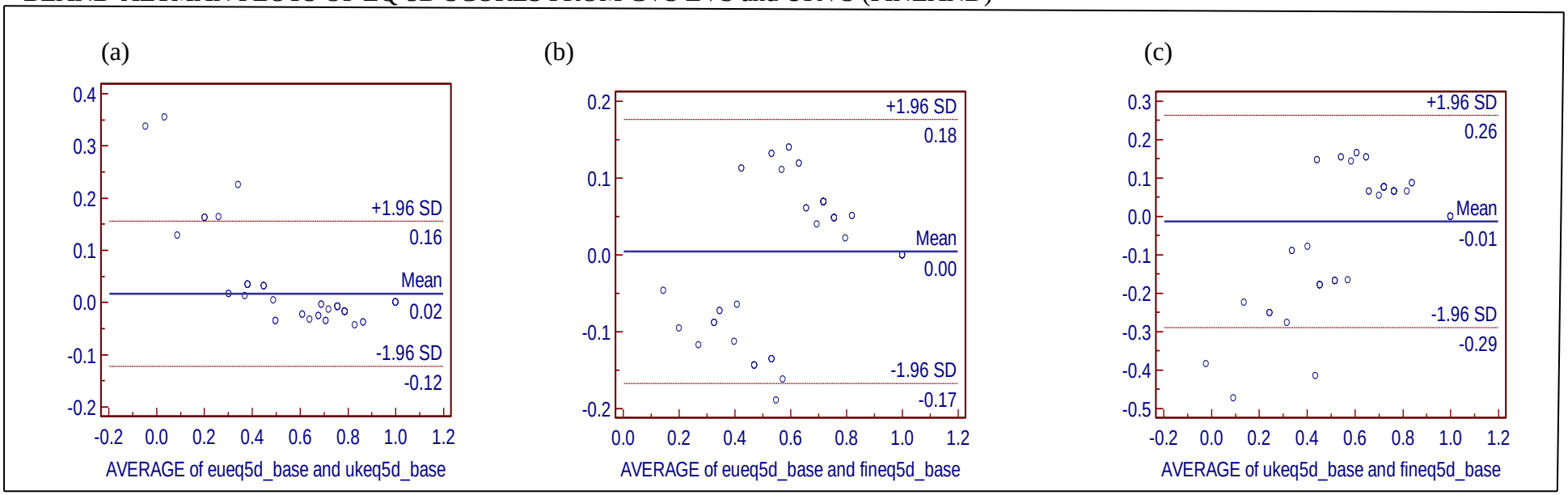
BLAND-ALTMAN PLOTS OF EQ-5D SCORES FROM CVS EVS and UKVS (GERMANY)



BLAND-ALTMAN PLOTS OF EQ-5D SCORES FROM CVS EVS and UKVS (BELGIUM)



BLAND-ALTMAN PLOTS OF EQ-5D SCORES FROM CVS EVS and UKVS (FINLAND)



CHAPTER 7 ANALYTICAL APPROACHES TO ECONOMIC EVALUATION ALONGSIDE MULTINATIONAL TRIALS

7.1 Background

Issues related to costing and outcomes in multinational studies were explored and discussed in the preceding chapters. The work presented in this chapter explores the alternative analytical approaches (pooled and split) available for the conduct of economic evaluations alongside multinational trials using Case study 1 and Case study 2 from the GRACE project. A brief background is presented initially, followed by the results of the various analytical approaches from both case studies.

In an attempt to reach a consensus on how economic evaluations should be conducted alongside multinational trials, researchers reviewed existing literature and suggested a number of analytical approaches based on the sources of resource use, cost and effectiveness data (Reed et al. 2005). These methods have been described in detail in Chapter 2 and they range from the fully pooled multicountry costing approach to the fully split one country costing approach. In Chapter 3, a review of the existing literature on economic evaluations alongside multinational trials was conducted. The results showed that the most common analytical method, adopted by 60% of included studies, was the fully pooled one country costing approach. The approach considers estimates of resource use and effectiveness from all participating countries whilst costs from just one of the participating countries are applied

to the data. Even though there are benefits associated with each of the approaches (see Chapter 2), it is still unclear what impact each approach would have on the cost-effectiveness of interventions under consideration. There is therefore the need for additional research to assess the various approaches in terms of their strengths and weaknesses and whether they lead to different conclusions.

The aims of this chapter are as follows:

(i) To compare analytical approaches to economic evaluation alongside multinational trials using the case studies (Case studies 1 and 2) from the GRACE study. The focus here will be on the fully pooled and the fully split analytical approaches (Table 7.1). A partially split analysis will not be considered in this chapter because the review undertaken in Chapter 3 did not identify many studies that used this method. However, a summary of results from this approach have been presented in Appendix 11.

Table 7.1 Analytic approaches to economic evaluation alongside multinational trials

1	Fully pooled with one country costing
2	Fully pooled with multicountry costing
3	Fully split with one country costing
4	Fully split with multicountry costing

7.2 Methods: Case study 1 (GRACE WP10A)

7.2.1 Study design

The economic evaluation took the form of a cost-utility analysis and was conducted alongside a randomised double blinded trial in which patients received either amoxicillin or placebo/control (Little et al. 2013a). The perspective adopted was health care including costs to the health service, and health care costs to the patient. A total of 2,060 eligible and consenting patients were recruited in 16 primary care networks across 12 countries in Europe: Belgium, France, Germany, Italy, Netherlands, Poland, Slovakia, Slovenia, Spain, Sweden and UK (England and Wales) (see Chapter 4). The study was approved by ethics committees in all participating countries and all patients provided written consent before participating in the study. Full details of the clinical trial have been presented in Chapter 4 and the clinical results have been published elsewhere (Little et al. 2013a). Some of the results from Case study 1 presented in this Chapter have been published in the following peer reviewed article: *Oppong et al. (2016) Cost-effectiveness of Amoxicillin for LRTI in Primary Care: An Economic Evaluation Accounting for the Cost of Antimicrobial Resistance British Journal of General Practice* ***DOI:*** *10.3399/bjgp16X686533.*

7.2.2 Resource use data

The main sources of resource use information were the case report form (CRF) completed by primary care physicians and a diary completed by patients. Resource use data were collected concerning:

- **Health professionals** - including information on the number of visits to the nurse, doctor and other medical professionals. This was obtained from the patient diary.

- **Medication** - including information on the type and volume of medication that primary care physicians prescribed to patients as well as information on over-the-counter medication purchased during the course of the study. This information was obtained from both the CRF and patient diary.

- **Referrals to specialists and procedures** - including information on the numbers and types of referrals (obtained from CRF).

7.2.3 Unit costs

Country-specific unit costs associated with resource use items were obtained mainly from national and international publications. In cases where country specific costs were not available, those obtained from the costing study carried out in Chapter 5 were used and inflated using the consumer price index for each country (World Bank). Where unit costs remained unavailable, a market basket approach (Schulman et al. 1998) was used to estimate the relationship between the UK and the country of interest and then estimate a cost (Appendix 8). The UK was used in this estimation because all unit costs were available for this setting.

Medication (prescribed and over the counter) were classified into 13 different groups (mucolytic, other remedies, antitussives, other medication, anti-inflammatories, bronchodilators, expectorants, antihistamine, tetracycline, amoxicillin, macrolide, other antibiotics and unknown). As it was not feasible to obtain unit costs for each individual drug for each country, a cost was generated for each of the 13 groups by estimating a mean price from a list of drugs within that group.

Table 7.2 gives a summary of the various sources of unit costs used to value resource use.

Table 7.2 Source of valuation data

	Belgium	France	Germany	Italy	Netherlands	Poland	Slovakia	Slovenia	Spain	Sweden	UK
GP Visits	1	2	1	1	1	1	1	2	1	1	3
Nurse Visits	1	2	1	1	1	1	1	2	1	1	3
Out of hours GP	12	12	12	12	12	12	12	12	12	12	3
Walk in centre	1	1	1	1	1	1	1	1	1	1	1
Hospital Admissions	1	2	1	1	1	1	1	2	1	1	11
Investigations	12	12	12	12	12	12	12	12	12	12	5
Medication	10	12	9	1,12	6	1,12	12	12	6,1	11	7

1= Previous study, 2= Derivation from group, 3= Netten and Curtis (www.pssru.ac.uk), 4= www.vademecum.es, 5= British National Formulary (www.bnf.org), 6= Dutch healthcare insurance board (www.medicijnkosten.nl), 7= Rote Liste (The red book) www.rote-liste.de, 8= www.bcfi.be, 9= Swedish Pharmaceutical Benefits Board (www.lfn.se), 10= www.http://riziv.fgov.be, 11= NHS Reference costs 12= Market basket approach, 13= NHF www.nfz.gov.pl

All costs were initially converted to Euros using PPPs using the Euro area 17 countries as a reference (OECD, 2012). In addition, costs were also converted to Pounds sterling using average exchange rates. All costs are presented in 2012 prices.

7.2.4 Health outcomes

For the economic analysis, health outcomes were measured using the 3-level version on the EQ-5D (EQ-5D-3L) (Rabin and de Charro 2001). Patients were asked to complete the EQ-5D questionnaire over the entire four week period (at baseline, and at the end of weeks 1, 2, 3 and 4), or until they felt better. EQ-5D index scores were generated using the European Harmonised Tariff (Greiner et al. 2003).

7.2.5 Data analysis

Data analysis was carried out on an intention to treat basis and took an incremental approach. Multiple imputation (Rubin, 1987) was used to impute missing EQ-5D and cost data over the four week period. An imputation model was fitted using Multiple Imputation by chained equations (MICE) in STATA version 12.1 (STATA, 2012) and included 25 imputed datasets. Predictive mean matching was used to model cost and EQ-5D (STATA, 2011). Predictive mean matching was chosen so that the imputed values remained on the same scale as their original outcome and because this method is particularly suited to modelling skewed data (White et al. 2011). For each patient, Quality Adjusted Life Years (QALYs) were estimated using the area under the curve approach and total cost was calculated by multiplying the resources used by the relevant country's unit cost for that item of resource use. Mean differences in costs and QALYs between trial arms were estimated. To avoid biased QALY scores, imbalances in baseline utility between the intervention and control groups were accounted for using a regression approach (Manca et al. 2005). Net benefits were also estimated for each patient using the formula (threshold x QALY- total cost) (Drummond et al. 2015).

As a result of the multinational nature of this study, hierarchical modelling (with explanatory variables stratified into patient and country levels) was used. Costs, QALYs and net benefits were included as dependent variables. Model estimates of the difference in costs, QALYs and net benefits between patients receiving amoxicillin and control were obtained and used to estimate a cost per QALY gained as well as incremental net monetary benefits. To determine the probability of antibiotics being cost-effective, a cost-effectiveness acceptability curve (CEAC) was constructed (Hoch et al. 2006). For the purpose of this study, the NICE recommended threshold of between £20,000 to £30,000 per QALY gained was used to judge the cost-effectiveness of the interventions (Appleby, 2007). All analysis was carried out in STATA 12 (STATA, 2012) and MS Excel. Due to the short length of the study period (4 weeks), discounting was not required.

7.2.6 Analytical approaches explored (Case study 1)

Four analytical approaches were considered to estimate the cost-effectiveness of amoxicillin vs placebo/control.

Approach I: Fully pooled one country costing approach: With this approach, all 2060 participants from all countries were included in the analysis and QALYs generated using EQ-5D values obtained from the European tariff (Greiner et al. 2003) was used as the main outcome measure. However, resource use items were valued using unit costs from the UK only.

Approach II: Fully pooled multicountry costing approach: Similar to the fully pooled one country costing approach described above, this approach considered QALYs and resource use from all participating countries. However, the unit costs derived from each country was used to value their resource use e.g. UK costs were applied to participants from the UK and Spanish costs were applied to Spanish resource use.

Approach III: Fully split one country costing approach: This analysis was limited solely to 329 participants from the UK and UK unit costs were applied to resource use data. For Case study 1, the UK value set was used to value EQ-5D scores when the FSOC approach was applied.

Approach IV: Fully split multicountry costing approach: This approach focused on the 904 participants who were recruited in UK, Germany, France, Netherlands and Belgium. These countries were chosen because they were representative of the western European countries that participated in this study. Similar to the fully pooled multicountry costing approach, country-specific unit costs were used to value resource use in each country.

7.3 Results from Case study 1

7.3.1 Baseline characteristics

Data were obtained from a total of 2060 patients who met the inclusion criteria. Of these, 1037 (50.3%) were randomised to receive amoxicillin and 1023 (49.7%) received a placebo/control. The contribution that countries made to the overall sample size was not equally distributed and was biased towards certain countries with 26.9% and 15.9% of the total sample size coming from Poland and the UK whilst other countries such as France, Germany and Italy only contributed 1.0%, 1.8% and 0.7% to the total sample size. Mean age was similar in both intervention and control groups and a total of 595 (28.8%) patients were aged 60 and above across all groups. With the exception of Germany, the proportion of females was higher than males in all countries (Table 7.3).

Table 7.3 Country contribution to sample size

Country	Intervention	Control	Total number of patients	Number of females	Aged 60 and Over
Belgium	133 (49.3%)	137 (50.7%)	270 (13.1%)	150 (55.6%)	68 (25.2%)
France	9 (42.9%)	12 (57.1%)	21 (1.0%)	14 (66.7%)	8 (38.1%)
Germany	20 (52.6%)	18 (47.4%)	38 (1.8%)	19 (50.0%)	12 (31.6%)
Italy	7 (46.7%)	8 (53.3%)	15 (0.7%)	10 (66.7%)	7 (46.7%)
Netherlands	124 (50.4%)	122 (49.6%)	246 (11.9%)	141 (57.3%)	87 (35.4%)
Poland	280 (50.5%)	274 (49.5%)	554 (26.9%)	337 (60.8%)	121 (21.8%)
Slovakia	67 (51.2%)	64 (48.9%)	131 (6.4%)	74 (56.5%)	18 (13.7%)
Slovenia	35 (49.3%)	36 (50.7%)	71 (3.5%)	43 (60.6%)	22 (30.9%)
Spain	161 (52.1%)	148 (47.9%)	309 (15.0%)	183 (59.2%)	120 (38.8%)
Sweden	36 (47.4%)	40 (52.6%)	76 (3.7%)	48 (63.2%)	31 (40.8%)
UK (England and Wales)	165 (50.2%)	164 (49.9%)	329 (15.9%)	205 (62.3%)	101 (30.7%)
Total	1037 (50.3%)	1023 (49.7%)	2060 (100%)	1224 (59.4%)	595 (28.9%)

7.3.2 Results: fully pooled one country costing (Approach I)

Resource use and costs (FPOC)

Resources used by the two groups (amoxicillin and control) are shown in Table 7.4. Patients in the control group had more visits to their general practitioner (GP) and nurse whilst those receiving amoxicillin had more out-of-hours GP visits and used more prescribed and over the counter medication. With the exception of GP visits and out-of-hours GP visits, the difference in all resource use items between the intervention and control groups was not statistically significant.

Costs associated with health professionals, over the counter medication, interventions and other health care costs were higher in the control arm whilst prescribed drug costs were higher in the amoxicillin arm. Overall, the amoxicillin group was associated with a lower cost per patient, £61.30 compared with £62.25 for the control arm (Table 7.5).

Health Outcomes (FPOC)

The mean EQ-5D score at baseline was higher in the amoxicillin group and increased over the four week period in both the intervention and control groups, showing (as expected) an improvement in health status in both trial arms. The arm was associated with a greater QALY gain over the 4 week period compared to the control. However, the difference was not statically significant (Table 7.6).

Table 7.4 Resource use mean (SD) (All approaches)

		Intervention	Control	Difference (95% CI)[a]
GP visits	**FPOC/ FPMC**	1.25 (0.59)	1.31 (0.71)	-0.06 (-0.11, -0.01)
	FSOC	1.16 (0.47)	1.12 (0.41)	0.04 (-0.06, 0.12)
	FSMC	1.19 (0.53)	1.23 (0.60)	-0.04 (-0.11, 0.03)
Nurse visits	**FPOC/ FPMC**	0.01 (0.15)	0.02 (0.16)	-0.007 (-0.02, 0.007)
	FSOC	0.01 (0.11)	0.02 (0.15)	-0.01 (-0.04, 0.02)
	FSMC	0.01 (0.08)	0.02 (0.12)	-0.009 (-0.02, 0.004)
Specialists visits	**FPOC/ FPMC**	0.02 (0.14)	0.02 (0.23)	-0.006 (-0.02, 0.008)
	FSOC	0	0	0
	FSMC	0.01 (0.12)	0.02 (0.23)	-0.01 (-0.04, 0.009)
Out of hours visits	**FPOC/ FPMC**	0.09 (0.09)	0.02 (0.04)	0.007 (0.0008, 0.014)
	FSOC	0.006 (0.08)	0	0.006 (0, 0.02)
	FSMC	0.007 (0.08)	0.002 (0.05)	0.004 (-0.004, 0.014)
Hospital emergency visits	**FPOC/ FPMC**	0.01 (0.10)	0.01 (0.11)	-0.001 (-0.01, 0.009)
	FSOC	0.006 (0.08)	0	0.006 (0, 0.02)
	FSMC	0.002 (0.05)	0.002 (0.05)	0
Prescribed medication	**FPOC/ FPMC**	1.10 (1.01)	1.09 (0.99)	0.01 (-0.08, 0.09)
	FSOC	0.35 (0.56)	0.24 (0.49)	0.11 (-0.002, 0.22)
	FSMC	0.90 (0.912)	0.85 (0.91)	0.04 (-0.08, 0.16)
Over the counter medication	**FPOC/ FPMC**	0.99 (1.18)	0.93 (1.13)	0.06 (-0.04, 0.16)
	FSOC	0.88 (1.15)	0.93 (1.25)	-0.04 (-0.30, 0.21)
	FSMC	0.92 (1.12)	0.93 (1.15)	-0.01 (-0.16, 0.13)

[a] Bootstrapped CI **FPOC/FPMC (n=2060), FSOC (n=329), FSMC (n=904)**

Table 7.5 Costs (£) mean (SD) (All approaches)

		Intervention	Control	Difference (95% CI)[a]
Staff costs	**FPOC**	47.27 (38.76)	48.56 (41.47)	-1.29 (-4.67, 2.09)
	FPMC	22.35 (35.78)	21.81 (33.61)	0.54 (-2.58, 3.59)
	FSOC	41.12 (28.27)	38.75 (15.39)	2.36 (-2.29, 7.54)
	FSMC	27.78 (34.93)	28.27 (36.97)	-0.49 (-5.39, 3.80)
Prescribed drug costs	**FPOC**	4.71 (9.48)	4.14 (8.21)	0.57 (-0.19, 1.32)
	FPMC	5.49 (10.63)	4.83 (8.98)	0.66 (-0.29, 1.47)
	FSOC	2.52 (6.95)	1.48 (5.88)	1.04 (-0.35, 2.50)
	FSMC	5.43 (10.19)	4.79 (8.96)	0.64 (-0.60, 1.94)
Over the counter drug costs	**FPOC**	1.54 (4.21)	1.66 (4.91)	-0.12 (-0.54, 0.23)
	FPMC	1.85 (4.99)	2.00 (5.73)	-0.14 (-0.58, 0.32)
	FSOC	0.83 (1.19)	0.83 (1.13)	-0.008 (-0.26, 0.24)
	FSMC	1.51 (3.46)	1.56 (3.99)	-0.05 (-0.52, 0.40)
Intervention/other drug costs	**FPOC**	5.64 (10.54)	6.56 (11.84)	-0.92 (-1.85, 0.03)
	FPMC	5.40 (10.12)	6.29 (11.37)	-0.89 (-1.78, 0.08)
	FSOC	7.03 (12.31)	7.83 (12.04)	-0.79 (-3.19, 2.09)
	FSMC	6.31 (10.55)	7.60 (12.42)	-1.29 (-2.73, 0.29)
Other healthcare costs	**FPOC**	0.94 (10.16)	1.32 (16.71)	-0.38 (-1.59, 0.87)
	FPMC	0.99 (10.78)	1.43 (17.89)	-0.44 (-1.88, 0.64)
	FSOC	0.58 (7.42)	0	0.58 (0, 2.22)
	FSMC	0.42 (6.42)	0.47 (9.53)	-0.05 (-1.29, 0.80)
Intervention costs	**FPOC**	1.20 (0)	0	1.2
	FPMC	2.23 (1.99)	0	2.23 (2.12, 2.36)
	FSOC	1.20 (0)	0	1.2
	FSMC	3.17 (2.23)	0	3.17 (2.98, 3.38)
Total cost	**FPOC**	61.30 (44.97)	62.25 (51.40)	-0.94 (-5.30, 3.24)
	FPMC	38.31 (41.04)	36.34 (44.49)	1.97 (-1.78, 5.30)
	FSOC	53.27 (32.39)	48.90 (21.32)	4.37 (-0.95, 11.16)
	FSMC	44.62 (39.09)	42.69 (46.83)	1.93 (-4.11, 7.47)

[a] Bootstrapped CI **FPOC/FPMC (n=2060), FSOC (n=329), FSMC (n=904)**

Cost-effectiveness (FPOC)

The results showed that amoxicillin was associated with a lower cost than the control. The difference in costs and QALYs between the groups was £-0.661 (CI: -4.69, 3.37) and 0.00037 (CI: -0.00029, 0.00097) respectively (Table 7.7). Amoxicillin thus dominated the control group i.e. it was less costly and more effective than placebo. The cost-effectiveness acceptability curve shows that at a willingness to pay threshold of £20,000 per QALY gained, there is an 88% chance that amoxicillin is cost-effective compared to the control (Figure 7.1). Incremental net benefit at £20,000 per QALY was £7.7 (CI: -5.38, 20.80) (Figure 7.2a).

Table 7.6 Health outcomes mean (SD) (All approaches)

		Intervention	Control	Difference (95% CI)[a]
EQ-5D baseline	**FPOC/ FPMC**	0·760 (0·185)	0·752 (0·192)	0·008 (-0·007, 0·024)
	FSOC	0.772 (0.180)	0.791 (0.169)	-0.02 (-0.07, 0.02)
	FSMC	0.779 (0.186)	0.777 (0.188)	0.002 (-0.02, 0.03)
EQ-5D week 1	**FPOC/ FPMC**	0·840 (0·173)	0·824 (0·176)	0·016 (0·002, 0·033)
	FSOC	0.832 (0.189)	0.824 (0.147)	0.002 (-0.04, 0.04)
	FSMC	0.840 (0.169)	0.826 (0.167)	0.014 (-0.007, 0.04)
EQ-5D week 2	**FPOC/ FPMC**	0·908 (0·134)	0·900 (0·134)	0·008 (-0·004, 0·018)
	FSOC	0.888 (0.163)	0.894 (0.121)	-0.01 (-0.05, 0.02)
	FSMC	0.908 (0.132)	0.892 (0.142)	0.017 (-0.0006, 0.03)
EQ-5D week 3	**FPOC/ FPMC**	0·929 (0·122)	0·925 (0·122)	0·004 (-0·006, 0·015)
	FSOC	0.908 (0.155)	0.920 (0.127)	-0.02 (-0.05, 0.02)
	FSMC	0.924 (0.132)	0.917 (0.136)	0.006 (-0.01, 0.02)
EQ-5D week 4	**FPOC/ FPMC**	0·936 (0.107)	0·936 (0·109)	0·0001 (-0·010, 0·008)
	FSOC	0.920 (0.126)	0.931 (0.10)	-0.01 (-0.04, 0.01)
	FSMC	0.930 (0.115)	0.931 (0.111)	-0.0006 (-0.16, 0.01)
QALYs gained over 4 weeks[b]	**FPOC/ FPMC**	0.0678 (0.009)	0.0672 (0.009)	0.0006 (-0.0002, 0.0014)
QALYs gained over 4 weeks[c]	**FPOC/ FPMC**	0.0938	0.0934	0.00037 (0.0002, 0.0009)
QALYs gained over 4 weeks[b]	**FSOC**	0.0669 (0.012)	0.0677 (0.009)	-0.0008 (-0.0031, 0.0015)
QALYs gained over 4 weeks[c]	**FSOC**	0.06728	0.06736	-0.00008 (-0.002, 0.002)
QALYs gained over 4 weeks[b]	**FSMC**	0.0678 (0.010)	0.0670 (0.010)	0.0008 (-0.0004, 0.0020)
QALYs gained over 4 weeks[c]	**FSMC**	0.0678	0.0671	0.0007 (-0.0002, 0.0016)

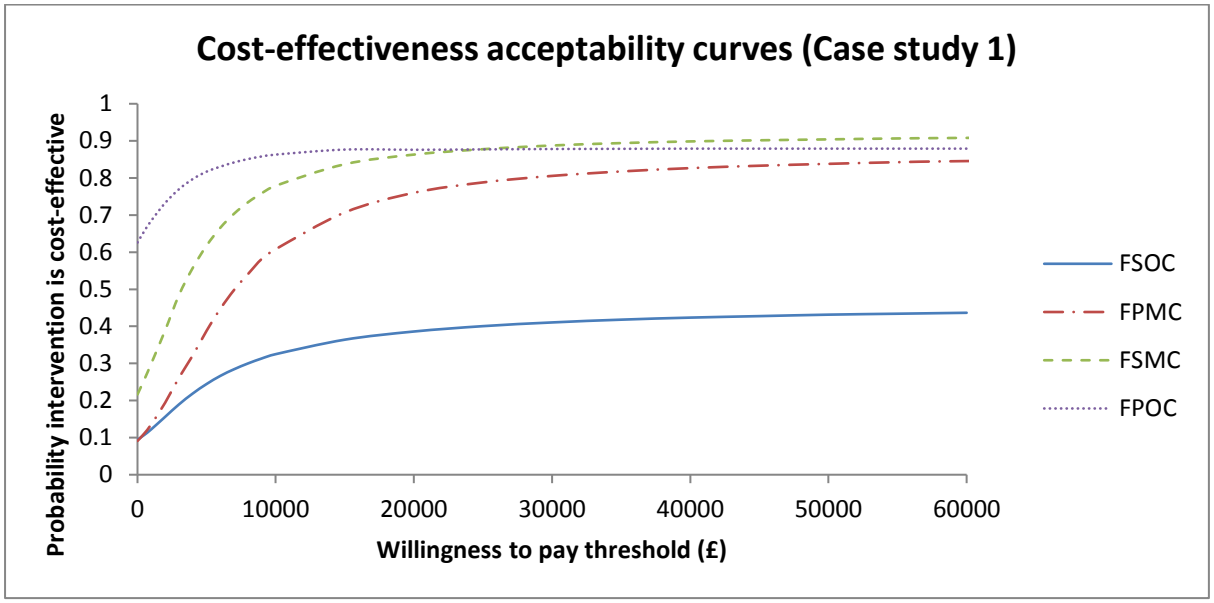
[a] Bootstrapped CI **FPOC/FPMC (n=2060), FSOC (n=329), FSMC (n=904) [b] Undajusted QALYs [c] Adjusted QALYs**

Table 7.7 Cost-effectiveness: Base case analysis (All approaches)

Difference in costs (CI) {CI width}	**FPOC**	£-0.661 (-4.69, 3.37)	{8.06}
	FPMC	£2.42 (-1.14, 5.98)	{7.12}
	FSOC	£3.84 (-1.98, 9.67)	{11.65}
	FSMC	£2.09 (-3.32, 7.49)	{10.81}
Difference in QALYs (CI) {CI width}	**FPOC/FPMC**	0.00037 (-0.0002, 0.0009)	{0.0011}
	FSOC	-0.00008 (-0.002, 0.002)	{0.004}
	FSMC	0.0007 (-0.0002, 0.0016)	{0.0018}
ICER	**FPOC**	Intervention dominant	
	FPMC	£6,540 per QALY gained	
	FSOC	Intervention dominated by placebo	
	FSMC	£2,986 per QALY gained	
Incremental net benefits {CI width}	**FPOC**	£7.7 (-5.38, 20.80)	{26.18}
	FPMC	£4.61 (CI: -7.23, 17.39)	{24.26}
	FSOC		
	FSMC	£11.07 (-8.74, 30.87)	{39.61}

FPOC/FPMC (n=2060), FSOC (n=329), FSMC (n=904)

Figure 7.1 Cost-effectiveness acceptability curves (Case study 1)



7.3.3 Results: fully pooled multicountry costing (Approach II)

Resource use and costs (FPMC)

Mean resources use was equivalent to that obtained with the FPOC approach (Table 7.4). Costs associated with health professionals and prescribed drugs were higher in the amoxicillin arm whilst over the counter costs and other health care costs were higher in the control group. Overall, the amoxicillin group was associated with higher costs per patient £38.32 compared with the control £36.34 (Table 7.5).

Health Outcomes (FPMC)

Health outcomes were similar to those obtained with the FPOC approach (Table 7.6).

Cost-effectiveness (FPMC)

The difference in cost between the amoxicillin and control group was £2.42 whilst the difference in QALYs was 0.00037. The ICER of amoxicillin compared to the control was £6,540 per QALY gained (Table 7.7) and at a cost-effectiveness threshold of £20,000 per QALY amoxicillin is associated with a 76% chance of being cost-effective and associated with positive net benefits (Figure 7.1 and Figure 7.2b).

7.3.4 Results: fully split one country costing (Approach III)

Baseline characteristics (FSOC)

For this approach, the analysis was limited to a total of 329 participants who met the inclusion criteria and were recruited in the UK. Of these, 165 (50.2%) were randomised to receive amoxicillin and 164 (49.85%) received the control. Approximately 62% of all UK participants were female and approximately 30% were aged 60 and above (Table 7.3).

Resource use and costs (FSOC)

The control group was associated with more nurse visits and over the counter medication than those in the amoxicillin group, whilst the amoxicillin group was associated with more GP visits and prescribed medication. The difference in all resource use items between the intervention and control groups was not statistically significant. There was no record of out of hours GP visits, or hospital emergency visits in the control group (Table 7.4). Costs associated with health professionals, prescribed drug costs, other health care costs and the intervention cost were higher in the amoxicillin arm whilst other drug and medical intervention costs were higher in the control group (Table 7.5). Overall, the intervention was associated with higher costs per patient (£53.27) compared with the control (£48.90).

Health Outcomes (FSOC)

The mean EQ-5D score at baseline was higher in the control group than in the amoxicillin group and similar to the pooled approaches, EQ-5D scores increased over the four week period in both the intervention and control groups. With the FSOC, although the difference between treatment arms was not statistically significant, the control arm was associated with a greater gain in QALYs over the 4 week period (Table 7.6).

Cost-effectiveness (FSOC)

The difference in cost and QALYs between trial arms was £3.84 and -0.00008 respectively suggesting that amoxicillin is dominated by the control (more expensive and less effective). At a threshold of £20,000 per QALY gained, there is about a 40% chance that amoxicillin is cost-effective when compared with the control group (Table 7.7 and Figure 7.1). Amoxicillin was also associated with negative incremental benefits at £20,000 per QALY threshold (Figure 7.2c).

7.3.5 Results: fully split multicountry costing (Approach IV)

Baseline characteristics

With this approach, the analysis was limited to a total of 904 participants who met the inclusion criteria and were recruited in the UK, Germany, France, Netherlands and Belgium. Of these, 451 (49.89%) were randomised to receive amoxicillin and 453 (50.1%) received a placebo. Approximately 57% of all participants were female.

Resource use and costs (FSMC)

Patients in the control group had more visits to the general practitioner, nurse and specialist whilst those receiving amoxicillin had more out-of-hours GP visits and prescribed medication (Table 7.4). Costs associated with prescribed drug and the intervention was higher in the amoxicillin arm whilst all other costs were higher in the control arm (Table 7.5). Overall, amoxicillin was associated with higher costs per patient £44.62 compared with placebo £42.69.

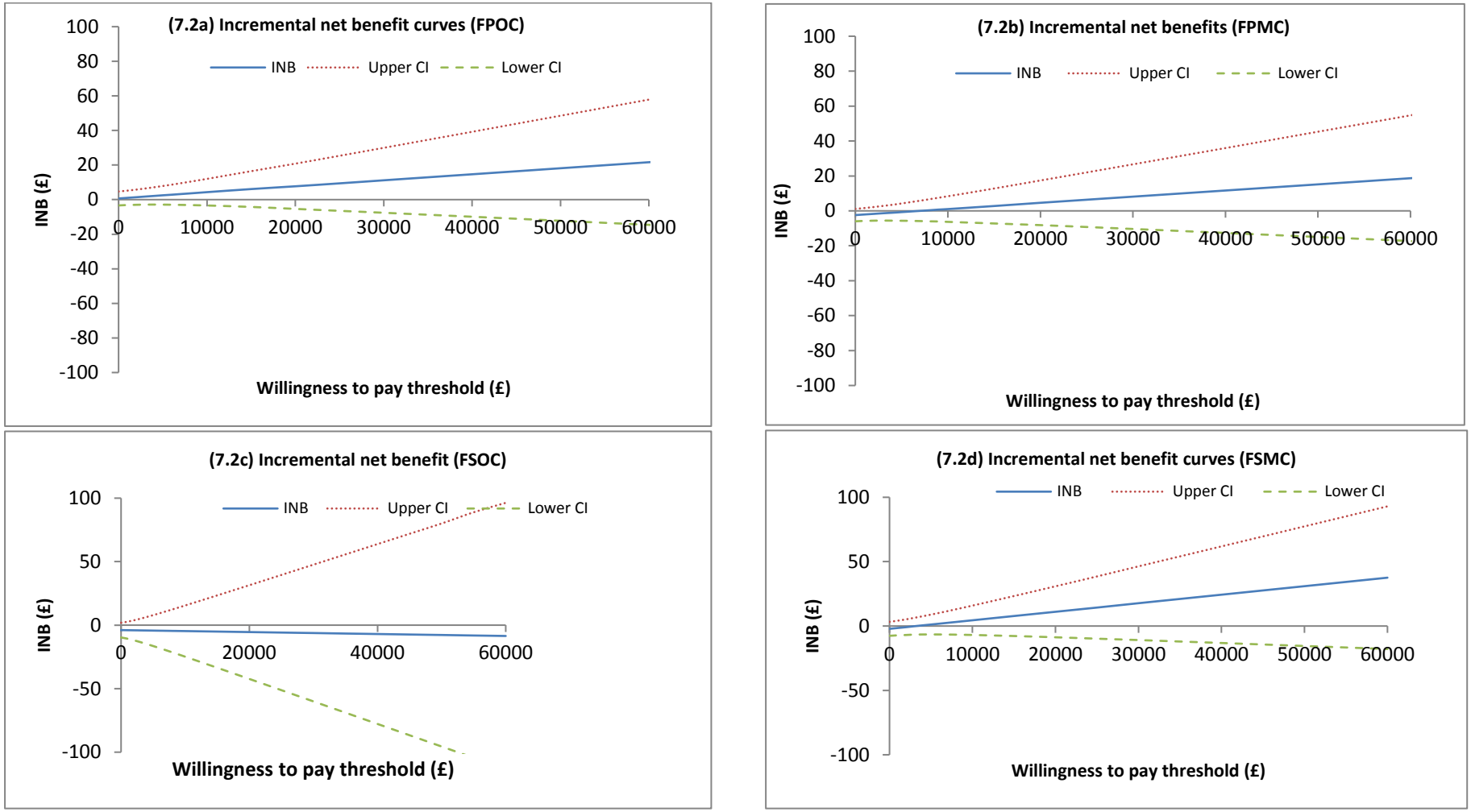
Health Outcomes (FSMC)

The mean EQ-5D score at baseline was higher in the amoxicillin group than in the placebo group and similar to the other approaches, increased over the four week period in both groups. The amoxicillin arm was associated with a greater QALY gain over the 4 week period compared to the placebo arm (Table 7.6).

Cost-effectiveness (FSMC)

Estimates of the difference in cost and QALYs between the amoxicillin and placebo groups was £2.09 and 0.0007 and this resulted in an incremental cost-effectiveness ratio of £2,986 per QALY gained (Table 7.7) and an 86% chance that amoxicillin is cost-effective at a willingness to pay threshold of £20,000 per QALY gained (Figure 7.1).

Figure 7.2 Net benefit curves (Case study 1)



7.4 Methods: Case study 2 (WP10 B)

7.4.1 Study design

The economic evaluation was conducted alongside a multinational, cluster, randomised, factorial controlled trial in which participating practices were randomised to receive one of four interventions (i) usual care, (ii) training in the use of CRP testing, (iii) training in communication skills and (iv) training in *both* CRP and communication skills (Little et al. 2013b). The perspective adopted was health care including costs to the health service, and health care cost to the patient. Consenting participants who presented with respiratory tract infections were recruited from primary care networks across 5 countries in Europe. The study was approved by ethics committees in all countries and all eligible individuals provided written consent before participating. Full details of the clinical trial and the interventions considered have been presented in Chapter 4 and the clinical results have been published elsewhere (Anthierens et al. 2012; Little et al. 2013b; Yardley et al. 2013; Anthierens et al. 2015).

7.4.2 Resource use

Resource use data were collected under the same headings and categories as Case study 1.

7.4.3 Unit costs

Unit costs for valuing resource use data were obtained from the same sources as Case study 1 (Details have been presented in Table 7.2). In order to cost the trial interventions, information on resource use and costs was obtained through consultations with the trial coordinators. For the CRP training intervention, the costs of the machine was obtained from the manufacturer (Orion Diagnostica) who quoted a value of €1,200. This cost was annuitized assuming that the machine has a lifespan of 3 years and at an interest rate of 3.5% (Drummond et al. 2015)

and a cost per patient estimated. The costs of the reagents used (£6 per patient) were obtained from the provider (Oxford Biosystems).

With respect to the communication skills training intervention, the cost of the leaflet given to patients (£0.29) was obtained from study coordinators and converted to country equivalent using the market basket approach (Schulman et al. 1998). For the combined intervention, the cost of the CRP machine and the cost of leaflet estimated above were included.

In order to estimate the cost of training, information on the amount of time GPs spent on internet training in each arm was obtained from a published study (Yardley et al. 2013) and used to estimate the total cost of time spent on training. This value was divided by the number of patients per GP in order to estimate the cost per patient. GPs spent on average 26.54 minutes, 37.44 minutes and 39.76 minutes on training in the CRP, communication skills and combined intervention arms respectively (Yardley et al. 2013). GPs also received face to face training in CRP and a similar approach to what has been described above was used to estimate a cost per patient in each arm. All costs were converted to Euros using PPPs using the Euro area 17 countries as a reference (OECD, 2012). The costs were then converted into UK Pounds Sterling for the purpose of this study. All costs are presented in 2012 prices.

7.4.4 Health outcomes

Patients were asked to complete the 3 level version of the EQ-5D questionnaire over the entire four week period, or until they felt better and the index scores were generated using the European Harmonised Tariff (Greiner, 2003). Since participants were specifically told to stop completing the patient diary when they got well, thus, the last value carried forward approach was used to estimate EQ-5D scores for patients who stopped completing the EQ-5D questionnaire before week 4.

7.4.5 Antibiotic prescribing

Physicians were asked to state whether they prescribed one of the following antibiotics: amoxicillin, co-amoxiclav, penicillin V, other penicillin, doxycycline, other tetracycline, erythromycin, clarithromycin, azithromycin, other macrolide, levofloxacine, moxifloxacine, other quinolone, cefaclor, cefuroxime, other cephalosporin as well as any other antibiotic that was prescribed. This information was used to estimate the rate of antibiotic prescribing in each of the trial arms.

7.4.6 Data analysis

Two types of economic evaluations: a cost-effectiveness analysis (cost per unit reduction in antibiotic prescribing) and a cost-utility analysis (cost per QALY gained) were carried out on an intention to treat basis. QALYs, costs and net monetary benefits were estimated using similar methods to those used in Case study 1 (i.e. multilevel modelling). Dependent variables included total cost, QALYs, antibiotic prescribing and net monetary benefits. Regression model estimates of the difference in costs, QALYs, antibiotic prescribing and net monetary benefits were used to derive an incremental cost per QALY gained, an incremental net monetary benefit and an incremental cost per unit reduction in antibiotic prescribing.

To account for the factorial nature of the trial, a 'within the table' analysis was adopted (Dakin and Gray 2010; Frempong et al. 2015; Oppong et al. 2015a). This method assumes that the interventions are not independent i.e. the costs and effects of training in communication skills are influenced by the inclusion of training in CRP testing and vice versa. All interventions were ordered in terms of increasing cost, in order for costs, QALYs and antibiotic prescribing for each treatment arm to be compared incrementally. The most cost-effective option was selected based on the principles of dominance (where an intervention is less costly and more effective than the appropriate comparator(s)) and

extended (weak) dominance (where an intervention is ruled out if the ICER is greater than that of a more effective intervention) (Cantor, 1994).

7.4.7 Analytical approaches explored

Case study 2 considered three main analytical approaches: (I) The fully pooled one country costing approach, (II) the fully pooled multicountry costing approach and (III) the fully split one country costing approach. With Case study 2, the EU tariff was used to obtain EQ-5D index scores for the FSOC approach. The fully split multi country costing approach which was used in Case study 1 was not adopted in this study mainly due to the relatively small number of countries that were included in the analysis (five countries) and there was no real basis for splitting the five countries into groups.

7.5 Results: Case study 2

A total of 2624 participants who completed the patient diary and EQ-5D questionnaire at baseline were included in the study. The country contribution to sample size is given in Table 7.8.

Table 7.8 Country contribution to sample size (Case study 2)

Country	Usual care	CRP no comm.	Comm noCRP	CRP comm	Total number of patients
Belgium	19 (10.6%)	56 (31.1%)	61 (33.9%)	44 (24.4%)	180 (6.9%)
Netherlands	41 (17.9%)	45 (19.7%)	80 (35.1%)	62 (27.2%)	228 (8.7%)
Poland	229 (24.4%)	223 (23.7%)	208 (22.1%)	280 (29.8%)	940 (35.9%)
Spain	129 (15.9%)	225 (27.6%)	251 (30.8%)	209 (25.7%)	814 (31.0%)
UK	97 (21.0%)	111 (24.0%)	140 (30.3%)	114 (24.7%)	462 (17.6%)
Total	515 (19.6%)	660 (25.2%)	740 (28.2%)	709 (27.0%)	2,624 (100%)

7.5.1 Fully pooled one country costing (FPOC) (Approach I)

Resource use and costs (FPOC)

Compared to the other interventions, visits to the GP, nurse and hospital admissions were lower in the usual care arm. Visits to the GP were highest in the CRP group whilst visits to the nurse were highest in the communication skills group. As it was expected, those in the CRP and combined intervention groups had more CRP tests performed. Approximately 59% of participants in the usual care arm had an antibiotic prescribed compared to approximately 34% in the combined intervention and CRP arms respectively (Table 7.9).

Table 7.9 Resource use (All approaches)

		Usual care	CRP no Comm	Comm no CRP	CRP comm
		PRIMARY CARE VISITS			
GP visits	**FPOC/ FPMC**	0.194 (0.472)	0.355 (0.762)	0.284 (0.713)	0.236 (0.596)
	FSOC	0.134 (0.399)	0.117 (0.399)	0.171 (0.549)	0.140 (0.396)
Nurse Visits	**FPOC/ FPMC**	0.016 (0.206)	0.045 (0.323)	0.103 (0.741)	0.039 (0.263)
	FSOC	0.010 (0.102)	0.009 (0.095)	0.079 (0.341)	0.026 (0.161)
Out hours GP visits	**FPOC/ FPMC**	0.015 (0.271)	0.006 (0.095)	0.023 (0.182)	0.016 (0.163)
	FSOC	0.010 (0.102)	0.009 (0.095)	0	0.018 (0.187)
		SECONDARY CARE VISTIS			
Hospital emergency visits	**FPOC/ FPMC**	0.002 (0.044)	0.003 (0.054)	0.018 (0.134)	0.016 (0.155)
	FSOC	0	0	0	0.009 (0.094)
Walk in centre visits	**FPOC/ FPMC**	0.004 (0.087)	0.002(0.039)	0.022 (0.186)	0.035 (0.383)
	FSOC	0	0	0.007 (0.085)	0.009 (0.094)
Specialist visits	**FPOC/ FPMC**	0.004 (0.062)	0.018 (0.155)	0.028 (0.222)	0.023 (0.218)
	FSOC	0	0	0.021 (0.118)	0.009 (0.094)
Admissions	**FPOC/ FPMC**	0.010 (0.182)	0.026 (0.379)	0.019 (0.320)	0.030 (0.394)
	FSOC	0.010 (0.102)	0	0	0.061 (0.656)
				PRESCRIPTIONS n (%)	
Antibiotic prescription	**FPOC/ FPMC**	307 (59.61%)	222 (33.64%)	303 (40.95%)	242 (34.13%)
	FSOC	72 (74.23%)	47 (42.345)	66 (47.14%)	37 (32.46%)
Over the counter medication	**FPOC/ FPMC FSOC**	346 (67.18%)	419 (63.48%)	451 (60.95%)	441 (62.20%)
CRP test	**FPOC/ FPMC**	12 (2.33%)	441 (66.82%)	57 (7.70%)	461 (65.02%)
	FSOC	2 (2.06%)	91 (81.89%)	1 (0.71%)	55 (48.25%)

FPOC/FPMC (n=2,624), FSOC (n=462)

GP costs were highest in the CRP group whilst nurse costs were highest in the communication skills group. Overall costs were highest in the combined intervention arm and lowest in the usual care arm. The difference in total cost between usual care and all the other intervention arms were statistically significant (Table 7.10).

Outcomes (FPOC)

There was an improvement in health of participants over the 4-week period as shown by the EQ-5D scores (Table 7.11). QALYs were highest in the CRP group initially; however, when baseline EQ-5D was accounted for, the usual care group had the greatest improvement in QALYs. The difference in QALYs between the usual care arm and the other intervention arms were very small and not statistically significant (Table 7.11). Antibiotic prescribing was highest in the usual care group and lowest in the combined intervention group (Table 7.11).

Cost-Effectiveness (FPOC)

With respect to the cost-effectiveness analysis (antibiotic prescribing as an outcome), the results showed that communication skills is the most cost-effective option (Table 7.12). Compared to usual care, the ICERs for the combined intervention, communication skills and CRP were £169.81, £173.02 and £212.01 per unit reduction in antibiotic prescribing respectively.

The results from the cost-utility analysis (QALYs as an outcome) indicate that the CRP, communication skills and combined interventions were dominated by usual care (Table 7.12). At a cost-effectiveness threshold of £20,000 per QALY gained, net benefits associated with usual care was higher than all other interventions and the probability of usual care being cost-effective was 98%, 96% and 99% when compared with CRP, communication skills and the combined intervention (Figures 7.3 to 7.5).

Table 7.10 Costs (£) (All approaches)

		Usual care	CRP no Comm	Comm no CRP	CRP comm
PRIMARY CARE VISITS					
GP visits	**FPOC**	7.77 (18.89)	14.18 (30.50)	11.35 (28.55)	9.42 (23.85)
	FPMC	2.79 (8.33)	4.68 (11.23)	4.60 (13.90)	3.65 (10.12)
	FSOC	5.36 (15.95)	4.68 (15.94)	6.85 (21.96)	5.61 (15.86)
Nurse Visits	**FPOC**	0.27 (3.64)	0.80 (5.70)	1.81 (13.09)	0.70 (4.64)
	FPMC	0.18 (2.53)	0.32 (3.01)	1.36 (9.95)	0.49 (4.71)
	FSOC	0.18 (1.79)	0.16 (1.68)	1.39 (6.02)	0.46 (2.84)
Out hours GP visits	**FPOC**	4.38 (76.55)	1.71 (26.58)	6.48 (51.46)	4.38 (45.99)
	FPMC	4.30 (75.30)	2.04 (32.27)	8.07 (63.65)	5.36 (56.01)
	FSOC	2.91 (28.63)	2.54 (26.77)	0	4.95 (52.82)
SECONDARY CARE VISTIS					
Hospital emergency visits	**FPOC**	0.22 (4.94)	0.34 (6.16)	2.11 (15.27)	1.73 (17.27)
	FPMC	0.22 (5.05)	0.33 (6.07)	2.11 (15.19)	1.75 (17.28)
	FSOC	0	0	0	0.98 (10.49)
Walk in centre visits	**FPOC**	0.16 (3.61)	0.06 (1.60)	0.89 (7.64)	1.45 (15.72)
	FPMC	0.07 (1.65)	0.02 (0.73)	0.42 (3.67)	0.63 (6.41)
	FSOC	0	0	0.29 (3.47)	0.36 (3.84)
Specialist visits	**FPOC**	0.61 (9.77)	2.85 (24.30)	4.46 (34.85)	3.54 (34.22)
	FPMC	0.69 (10.98)	3.04 (25.72)	4.53 (36.18)	3.92 (37.88)
	FSOC	0	0	3.36 (29.58)	1.38 (14.70)
Admissions	**FPOC**	4.01 (75.00)	10.64 (156.45)	7.81 (132.21)	12.23 (163.07)
	FPMC	3.88 (72.65)	9.90 (145.37)	7.37 (122.15)	11.29(151.54)
	FSOC	4.26 (41.93)	0	0	25.36 (270.77)
OTHER COSTS					
Prescription	**FPOC**	3.07 (6.75)	3.21 (8.10)	4.06 (9.88)	3.65 (8.51)
	FPMC	9.70 (21.79)	7.09 (15.67)	7.94 (15.45)	9.73 (28.10)
	FSOC	1.24 (2.94)	1.54 (3.38)	1.68 (3.86)	1.58 (3.76)

		Usual care	CRP no Comm	Comm no CRP	CRP comm
OTC medication	**FPOC**	2.67 (6.20)	2.22 (5.38)	2.32 (5.72)	2.84 (7.01)
	FPMC	5.31 (14.08)	3.63 (10.51)	3.67 (10.26)	5.01 (14.05)
	FSOC	1.21 (1.25)	1.74 (3.59)	1.40 (1.74)	1.78 (3.56)
CRP test	**FPOC**	0.13 (0.87)	3.85 (2.71)	0.44 (1.53)	3.75 (2.74)
	FPMC	0.15 (1.00)	4.26 (3.03)	0.23 (0.87)	3.96 (3.07)
	FSOC	0.12 (0.82)	4.72 (2.22)	0.04 (0.48)	2.78 (2.89)
Trial intervention cost	**FPOC**	0	22.35 (0)	10.61 (0)	26.33 (0)
	FPMC	0	9.26 (6.04)	4.56 (2.99)	10.89 (6.92)
	FSOC	0	22.35 (0)	10.61 (0)	26.33 (0)
		TOTAL COSTS			
Total cost[a]	**FPOC**	23.68 (114.84)	62.45 (168.77)	52.78 (160.28)	70.58 (208.93)
Total cost[b]		24.93	60.40	51.76	70.30
Difference (95% CI)[c]			35.46	26.83	45.37
			(11.42, 59.51)	(3.40, 50.26)	(21.88, 68.86)
CI Width			48.09	46.86	46.98
Total cost[a]	**FPMC**	27.46 (111.05)	43.42 (153.93)	42.47 (144.73)	53.13 (195.92)
Total cost[b]		27.77	43.77	41.79	55.16
Difference (95% CI)[c]			15.99	14.02	27.39
			(-4.22, 36.21)	(-5.71, 33.74)	(7.57, 47.22)
CI Width			40.43	39.45	39.65
Total cost[a]	**FSOC**	15.68 (53.03)	38.45 (39.00)	26.77 (46.23)	72.28 (334.11)
Total cost[b]		17.12	27.00	36.10	73.07
Difference (95% CI)[c]			18.98	9.88	55.95
			(-26.70, 64.65)	(-33.50, 53.26)	(10.60, 101.30)
CI Width			91.35	88.76	90.70

[a] Unadjusted costs [b] Adjusted costs (3-level model) excludes cost of resistance [c] Difference with reference to usual care

FPOC/FPMC (n=2,624), FSOC (n=462)

7.5.2 Results: fully pooled multicountry costing (FPMC) (Approach II)

Resource use and costs (FPMC)

Similar to what was observed in Case study 1; resource use was similar to that obtained with the FPOC approach. GP costs were highest in the CRP group whilst nurse costs were highest in the communication skills group. Costs associated with over the counter medication were highest in the usual care arm and total costs were lowest in the usual care arm. (Table 7.10)

Outcomes (FPMC)

Health outcomes in terms of EQ-5D, QALYs and antibiotic prescribing were similar to the FPOC approach (Table 7.11).

Cost-Effectiveness (FPMC)

With respect to the cost-effectiveness analysis, training in CRP testing proved to be the most cost-effective intervention (Table 7.13). Compared to usual care, the ICERs for the combined intervention, communication skills and CRP were £127.98, £88.69 and £78.00 per unit reduction in antibiotic prescribing respectively.

The cost-utility analysis yielded similar results to that which was obtained with the FPOC approach i.e. CRP, communication skills and combined interventions were dominated by usual care (Table 7.13). At a cost-effectiveness threshold of £20,000 per QALY gained, the probability of usual care being cost-effective was 83%, 94% and 97% when compared with CRP, communication skills and the combined intervention (Figures7.3 - 7.5).

Table 7.11 Mean EQ-5D scores over 4 weeks (All approaches)

		Usual care	CRP no Comm	Comm no CRP	CRP comm
			EQ-5D		
Baseline	FPOC/ FPMC	0.718 (0.216)	0.729 (0.212)	0.693(0.228)	0.710 (0.223)
	FSOC	0.720 (0.205)	0.692 (0.218)	0.711 (0.199)	0.715 (0.233)
Week 1	FPOC/ FPMC	0.816 (0.197)	0.817 (0.207)	0.786 (0.214)	0.792 (0.210)
	FSOC	0.759 (0.202)	0.761 (0.220)	0.772 (0.219)	0.763 (0.228)
Week 2	FPOC/ FPMC	0.884 (0.176)	0.881 (0.182)	0.864 (0.185)	0.896 (0.186)
	FSOC	0.835 (0.203)	0.832 (0.196)	0.830 (0.206)	0.826 (0.221)
Week 3	FPOC/ FPMC	0.898 (0.170)	0.899 (0.176)	0.894 (0.176)	0.893 (0.174)
	FSOC	0.851 (0.200)	0.845 (0.210)	0.873 (0.186)	0.842 (0.223)
Week 4	FPOC/ FPMC	0.906 (0.165)	0.906 (0.170)	0.903 (0.168)	0.899 (0.169)
	FSOC	0.869 (0.197)	0.858 (0.205)	0.882 (0.187)	0.849 (0.209)
			QUALITY ADJUSTED LIFE YEARS		
QALYs[a]	FPOC/ FPMC	0.0656 (0.012)	0.0657 (0.013)	0.0643 (0.012)	0.0646 (0.013)
QALYs[b]		0.0649	0.0647	0.0648	0.0643
Difference (95% CI)[c]			-0.0002 (-0.002, 0.001)	-0.0001 (-0.002, 0.001)	-0.0006 (-0.002, 0.001)
CI Width			0.003	0.003	0.003
QALYs[a]	FSOC	0.062 (0.014)	0.062 (0.015)	0.063 (0.013)	0.062 (0.015)
QALYs[b]		0.062	0.063	0.063	0.061
Difference (95% CI)[c]			0.001 (-0.002, 0.004)	0.001 (-0.002, 0.004)	-0.003 (-0.003, 0.002)
CI Width			0.006	0.006	0.005

		Usual care	CRP no Comm	Comm no CRP	CRP comm
ANTIBIOTIC PRESCRIBING					
Antibiotic Prescribing [a]	FPOC/ FPMC	0.596 (0.491)	0.336 (0.473)	0.409 (0.492)	0.341 (0.474)
Antibiotic Prescribing [b]		0.555	0.350	0.397	0.341
Difference (95% CI)[c]			-0.204 (-0.304, -0.103)	-0.157 (-0.260, -0.054)	-0.213 (-0.309, -0.118)
CI Width			0.201	0.206	0.191
Antibiotic Prescribing [a]	FSOC	0.742 (0.440)	0.423 (0.496)	0.471 (0.501)	0.325 (0.470)
Antibiotic Prescribing [b]	FSOC	0.744	0.416	0.423	0.382
Difference (95% CI)[c]			-0.328 (-0.547, -0.109)	-0.321 (-0.534, -0.107)	-0.362 (-0.578, 0.145)
CI Width			0.438	0.427	0.723

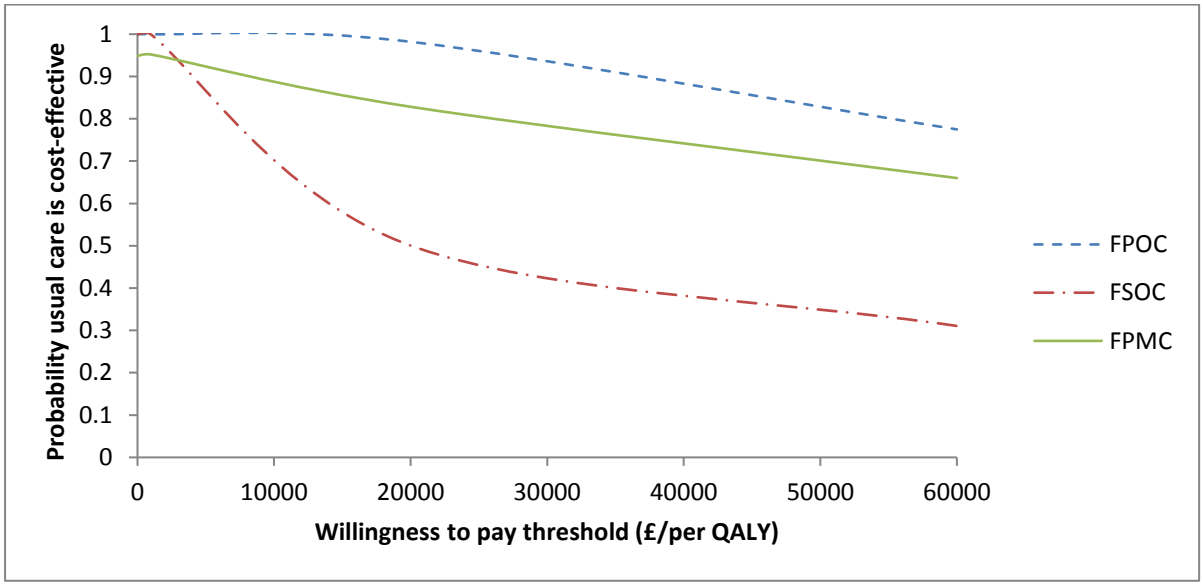
[a] Unadjusted [b] Adjusted (3-level model) [c] Difference with reference to usual care **FPOC/FPMC (n=2,624), FSOC (n=462)**

Table 7.12 Cost-effectiveness analysis (FPOC)

	Mean cost (£)	QALYs	NMB (£20,000)	ICER (£ per QALY)
COST-UTILITY ANALYSIS				
CRP comm (709)	70.30	0.0643	1219.39	Dominated by usual care
CRP no comm (660)	60.40	0.0647	1235.36	Dominated by usual care
Comm no CRP (740)	51.76	0.0648	1247.55	Dominated by usual care
Usual care (515)	24.93	0.0649	1274.57	N/A
COST-EFFECTIVENESS ANALYSIS			**ICER (£ per unit reduction in prescribing)**	
CRP comm (709)	70.30	0.341		1,100 [a]
CRP no comm (660)	60.40	0.350		183.83[b]
Comm no CRP (740)	51.76	0.397		169.81[c]
Usual care (515)	24.93	0.555		N/A

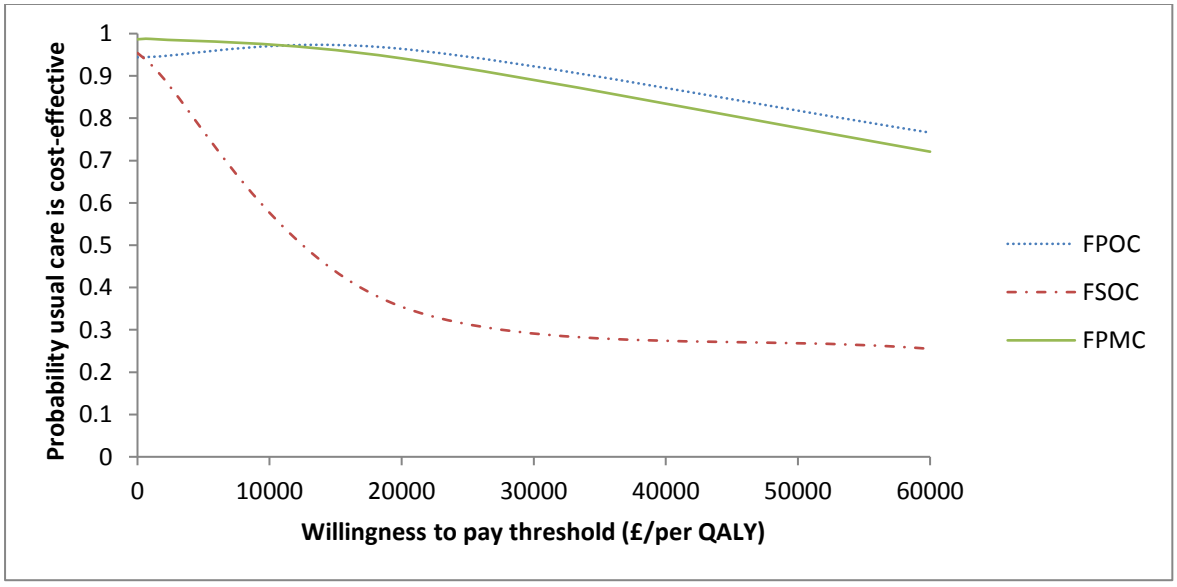
[a] ICER derived from a comparison of CRP comm with CRP no comm [b] ICER derived from a comparison of CRP no comm with Comm no CRP [c] ICER derived from a comparison of CRP no comm and Usual care

Figure 7.3 Cost-effectiveness acceptability curves for usual care vs CRP (Case study 2)



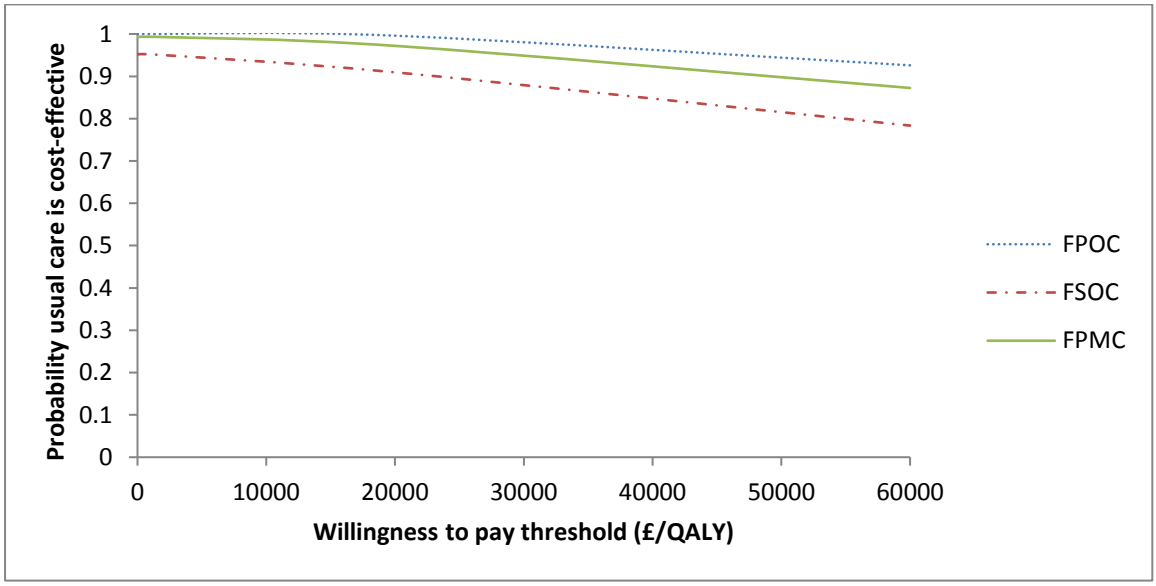
FPOC: Fully pooled one country costing FSOC: Fully split one country costing FPMC: Fully pooled multicountry costing

Figure 7.4 Cost-effectiveness acceptability curves for usual care vs comm skills (Case study 2)



FPOC: Fully pooled one country costing FSOC: Fully split one country costing FPMC: Fully pooled multicountry costing

Figure 7.5 Cost-effectiveness acceptability curves for usual care vs combined intervention (Case study 2)



FPOC: Fully pooled one country costing FSOC: Fully split one country costing FPMC: Fully pooled multicountry costing

7.5.3 Results: fully split one country costing (FSOC) (Approach III)

This analysis was limited to the 462 participants who were recruited from the UK.

Resource use and costs (FSOC)

Compared to the other interventions, visits to the GP, nurse and specialist were higher in the communication skills arm whilst hospital admissions were highest in the combined intervention arm. Approximately 74% of participants in the usual care arm had an antibiotic prescribed compared to approximately 32% in the combined intervention arms (Table 7.9).

Costs associated with resource use items are presented in Table 7.10 below. Total costs were lowest in the usual care arm and with the exception of the combined intervention; all other differences in costs compared to usual care were not statistically significant.

Outcomes (FSOC)

QALYs were equivalent in the usual care, CRP and combined intervention groups initially and when baseline EQ-5D was accounted for, the CRP and communication skills group recorded the greatest improvement in QALYs (Table 7.11). Antibiotic prescribing was highest in the usual care group and lowest in the combined intervention group (Table 7.11).

Table 7.13 Cost-effectiveness analysis (FPMC)

	Mean cost (£)	**QALYs**	**Net Monetary Benefits**	**ICER (£ per QALY)**
		Cost-Utility Analysis		
CRP comm (709)	55.16	0.0643	1235.17	Dominated by usual care
CRP no comm (660)	43.77	0.0647	1252.70	Dominated by usual care
Comm no CRP (740)	41.79	0.0648	1256.97	Dominated by usual care
Usual care (515)	27.77	0.0649	1273.75	N/A
		Cost-Effectiveness Analysis		
				ICER (£ per unit reduction in prescribing)
CRP comm (709)	55.16	0.341		1,266[a]
CRP no comm (660)	43.77	0.350		42.13[b]
Comm no CRP (740)	41.79	0.397		88.73[c]
Usual care (515)	27.77	0.555		N/A

[a] ICER derived from a comparison of CRP comm and CRP no comm [b] ICER derived from a comparison of CRP no comm with Comm no CRP [c] ICER derived from a comparison of Comm no CRP with usual care

Cost-effectiveness (FSOC)

Results from the cost-effectiveness analysis show that communication skills was the most cost-effective intervention (Table 7.14). Compared to usual care, the ICERs for the combined intervention, communication skills and CRP were £154.56, £30.78 and £57.87 per unit reduction in prescribing respectively.

With respect to the cost-utility analysis, when compared to usual care, communication skills was associated with an ICER of £9,880 per QALY gained whilst CRP and the combined intervention were dominated by communication skills (Table 7.14). The probability of usual care being cost-effective was 50%, 35% and 90% when compared with CRP, communication skills and the combined intervention at the £20,000 threshold (Figures 7.3 to 7.5).

Table 7.14 Cost-effectiveness analysis (FSOC)

	Mean cost (£)	**QALYs**	**NMB (20,000)**	**ICER**
COST-UTILITY ANALYSIS				
CRP comm (114)	73.07	0.061	1156.67	Dominated by Comm noCRP
CRP no comm (111)	36.10	0.063	1216.60	Dominated by Comm noCRP
Comm no CRP (140)	27.00	0.063	1229.39	£9,880 per QALY gained
Usual care (97)	17.12	0.062	1218.84	N/A
COST-EFFECTIVENESS ANALYSIS				
				ICER £ per unit reduction in prescribing
CRP comm (114)	73.07	0.382		1,087[a]
CRP no comm (111)	36.10	0.416		1,300[b]
Comm no CRP (140)	27.00	0.423		30.78[c]
Usual care (97)	17.12	0.744		N/A

[a] ICER derived from a comparison of CRP comm and CRP no comm [b] ICER derived from a comparison of CRP no comm with Comm no CRP [c] ICER derived from a comparison of Comm no CRP with usual care

7.6 Comparison of results from analytical approaches (Case studies 1 and 2)

The previous section presents the base case analysis exploring the various pooled and split approaches to economic analysis alongside multinational trials. Table 7.15 provides a summary of the main results from each of the analytical approaches. A comparison of the approaches in terms of resource use, costs, outcomes and cost-effectiveness is now presented

in the section below in order to determine the similarities and differences between the approaches. It should be noted that a comparison of the FSMC and FSOC approaches could not be made for Case study 2 since the FSMC approach was not undertaken in Case study 2.

Table 7.15 Summary of results obtained using various approaches to economic evaluation alongside multinational trials

Approach	Case study 1 (Cost-utility analysis)	Case study 2 (Cost-utility analysis)	Case study 2 (Cost-effectiveness analysis)
Fully pooled one country costing	Amoxicillin dominant	Usual care dominant	Communication skills cost-effective £169.81 per prescription avoided
Fully pooled multicountry costing	Amoxicillin cost-effective £6,540 per QALY	Usual care dominant	CRP cost-effective £42.13 per prescription avoided
Fully split one country costing	Control dominant	Communication skills cost-effective £9,880 per QALY gained	Communication skills cost-effective £30.78 per prescription avoided
Fully split multicountry costing	Amoxicillin cost-effective £2,986 per QALY	N/A	N/A

7.6.1 Resource use

Fully pooled multicountry costing versus fully pooled one country costing

The values obtained with the FPOC and FPMC approaches were the same therefore; adopting either approach would not make a difference to the conclusions drawn about health care resources that are used up as a result of the interventions. This was true for both case studies (Table 7.4 and Table 7.9).

Fully split multicountry costing versus fully split one country costing

For Case study 1, out of hour GP visits and prescriptions were higher in the amoxicillin arm with the FSMC approach whilst GP visits, out of hour GP visits, hospital emergency visits and prescribed medication were higher in the intervention arm with the FSOC approach (Table 7.4). This suggests that a choice between these two approaches could lead to conflicting conclusions with respect to whether the intervention or control is associated with more of a particular resource item.

Fully pooled one-country costing versus fully split one country costing

With Case study 1, the use of over the counter medication was highest in the amoxicillin arm with the FPOC analysis whilst the opposite was true for the FSOC analysis. In some cases, such as visits to the specialist, there was no record of any visit with the FSOC approach (Table 7.4). For Case study 2, GP visits were highest in the communication skills group for the FPOC approach and highest in the CRP group for the FSOC approach. Some intervention arms had no record of some resource items such as hospital emergency visits, specialist visits and hospital admissions for the split analysis (Table 7.9). Both case studies suggest that the FSOC approach might lead to a situation in which some items may be missed out possibly due to the sample size or the health system.

7.6.2 Costs

Fully pooled multicountry costing versus fully pooled one country costing

Results from Case study 1 showed that total staff costs were higher in the control group with the FPOC approach but higher in the intervention group with the FPMC approach. Most other differences in costs were in the same direction, although the magnitude of the differences was not the same. Overall, the amoxicillin group was associated with higher costs per patient with the FPMC whilst the opposite was observed with the FPOC approach (Table 7.5). In both

cases, the difference in overall cost between the intervention and control groups was not statistically significant.

With Case study 2, costs were similar in terms of direction for both analytical approaches for most resource use items. The only differences were observed for prescribed medication and over the counter medication where the highest costs were recorded in different intervention arms (Table 7.10). Total costs were highest in the combined intervention arm and lowest in the usual care arm for both the FPMC and FPOC approaches.

Fully split multicountry costing versus fully split one country costing

Staff related costs were higher in the intervention group with the FSOC approach whilst the FSMC approach showed that the control group was associated with a higher staff cost for the FSMC approach (Table 7.5).

Fully pooled one-country costing versus fully split one country costing

With Case study 1, the cost differences between the various resources use items were similar for most resource use items (i.e. they were in the same direction). However, the overall mean cost per patient was higher in the intervention arm with FSOC whilst the control arm was the most expensive overall with FPOC (Table 7.5). With Case study 2, out of hours GP was associated with a zero cost in the communication skills arm with FSOC approach whilst the highest cost value was recorded in the communication skills arm with the FPOC (Table 7.10).

7.6.3 Health outcomes

Fully pooled multicountry costing versus fully pooled one country costing

Similar to what was seen with resource use, both case studies showed that EQ-5D and QALYs were the same with both approaches.

Fully split multicountry costing versus fully split one country costing

EQ-5D values were higher in the amoxicillin group over the four week period with FSMC approach whilst with the exception of week one, EQ-5D scores in the control group were higher with the FSOC approach (Table 7.6).

Fully pooled one-country costing versus fully split one country costing

With Case study 1, health outcomes between the two approaches were quite similar with the exception of week 4 where EQ-5D scores were the same in both trial arms for the FPOC approach but greater in the control arm for FSOC (Table 7.6). For Case study 2, QALYs were highest in the CRP and communication skills arms for the split analysis and highest in the usual care arm for the pooled analysis indicating that different strategies are considered as the most effective with the different approaches (Table 7.11).

7.6.4 Cost-effectiveness

Fully pooled multicountry costing versus fully pooled one country costing

With Case study 1, the ICER for the fully pooled multicountry costing approach was £6,540 per QALY gained whilst the intervention was dominant (less costly and more effective) with the FPOC approach. At a threshold of £20,000 per QALY gained the CEACs showed that there was an 86% chance that amoxicillin is cost-effective with the FPMC approach and an 88% chance that the intervention is cost-effective with the FPOC approach (Figure 7.1 and Table 7.7). For Case study 2, the cost-effectiveness analysis showed that CRP was the most cost-effective strategy with the FPMC approach whilst communication skills was cost-effective with the FPOC approach. With the cost-utility analysis, both approaches (FPOC and FPMC) resulted in a conclusion that usual care is the most cost-effective strategy. In both cases, usual care dominated all other strategies that were considered (Table 7.12 and Table 7.13).

Fully split multicountry costing versus fully split one country costing

The ICERs obtained with the FSMC approach in Case study 1 for amoxicillin was £2,986 per QALY gained whilst the control was dominant (less costly and more effective) with the FSOC approach. The result from this analysis suggests that at a cost-effectiveness threshold of £20,000 per QALY gained different conclusions are reached with each approach (Table 7.7). There was a 50% and 76% chance that amoxicillin is cost-effective with the FSOC and FSMC approaches respectively (Figure 7.1).

Fully pooled one-country costing versus fully split one country costing

With Case study 1, two very different results were obtained with the approaches. The intervention (amoxicillin) was dominant with the FPOC approach whilst the control was dominant with the FSOC approach (Table 7.7). For Case study 2, a similar result was obtained with the cost-effectiveness analysis that was conducted with both approaches (communication skills was once again the most cost-effective option). With respect to the cost-utility analysis however, usual care was dominant with the FPOC approach whilst communication skills was cost-effective (£9,880 per QALY gained) with the FSOC. At a cost-effectiveness threshold of £20,000 per QALY gained, different conclusions can be drawn from the different approaches.

7.7 Sensitivity analysis

A number of sensitivity analyses were carried out with both case studies (Case study 1 and Case study 2).

Sensitivity analysis had two main foci:

(I) To explore the impact of the choice of EQ-5D tariff by estimating cost-effectiveness using the UK EQ-5D tariff (Dolan, 1997) and the EU tariff (Greiner et al. 2003). This

analysis was limited to the cost-utility analysis conducted alongside Case study 1 and Case study 2.

(II) Explore statistical approaches for the estimation of country-specific cost-effectiveness.

(III) To explore the effects of imputation by limiting the analysis to participants who returned the patient diary.

(I) The impact of choice of tariff for obtaining EQ-5D index scores

Issues relating to the choice of value set for valuing EQ-5D were discussed in Chapter 6 and this analysis explores this issue further by assessing the impact a different tariff would have on the results of the study. To achieve this, participant responses to the EQ-5D questionnaire were valued using the UK value set (Dolan, 1997) and EU value set (Greiner et al. 2003) and QALYs were generated using the same approach described earlier. This was used to generate incremental cost-effectiveness ratios. The results obtained were then compared to those from the base case analysis to see if there were any important differences.

Results (I) The impact of choice of tariff for obtaining EQ-5D index scores

This analysis was carried out using both case studies and the results are presented in Table 7.16 and Table 7.17 and they clearly show a very similar finding with Case study 1 and Case study 2 thus strengthening the finding that either the UK or EU value set can be used interchangeably within multinational trials. The only exception was with the FSOC approach in Case study 1 where the difference in QALYs obtained with the EU and UK tariffs were in the opposite direction. However, the conclusions with both tariffs indicate that amoxicillin was not cost-effective.

Table 7.16 Cost-effectiveness UKVS and EVS (Case study 1)

		UK tariff	EU tariff
Difference in costs (CI)	**FPOC**	£-0.661 (-4.69, 3.37)	£-0.661 (-4.69, 3.37)
	FPMC	£2.42 (-1.14, 5.98)	£2.42 (-1.14, 5.98)
	FSOC	£3.84 (-1.98, 9.67)	£3.84 (-1.98, 9.67)
	FSMC	£2.09 (-3.32, 7.49)	£2.09 (-3.32, 7.49)
Difference in QALYs (CI)	**FPOC/FPMC**	0.00034 (-0.0003, 0.001)	0.00037 (-0.0002, 0.0009)
	FSOC	-0.00008 (-0.002, 0.002)	0.00019 (-0.002, 0.002)
	FSMC	0.00071 (-0.0003, 0.002)	0.0007 (-0.0002, 0.0016)
ICER	**FPOC**	Intervention dominant	Intervention dominant
	FPMC	£7,118 per QALY gained	£6,540 per QALY gained
	FSOC	Intervention dominated by placebo	£20,210 per QALY gained
	FSMC	£2,944 per QALY gained	£2,986 per QALY gained

FPOC/FPMC (n=2060), FSOC (n=329), FSMC (n=904)

Table 7.17 Cost-effectiveness UKVS and EVS (Case study 2)

	Mean cost (£)	QALY (UK tariff)	QALY (EU tariff)	ICER (UK tariff)	ICER (EU tariff)
FPOC approach					
CRP comm (709)	70.30	0.0629	0.0643	Dominated by usual care	Dominated by usual care
CRP no comm (660)	60.40	0.0695	0.0647	Dominated by usual care	Dominated by usual care
Comm no CRP (740)	51.76	0.0658	0.0648	Dominated by usual care	Dominated by usual care
Usual care (515)	24.93	0.0697	0.0649	N/A	N/A
FPMC approach					
CRP comm (709)	55.16	0.0629	0.0643	Dominated by usual care	Dominated by usual care
CRP no comm (660)	43.77	0.0695	0.0647	Dominated by usual care	Dominated by usual care
Comm no CRP (740)	41.79	0.0658	0.0648	Dominated by usual care	Dominated by usual care
Usual care (515)	27.77	0.0697	0.0649	N/A	N/A
FSOC approach					
CRP comm (114)	73.07	0.0601	0.061	Dominated by Comm noCRP	Dominated by Comm noCRP
CRP no comm (111)	36.10	0.0651	0.063	Dominated by Comm noCRP	Dominated by Comm noCRP
Comm no CRP (140)	27.00	0.0652	0.063	£4,940 per QALY gained	£9,880 per QALY gained
Usual care (97)	17.12	0.0629	0.062	N/A	N/A

(II) Exploration of statistical approaches for obtaining country-specific cost-effectiveness

From earlier chapters, it was mentioned that decision makers are interested in results that are limited to their own countries and as a result, they might not be interested in the results that are obtained from multinational trials. A number of approaches for estimating this country-specific cost-effectiveness have been suggested in the literature which includes the following: (i) Hypothesis tests of homogeneity of results across countries (ii) multivariable cost or outcome regressions to adjust for country effects and (iii) Hierarchical models with shrinkage estimators (Manca et al. 2007; Ramsey et al. 2015). These and other approaches have been assessed and it has been shown that hierarchical modelling can facilitate the estimation of country-specific results through shrinkage estimation which assumes that even though data is collected from different countries, to some extent, there are some similarities which allow country-specific estimates to borrow strength from each other. Thus, country-specific estimates are shrunken towards the overall estimate based on the sample size in the country of interest and the sampling variability around the country-specific estimate (Manca et al. 2007).

This approach was applied to estimate the country-specific cost-effectiveness for all countries that participated in Case study 1. However, the focus was on the UK since the FSOC approach presented above only focused on UK participants. Two different scenarios were considered. The first was a one country costing approach where UK costs were used in the model and the second was a multicountry costing approach where country-specific costs were applied. The model was implemented in Winbugs 14 (Appendix 12).

Results (II) Statistical approaches to obtaining country-specific cost-effectiveness

The results from the Bayesian hierarchical model showed that overall amoxicillin was dominant (less expensive and more effective) with one country costing but more costly and

more effective (cost-effective with an ICER of £4,200 per QALY gained) with multicountry costing. Overall, the probability of antibiotics being cost-effective at a threshold of £20,000 per QALY gained ranged from 65% with the multicountry costing approach to 76% with the one country costing approach (Table 7.18 and Appendix 13). Country-specific results are presented in Table 7.18 and they show that in more than 50% of countries (6 out of 11 countries), the results were similar to the overall result with the multicountry costing approach, however, with the one country costing approach, only 4 countries had results that matched the overall. The results from the UK showed that amoxicillin was more costly and more effective with both the one country and multicountry costing approaches. This was similar to the results that were obtained with the FPMC approach. It should be noted that the FSOC approach which considered only UK participants yielded results which showed that the control was dominant. Eight out of the eleven countries also obtained similar results in terms of the direction of difference in costs and difference in outcomes with both the multicountry and one country costing. The only exceptions were Belgium, Spain and Sweden (Table 7.18).

Table 7.18 Country-specific cost-effectiveness results obtained from Bayesian hierarchical model

Country		Incremental costs	Incremental QALY (EVS)	Incremental QALY (UKVS)	ICER Per QALY (EVS)	Net benefit (£20,000 per QALY) (EVS)	Probability cost-effective at £20,000 per QALY (EVS)
Belgium	OC	-2.14	0.002	0.002	N/A	32.28	97%
	MC	4.83	0.002	0.002	£2,415	25.19	94%
France	OC	-7.82	-0.00022	-0.00017	N/A	3.36	54%
	MC	-6.06	-0.00022	-0.00017	N/A	1.29	53%
Germany	OC	1.20	0.00041	0.0001	£2,926	7.05	58%
	MC	6.15	0.00041	0.0001	£15,000	2.57	51%
Italy	OC	-6.01	0.0011	0.0010	N/A	27.65	77%
	MC	-3.50	0.0011	0.0010	N/A	24.56	75%
Netherlands	OC	-6.00	0.00051	0.00090	N/A	16.15	83%
	MC	-2.55	0.00051	0.00090	N/A	12.61	78%
Poland	OC	3.07	0.00068	0.00061	£4,517	10.69	81%
	MC	4.79	0.00068	0.00061	£7,044	8.76	76%
Slovakia	OC	-9.97	-0.00076	-0.00058	N/A	-5.29	41%
	MC	-3.70	-0.00076	-0.00058	N/A	-11.07	33%
Spain	OC	-1.04	-0.0011	-0.0014	N/A	-21.36	8%
	MC	1.44	-0.0011	-0.0014	N/A	-23.67	7%
Sweden	OC	-6.12	0.00043	0.00052	N/A	14.76	72%
	MC	3.72	0.00043	0.00052	£8,651	4.31	56%
UK	OC	2.77	0.00035	0.00001	£7,914	4.17	61%
	MC	5.28	0.00035	0.00001	£15,085	1.66	54%
Slovenia	OC	4.83	0.0016	0.0014	£3,019	26.5	85%
	MC	8.81	0.0016	0.0014	£5,506	21.2	77%
Overall	OC	-2.48	0.00040	0.00034	N/A	10.66	76%
	MC	1.68	0.00040	0.00034	£4,200	6.19	65%

OC= One country costing MC=multicountry costing EVS=European value set UKVS=UK value set

(III) Limiting the analysis to participants with complete diaries

Sensitivity analysis was carried out to validate the multiple imputation technique used in this analysis. To achieve this, the analysis comparing the various approaches to economic evaluation was limited to participants who completed the patient diary over the four week period. Similar methods described in the methods section was used to estimate incremental costs, incremental QALYs and ICERs for each of the approaches.

Results (III) Limiting the analysis to participants with complete diaries

For the FPOC and FPMC approaches, a total of 1687 (81%) participants completed the patient diary. A similar proportion 243 (74%) and 720 (79%) completed the diaries for the FSOC and FSMC approaches respectively. The results obtained from each approach were in line with that which was obtained in the main analysis i.e. amoxicillin was dominant with the FPOC approach, cost-effective with the FPMC and FSMC approaches and was dominated by the control with the FSOC approach (Table 7.19).

Table 7.19 Sensitivity analysis: Participants with complete cases (All approaches)

Difference in costs (CI) {CI width}	**FPOC**	£- 1.20 (-5.98, 3.58)
	FPMC	£2.73 (-2.22, 8.26)
	FSOC	£4.97 (-2.49, 12.43)
	FSMC	£2.50 (-5.59, 10.60)
Difference in QALYs (CI) {CI width}	**FPOC/FPMC**	0.00035 (-0.00036, 0.0011)
	FSOC	-0.00015 (-0.002, 0.002)
	FSMC	0.0006 (-0.0005, 0.0017)
ICER	**FPOC**	Intervention dominant
	FPMC	£7,800 per QALY gained
	FSOC	Intervention dominated by placebo
	FSMC	£4,166 per QALY gained

FPOC/FPMC (n=1687), FSOC (n=243), FSMC (n=720)

7.8 Discussion

7.8.1 Summary of main findings

This chapter has compared various analytical approaches for the conduct of economic evaluation alongside multinational trials using two case studies (Case study 1 and Case study 2) from the GRACE project. The main results, which were confirmed by both case studies, show that the choice of analytical approach could possibly lead to different conclusions about the cost-effectiveness of the interventions under consideration. For the comparison between amoxicillin and control, (Case study 1) amoxicillin was cost-effective at the £20,000 per QALY threshold with three of the approaches (FPOC, FPMC and FSMC). However with the FSOC approach where the analysis was limited to UK participants, the control was dominant. Even though three approaches confirmed the cost-effectiveness of the intervention, the probability of the intervention being cost-effective varied (Figure 7.1), a result which might possibly have an influence on a decision maker who is tasked with deciding on whether to adopt the health technology in their country/jurisdiction.

With Case study 2, more variation was observed with the different analytical approaches. For example, the cost-utility analysis showed that usual care was dominant with the FPOC and FPMC costs and communication skills cost-effective with the FSOC approach whilst the cost-effectiveness analysis showed that communication skills was cost-effective with the FPOC and FSOC costing approach (Table 7.15). Case study 2 considered both cost-effectiveness and cost-utility analysis and there was no evidence

that conducting one approach over the other would reduce the variation in the results that are obtained with the various approaches.

The study also considered differences in *resource use* obtained from the various approaches and found that there was no record of some resource use items when the FSOC approach was used. This is quite worrying because quite a lot of information about potential resource use information could be lost with this approach. However, it could be possible that the non-use of certain resource items in a particular country/health system might be responsible for this. Even though there was the same amount of resource use for the fully pooled approaches (one and multicountry costing), this study has shown that the price weights used can determine whether an intervention is cost-effective or not (see results from FPOC and FPMC). It could be argued that even when the one country approach is used the results might still not be relevant to decision makers in a particular country due to the fact that resource use in certain countries may not really reflect that of the country of interest. Possible solutions that have been suggested in the literature include adjusting the resource use in the study to reflect those in the country of interest (Koopmanschap et al. 2001). However, it can be argued that unless the relationship between resource use in the country of interest and other countries is well established, adjusting resource use in this way can equally lead to overestimation or underestimation of resources consumed. From the review of the literature in Chapter 3, it was noted that within the context of a clinical trial, protocols should be followed strictly and this could potentially remove variation in resource use between countries (Lofdal et al. 2005). If this is the case however, then the question that should be asked is which country is the protocol most relevant to?

In terms of *health outcomes*, there were some differences between the analytical approaches e.g. with Case study 1, the FSOC approach suggested that the intervention is associated with a lower health gain in terms of QALYs whilst the opposite was true with the other approaches. This result suggests that interventions might be more effective in particular settings as a result of adjunct treatments, training and socio-demographic factors (Koopmanschap et al. 2001). Thus, care must be taken when generalizing the effectiveness of an intervention across jurisdictions. With the exception of the FSOC approach, sensitivity analysis yielded similar results with the EVS and UKVS and this supports the finding that the EVS and UKVS can be used interchangeably within the context of a multinational trial (Bernert et al. 2009). However, as was suggested in Chapter 6, the impact of both should be considered within sensitivity analysis. The CEA conducted alongside Case study 2 resulted in a consistent result across all analytical approaches i.e. usual care was associated with the highest number of antibiotic prescriptions whilst the combined intervention was associated with the lowest prescriptions. Therefore, it may seem that the effectiveness of an intervention measured with a disease specific measure may be more generalizable across jurisdictions/countries than those measured with a generic measure such as the EQ-5D. However, care must be taken when interpreting this finding since this study was limited to a specific disease area. The interventions considered in Case study 2 (training in communication skills and CRP) are primarily aimed at reducing the prescription of antibiotics by GPs and a potential question is whether the QALY, which is focused primarily on measuring health gain, should be the main outcome measure for interventions of this sort. This is because withholding antibiotics may actually lead to a reduction in health in the short-run (Cals et

al. 2011) and as this study has shown, the interventions led to a slight reduction in current health gain (QALYs). However, this reduction in health gain was very small and not statistically significant and as a result we cannot conclude that patient outcomes have been affected adversely by the intervention. On the other hand, it has been well established that a reduction in antibiotic prescribing has the potential to reduce antibiotic resistance and as a direct result lead to health gains in the future. Thus, if the QALY is to be used as an outcome measure, it is suggested that the impact of antibiotic resistance should be accounted for. This suggestion could also be extended to other clinical areas, where the cost per QALY approach has been questioned.

The differences in *costs* between interventions were generally small and not statistically significant and in most cases the approach that was adopted did not really change the significance of the findings one might be tempted to say that the approaches lead to similar findings in terms of costs and QALYs. However, it should be made very clear that hypothesis/significance testing is not the focus of economic evaluation in health care. The focus is on estimation (Briggs et al. 2001; Dakin et al. 2013). Therefore a small non-significant difference in costs and outcomes could make a difference to the results of the study. This is evidenced by the results from Case study 1 where adopting the FPOC approach led to a situation where there was a change in the direction of the difference and a conclusion that the intervention is cheaper than the control whilst all the other approaches showed that the intervention is more costly. Sensitivity analysis sought to estimate country-specific cost-effectiveness with approaches that have been suggested in the literature and in some cases; the results were shown to be sensitive to whether a one country or multicountry costing approach was used.

7.8.2 Strengths and limitations of the study

There are a number of strengths of this study: First, this is the first study of interventions that are aimed at tackling antibiotic resistance that have considered the pooled and split approaches to the conduct of economic evaluation alongside multinational trials. Second, one of the main utilities that has been derived from this thesis is that in addition to showing the possible differences in cost-effectiveness of interventions that may occur as a result of using different approaches, this study also tried to explain why these differences occur by considering resource use, health outcomes and unit costs individually. Third, this study also used two different case studies one normal two arm trial (Case study 1) and a more complex design (Case study 2) in assessing the approaches to economic evaluation alongside multinational trials.

There are a few limitations associated with this study. First, the case studies that were considered were all related to the area of antibiotic resistance and as such the results obtained might not be generalisable to other disease areas. However, it is believed that most of the issues such as those relating to costs and outcomes can be generalised to other disease areas. Second, Case study 1 compared amoxicillin with placebo which may not be the most appropriate comparator in an economic evaluation. However, it has been noted that if usual care is to do nothing, which is the case here, then placebo controlled trials can be useful for economic evaluation (Drummond 1996). Third, Case study 2 is conducted alongside a cluster randomised, factorial controlled trial and therefore presents additional complexities with respect to the analysis of the data. The factorial nature has the effect of reducing the sample size and also increasing the degree of uncertainty in the economic data (Oppong et al. 2015a; Frempong et al. 2015). In this study, randomisation

took place at the cluster/practice level whilst health economics outcomes such as QALYs were measured at the level of the individual. However, an effort has been made to account for this through the use of methods that account for the hierarchical nature of the data.

7.8.3 Comparison with existing literature

Two studies have looked at some of consequences of pooling and splitting data within multinational trials (Willke et al. 1998; Cook et al. 2003) and they obtained similar results to those that were obtained in our study i.e. pooling and splitting the data might lead to different results. However, these studies did not explicitly consider the analytical approaches and they were also in different disease areas i.e. cardiovascular disease. Also the focus of both studies was solely on estimating country-specific cost-effectiveness. These studies also did not consider the cost-per QALY approach something which the work presented in this chapter adds. Another study (Reinhold et al. 2010) discussed the consequences of pooling and splitting data within multinational trials however, like the others studies, this study did not explicitly consider the various analytical approaches and did not use empirical data to support the findings. In addition, the work discussed in this chapter pays close attention to issues relating specifically to effectiveness data, resource use and costs.

7.8.4 Implications for policy and practice

For a researcher who is faced with the analysis of a multinational trial, it is important to understand the pros and cons associated with pooling and splitting data. From the literature, there are clear advantages associated with pooling and splitting the data which can be clearly inferred (Wilke et al. 1998, Cook et al. 2003). Pooling data across

countries results in less transferability to a single country and greater statistical power whilst splitting the data results in increased transferability to a single country but results in a loss of statistical power (Reinhold 2010). By conducting the fully pooled one country costing approach which is the most common method used, the analyst assumes that resource use and effectiveness data are transferable between countries (Manca et al. 2007). With the fully split one country costing approach, there is also the assumption that data are not transferable between countries and potential problems could occur if this approach were adopted in a case where the recruitment from the country is low (Bachert et al. 2007; Canonica et al. 2007). For example, with Case study 1, recruitment was very low in countries like France (28 participants), Germany (18 participants) and Italy (8 participants) therefore adopting the FSOC approach would be almost impossible. Statistically, the pooled approaches, specifically the FPMC approach seems to be the most appropriate since it has the least uncertainty (smallest confidence intervals) around the results.

For results from multinational trials to be practically useful, it is important that they are relevant at the national level i.e. estimate acceptable country-specific cost-effectiveness (Manca et al. 2010) and it would seem that the simplest way to achieve this is to conduct the FSOC approach. However, results from this study have identified many shortcomings of the approach such as small sample size, missing out important resource use items and inefficient use of data. The approach which has been recommended by the ISPOR taskforce is hierarchical modelling (Manca et al. 2007; Drummond et al. 2009) and the country-specific results for the UK obtained with this approach were similar to the FPMC and FPOC approaches depending on the costing approach that was adopted. This further

strengthens the case against the FSOC approach. It should be pointed out that there are several shortcomings with the bivariate hierarchical models. The most important is the difficulty associated with its implementation and recent reviews have shown that uptake of this approach is low (Vermer et al. 2013; Oppong et al. 2015b). The approach also assumes that there is exchangeability between countries and that all differences between countries are random. In reality however, these differences are often systematic thus the exchangeability assumption would not be plausible (Grieve et al. 2007).

One possible solution would be to split data into similar countries based on factors such as health system, GDP, geographical proximity or some form of test for homogeneity and then apply the bivariate hierarchical models to estimate the country specific cost-effectiveness. Thus the FSMC approach may offer a suitable compromise between the fully pooled approaches and the split approach in cases where there is a lot of uncertainty about pooling or splitting data.

7.8.5 Future research

Even though the case against FSOC approach has been made in this thesis, the fully pooled approaches might not be acceptable to a decision maker who is only interested in results that are applicable to their own setting. A study assessing factors that are considered to be transferable showed that most countries do not consider unit cost and resource use transferable or generalisable to their settings (Barbieri et al. 2010) and as a result there may be an issue with the fully pooled approaches. To fully assess the impact of the various analytical approaches, it is suggested that decision makers should be presented with results from the various approaches and asked to judge the extent to which they would consider the results from each approach in their jurisdiction/country. In

addition to this, the potential for a modelling approach should be considered in order to help solve some of these issues. Current ISPOR guidelines recommend that modelling could be used in cases where a country/jurisdiction did not participate in a trial (Drummond et al. 2009). However, it is believed that modelling could potentially play an important role even in cases where a country participated in a study.

In the area of economic evaluation alongside multinational trials, estimating country-specific cost-effectiveness is one of the areas of contention that remains. Although there are several approaches that have been suggested in the literature, there is yet to be full consensus with respect to the approach that should be used for this purpose. Thus, future research should focus on reaching a consensus with respect to estimating country-specific results. A possible suggestion would be to pool data from countries that are similar and apply the hierarchical models to estimate the country-specific values. Further research needs to be carried out to confirm these approaches.

7.9 Conclusion

The various analytical approaches to economic analysis alongside multinational trials have been considered and the results appear to weigh heavily against the FSOC approach. Therefore it is recommended that considerable caution should be exercised when using this approach. Theoretically, the FPMC approach could be the most plausible approach since it considers data from all countries and applies appropriate unit costs to the data. However, as we have seen, there is the possibility that decision makers in a particular

jurisdiction/country may not accept results from this approach and in such cases; the FSMC approach could offer a suitable compromise after making a judgement about the similarities between the selected countries. This study also showed the importance of other issues such as unit costs resource use and outcomes. To date, most research in this area has focused on developing advanced statistical methods which in most cases are complex and difficult to understand and implement and less emphasis has been on issues such as unit cost, resource use and health outcomes which could equally have an impact on the validity, generalisability and transferability of the results. It is therefore recommended that these issues are considered at the trial design stages.

7.10 Summary

Using data from two randomised controlled trials (Case study 1 and Case study 2) from the GRACE project, the work presented in this chapter compared the pooled and split approaches to economic evaluation alongside multinational trials with respect to resource use, costs, outcomes and cost-effectiveness. The results from both case studies clearly show that the results from economic evaluations alongside multinational trials are sensitive to the methodological approach that is adopted. The overall conclusions and recommendations from the PhD research are presented in the next chapter.

CHAPTER 8 GENERAL DISCUSSION AND CONCLUSION

8.1 Introduction

The purpose of this Chapter is to summarise the findings from the literature review and empirical work presented in this thesis and is structured as follows. Firstly, the aims and objectives of the research are initially stated and the principal findings of the study presented. This is followed by a discussion of the strengths and limitations of the study and a comparison with other studies. The implications for policy and practice and suggestions for future research are discussed and finally, the overall conclusions from the research work are presented.

The general objective of this research was to explore the implications of conducting economic evaluations alongside multinational studies/trials, to document the challenges and to explore various analytical approaches (pooled and split) used. To attain these goals, a literature search was conducted to explore the appropriate approaches and identify the various challenges reported by researchers. Case studies from the Genomics to combat Resistance against Antibiotics in Community-acquired LRTI in Europe (GRACE) project were then used to investigate issues related to the conduct of trial based economic evaluations in multinational settings.

Specific objectives of the research were: (i) To document challenges that have been reported in published economic evaluations alongside multinational trials (ii) To explore

various issues related to estimating costs in multinational studies (iii) To explore the impact of using different tariffs to value EQ-5D health state descriptions in economic evaluation alongside multinational trials using case studies (iv) To compare methods that have been used to conduct economic evaluation alongside multinational trials using case studies and (v) To make recommendations with respect to the conduct of economic evaluation alongside multinational trials. Chapter 2 of this thesis presented a background to economic evaluation in healthcare and outlined issues that are associated with economic evaluation alongside multinational trials. The systematic review of published economic evaluations alongside multinational trials was presented in Chapter 3, and Chapter 4 gave a detailed overview of the case studies that were used to empirically explore the research questions. Costing alongside multinational studies was explored in Chapter 5, while Chapter 6 looked at outcomes in multinational trials with a focus on the choice of EQ-5D-3L value sets. Chapter 7 focused on economic evaluation in general and compared the pooled and split approaches to economic evaluation of multinational trials.

The next section provides a summary of the main findings of the research work in relation to the specific objectives mentioned above.

8.2 Principal findings of the research

Aim 1: Challenges associated with economic evaluation alongside multinational trials

One of the main aims of the systematic review of the literature (Chapter 3) was to identify challenges that have been specifically mentioned by researchers who have

conducted economic evaluations alongside multinational trials. The main challenge identified was related to dealing with the differences (resource use, price, health systems and practice patterns) between countries, and it was particularly noted that there was inadequate guidance to help deal with these differences. Other challenges that were mentioned included: the lack of readily available unit cost data in all participating countries, inadequate sample size in some countries, and how best to estimate country-specific cost-effectiveness. However, not all identified studies (approximately 35%) reported any difficulties that were encountered. From the analysis that was subsequently conducted in this thesis using the case studies, some of the main challenges such as dealing with differences between countries and lack of readily available cost data were also encountered, suggesting that these are some of the areas where additional research is needed.

Aim 2: Costing alongside multinational trials

Exploring the approaches used for costing, with a focus on methods for obtaining unit cost data alongside multinational trials was one of the aims of this research work and to achieve this, a costing exercise using a four stage approach to obtaining unit cost was undertaken in Chapter 5: (i) An internet search was carried out, (ii) UK Health Economists Study Group (HESG) members were contacted by email, (iii) Network coordinators and facilitators in participating countries were contacted and (iv) Health economists and researchers in participating countries were directly contacted. Even though the costing exercise was ultimately successful, there were still instances where unit cost data were unavailable, confirming the difficulty associated with collecting these data, particularly when researchers conducting the economic analysis are based in just

one of the participating countries. It is therefore not surprising that all unit costs associated with resource use from the UK, where the health economists were typically based, were most easily obtained. The study revealed that the most effective approach to collecting unit cost data from countries other than the one in which the researcher was based was by collaborating/direct contact with project partners and researchers/health economists from participating countries. In cases where unit cost data were not available, a number of assumptions, including splitting countries into groups and assuming that they are similar were made in order to estimate unit costs.

The challenge is therefore choosing between good quality unit costs from one country and disrupting the relationship between unit prices and resource use or using less accurate unit costs but maintaining the relationship between price and resource use. However, from the literature, the latter seems to be a more accepted approach (Manca et al. 2010). In terms of methods used to ensure the comparability of costs across countries, the results in Chapter 2 also revealed that the approach used for the conversion of costs into a common currency is important and different conclusions, in terms of the country associated with the highest and lowest cost could be reached if either exchange rates or PPPs are used. Based on the results and the literature, it appears that PPPs would be the most appropriate approach for converting costs into a single currency and ensuring comparability.

Aim 3: Impact of choice of value set to value EQ-5D-3L health states

Chapter 6 explored health outcomes for economic evaluation alongside multinational trials and with a specific focus on the use of the EQ-5D-3L. The study had two main

aims, the first was to assess the construct validity of the EQ-5D-3L in the study population (patients presenting with acute cough/lower respiratory tract infections). The second was to compare EQ-5D-3L scores obtained with the European value set (EVS) with those obtained from country specific value sets (CVS) and the United Kingdom's value set (UKVS) and to explore the impact of between-value-set discrepancies on the estimation of cost-effectiveness in multinational studies.

The results from the validation exercise showed that the EQ-5D-3L is valid for use in patients presenting with acute cough/lower respiratory tract infections since most of the results obtained were in line with the a priori hypotheses that were formulated. With respect to the question about which value set should be used to value EQ-5D-3L health states in multinational trial settings, the results presented in Chapter 6 showed that there were some differences in the EQ-5D-3L scores obtained with different value sets. In some cases, the EVS yielded higher EQ-5D-3L scores than the CVS in some countries (e.g. Wales, England) and lower scores than the CVS in other countries (e.g. Netherlands, Germany). This was attributed to differences between the algorithms used for the various value sets such as the exclusion of the disutility values of dimensions of the EQ-5D-3L e.g. in the German value set. This research also found that the choice of value set could potentially lead to situations where the cost-effectiveness of an intervention is affected. For example, in some countries like Belgium and Finland, the choice led to a change in the decision that would be taken based on the results of the cost-utility analysis. However, in most cases, a similar cost-effectiveness result was obtained irrespective of the value set that was chosen. Testing the choice of value sets within the real trial settings (sensitivity analysis with Case study 1 and Case study 2 in Chapter 7) also showed that,

although different EQ-5D-3L scores were obtained with the UK and European value sets, in most cases, this was not enough to affect the cost-effectiveness of the interventions under consideration. Thus, using either the UK value set or the European value set may not present a problem in multinational trials, although the magnitude of the differences in costs or the cost-effectiveness threshold used could affect the results.

Aim 4: Comparison of analytical approaches to economic evaluation of multinational trials

Chapter 7 of this thesis tested the main analytical approaches to economic evaluation alongside multinational trials that were outlined in Chapter 2 in order to investigate the implications of pooling and splitting data in terms of resource use, costs, health outcomes and cost-effectiveness. Two case studies (Case study 1 and Case study 2) were used for this purpose and the results revealed differences in the cost-effectiveness of the interventions being considered based on the particular analytical approach that was adopted. For example, as shown in Case study 1, the results obtained ranged from a case where the intervention (amoxicillin) was dominant with the fully pooled one country approach to a situation where it was dominated by the control (fully split one country approach). A similar trend was also observed with Case study 2 confirming the view that different analytical approaches may lead to different results. Even for cases where different approaches led to similar results, the cost-effectiveness acceptability curves revealed differences in the probability that the intervention was cost-effective. Assessing resource use, costs and outcomes obtained from the various approaches also revealed vast differences which could potentially lead to inaccurate conclusions about the cost of interventions, resource use and ultimately wrong resource allocation decisions. The only

instance when there was no variation in the outcomes was with the cost-effectiveness analysis in Case study 2 where antibiotic prescribing was the outcome of interest. The results showed that antibiotic prescribing was highest in the usual care group and lowest in the combined intervention group irrespective of the analytical approach adopted. However, when the outcomes are combined with costs to obtain cost-effectiveness ratios and net benefits, the approaches based on pooling and splitting led to contradictory results.

Although from an economic point of view, there are several arguments against pooling data, it has been shown that it is important that the dangers of sub-group analysis and splitting the data are properly highlighted. Smaller confidence intervals (for costs, outcomes and net monetary benefits) were obtained with the pooled approaches when compared to the split approaches. This suggests that from a statistical point of view, pooled approaches may reduce the uncertainty around the results. The FPMC costing approach consistently yielded smaller confidence intervals than the FPOC approach suggesting that this is the approach that yields the least uncertainty. Although obtaining unit costs in every country is difficult, multicountry costing has been shown to be the superior approach (Reed et al. 2005; Manca et al. 2010). As part of sensitivity analysis, a statistical approach (Bayesian hierarchical modelling) to estimating country-specific cost-effectiveness was explored with Case study 1 and the overall results were shown to be sensitive to whether a one country or multicountry costing approach was used. The country-specific results from the UK obtained from the Bayesian hierarchical modelling showed that amoxicillin was more costly and more effective with both the one country and multicountry costing approaches. This was similar to the results that were obtained

with the FPMC approach but very different from those obtained from with the FSOC approach which considered only UK participants (the control was dominant with the FSOC approach). Although, this approach could be quite useful in terms of maximising the available data, it has been noted that some of the assumptions that are made by the model such as the exchangeability assumption which implies that the differences between countries are random have been questioned (Grieve et al. 2007). In practice however, the differences between countries are systematic (Grieve et al. 2007).

8.3 Strengths and limitations of the research

This study has a number of strengths. First, both randomised trials (Case study 1 and Case study 2) and an observational study were used to answer the questions raised in this thesis which enabled testing in real life settings (observational study) and in controlled settings (randomised trials). In addition to this, the two randomised trials had very different designs. Case study 1 was a standard 2 arm trial and Case study 2 had a more complex design i.e. a 2x2 factorial, cluster randomised trial. Second, this study sought to explore economic evaluation alongside multinational trials by looking at the effects pooling and splitting data from multinational trials has not only on the overall cost-effectiveness of interventions, but also on the individual components of resource use, costs and health outcomes. In addition, the study was also able to assess the fully pooled and fully split approaches to economic evaluation of multinational trials. Third, the study provided an updated systematic review of published economic evaluations alongside multinational trials. Lastly, even though the study was limited to European countries, a

total of fourteen countries were included in the case studies, which allowed a good assessment of issues such as costing in multinational trials.

One limitation of the thesis is the use of the 3-Level version of the EQ-5D. A 5-level version of the EQ-5D has now been developed (Herdman et al. 2011) which would have been more appropriate to use. However, at the time the case studies used for this thesis was being conducted, the 5-level version was not readily available. At present, with the exception of the UK (Devlin et al. 2016) and Netherlands (Versteegh et al. 2016), the value sets for the 5-level version of the questionnaire have not been fully developed for most countries, with only an interim cross-walk data algorithm available. Therefore, this implies that even if data on the 5-level version of the questionnaire were collected, it would not be possible to test the impact of various country-specific value sets on the results from cost-effectiveness alongside multinational trials. It is however suggested that future studies should explore the impact of using different tariffs with the 5-level version of the questionnaire in order to determine whether it has an impact on the results of the study. Second, all the participating countries in this study were from the European continent and as a result, issues related to conducting economic evaluation in developing countries as well as countries in other continents were not explored in this study. This has the potential of limiting the generalisability of the findings of this thesis, although most of the issues such as problems of obtaining data and generalisability of findings can be applied to all countries and continents. Third, all the case studies were limited to specific disease area i.e. acute cough/lower respiratory tract infections and most of the interventions were those aimed at combating antibiotic resistance. In addition, the case studies used were carried out over a short-period of time i.e. four weeks and as a result,

issues such as discounting were not explored. Thus, it may seem that some of the findings of this thesis may not be generalisable to other disease areas. However, findings such as those related to costing, choice of value sets and analytical approaches can be easily applied to other disease areas.

8.4 Comparison with other studies

Other studies have considered economic evaluations alongside multinational trials in general or looked at specific issues such as costing or outcomes. This study has looked at costs and outcomes individually and also looked at economic evaluation more generally. In terms of outcomes, one study (Bernert et al. 2009) compared EQ-5D-3L scores that are obtained from different value sets and found results that were similar to those that were obtained in this thesis i.e. different value sets result in different EQ-5D scores. Other studies have also considered the impact of using different value sets for cost-utility analysis using illustrative examples and found mixed results. One study (Sakthong et al. 2008) found that the choice of value set for cost-utility analysis does not matter whilst others (Huang et al. 2007; Johnson et al. 2005) suggest that the choice of value set could make a difference to results. This PhD thesis compared value sets within a proper multinational trial setting which the other studies did not do. In terms of analytical approaches, other studies have considered the implications of pooling and splitting data from multinational trials (Willke et al. 1998; Cook et al. 2003; Reinhold et al. 2010), however, no study has explicitly considered the various analytical approaches (FPMC, FPOC). Challenges associated with multinational trials have been outlined in a number of

studies (Knies et al. 2009; Koopmanscap 2001; Grieve et al 2005; Magnell et al. 2005; Thompson et al. 2006). However, no other study has specifically discussed challenges that have been reported by researchers who have conducted economic evaluations alongside multinational trials.

8.5 Recommendations, implications for policy and future research

This section outlines the various recommendations that have arisen from this piece of work, discusses some implications for policy and suggests some areas where future work can be undertaken.

Aim 1: Challenges associated with economic evaluation alongside multinational trials

The main challenge that was identified in this PhD thesis was related to how to address the differences between countries, which could be attributed to a lack of consensus on many aspects such as how to estimate country-specific cost-effectiveness. Future research should therefore focus on reaching a consensus about how to address these challenges associated with economic analysis alongside multinational trials. This may be achieved by comparing the possible solutions to the challenges and agreeing on the most appropriate solution. Researchers should also be encouraged to report the challenges that they encounter, since this not only serves as a guide to others but may also give an

indication about areas where there is a need for consensus or where additional research is needed.

Aim 2: Costing alongside multinational trials

From the results that have been presented in this thesis, the main problem with respect to costing was related to the identification of unit cost data in all participating countries in a multinational study. To minimise this problem, it is suggested that health economists/researchers make efforts to deal with this difficulty right from the design stages of the study by collaborating with study coordinators and other health economists/researchers in the countries of interest as the work presented in this thesis has shown that this is the most effective method for obtaining unit costs, and most importantly, by getting involved with the recruitment of participating countries for the trial since this may enable them to have a say about which country should be included in the trial. It should be recognised that in most cases, economic questions are not the main motivation for conducting the study and as such, countries that are preferred by health economists due to the availability of unit cost data might not be selected. It is also suggested that analysts endeavour to publish sources of unit cost data that were used in their respective studies so that others would be able to use this data in the future. From the review that was presented in Chapter 3, it was established that the sources of unit costs in many studies were unknown. This is a concern and has implications in terms of transparency of results. One of the aims of the costing study carried out in Chapter 5 was to be very clear about sources of data, such that others are able to locate these sources. However, given the number of assumptions that had to be made in order to apply

country-specific unit costs, there is a need for some additional research in order to solve the problem associated with obtaining unit costs in all participating countries.

With regards to longer term solutions, a recommendation from this work is that greater efforts are made by the health economics community within Europe and around the world to develop a central resource, such as a website, where unit costs are made freely available. This would substantially reduce the burden associated with identifying unit cost data in all countries. In addition, it would help boost transparency of results because researchers would be able to reference sources of cost data rather than relying on strong assumptions to derive the cost data in all centres/countries. A potential starting point for obtaining the data that would be published on this website would be to carry out a study that would compile all published sources of unit cost data across countries. This could be achieved by conducting a systematic review of costing and economic evaluation studies. The information that is obtained from this search can then be published on the website. Also, respective bodies in various countries such as statistical services that may have access to valuable unit cost data should be encouraged to provide this data free of charge. It is generally thought that cost data collection is an under researched area and thus, there is an opportunity for some further work to be carried out here. A potential problem with this suggestion would be obtaining funding to develop such a resource. However, if the problem is highlighted and additional awareness created, funding to support such a project may become available.

Aim 3: Impact of choice of value set to value EQ-5D-3L health states

In terms of using the EQ-5D-3L as an outcome in multinational trials, the current guidance produced by the EuroQol group states that the most appropriate tariff should be used (Szende et al. 2007). However, in a multinational trial setting, this advice may not be sufficient since it would be difficult to determine what the most appropriate tariff is. From the discussions presented in Chapter 6, it is suggested that both the EVS and UKVS should be explored within sensitivity analysis as this would minimise any potential issues with generalisability of the results. A 5-level version of the EQ-5D has recently been developed, and lessons learnt from the issues raised with the 3-level version of the questionnaire, particularly in relation to the development of value sets should not be ignored. First, it is suggested that a similar valuation protocol should be used when developing each value set. This would ensure that there is conformity across all value sets. Recent studies have shown that differences in value sets are not only as a result of the differences in regression coefficients, but also as a result of independent variables that are included in the model (Pullenayegum et al. 2015). There is therefore a need for standardisation of protocols if value sets are to be comparable. Second, to solve the problems associated with the choice of cost per QALY thresholds, it is suggested that a common threshold which can be applied to a number of countries is developed and this threshold would have to be in a common currency for example, Euros, US Dollars or UK Pounds. This would not only help researchers in their reporting of results from multinational trials, but also help local decision maker's judge whether an intervention is cost-effective in their own local setting, particularly in cases where there is no locally accepted threshold. Thirdly, it is suggested that cost-effectiveness analysis which

measures outcomes in terms of natural units should also be used in addition to cost-utility analysis as cost-effectiveness analysis does not depend on the use of tariffs/value sets. However, it must be pointed out that there are no acceptable thresholds for use with outcomes from cost-effectiveness analysis and for this to work out; each outcome would need its own threshold which would not be feasible. Using the example that was presented in this thesis, there is no identifiable threshold at which antibiotic prescribing would be considered cost-effective.

Most importantly, it would be beneficial if the EuroQol group provided specific guidance about the use of value sets in multinational trial settings which would help researchers make the right choice when analysing their data. It is hoped that when the 5-level version of the questionnaire has country-specific value sets developed (rather than the interim cross-walk value sets), most of the problems associated with the use of the 3-level version of the questionnaire for the economic analysis of multinational trials will be resolved through suggestions from research, such as that which has been presented in this thesis.

Aim 4: Comparison of analytical approaches to economic evaluation of multinational trials

With respect to deciding on the right analytical approach i.e. whether to pool or split the data, this research has shown that different approaches lead to different conclusions. Statistically, the pooled approaches, specifically the FPMC approach seems to be the most appropriate since it has the least uncertainty (smallest confidence intervals) around the results. However, this may not be the best approach when the differences between countries and the needs of local decision makers are taken into consideration. Choosing

between the FPOC and FPMC approaches may not be an issue if resource use and outcomes are the most important considerations. However, if costs, net benefits and ICERs are important, then as the results in Chapter 7 shows, differences between the FPOC and FPMC can be expected. Overall, the FPMC approach seems to be the most appropriate because it maintains the relationship between resource use and costs and provides the least uncertainty. Future research should focus on developing an algorithm that would enable researchers choose the most appropriate analytical approach for a particular situation. This could involve running several models simultaneously and selecting the most appropriate, based on a predefined criterion.

The ultimate aim of the results from economic evaluations studies including those from economic evaluations alongside multinational trials is to inform decision makers about whether an intervention represents value for money or not (Drummond et al. 2015) and it is also well established that decision makers are mostly interested in results from their own countries/results that can be applied to their local setting (Manca et al. 2010; Reed, 2012). As a result of this, it may be difficult for decision makers to consider results from pooled approaches e.g. FPOC and FPMC useful. However, the result from the thesis has shown that splitting the data may lead to inaccurate results with respect to resource use, costs and overall cost-effectiveness. One potential solution to the problem is to split countries into similar groups based on factors such as health system, GDP, geographical proximity or some form of test for homogeneity and apply the multilevel approaches for estimating country-specific cost-effectiveness. However, this solution may not be enough since valuable data may be lost as a result of the splitting, but on the other hand, the data may potentially be more applicable to the country of interest. This research also showed

that the multilevel models that are used to derive country-specific cost-effectiveness estimates may be sensitive to the costs that are used i.e. one country versus multicountry costing. This implies that even though this approach utilises the available data, different conclusions may be reached depending on the costs that are applied. As a result of the work carried out in this thesis, some general recommendations and suggestions for future research were identified and they are presented in the section below.

Recommendations for the design conduct and analysis of economic evaluations alongside multinational trials

Based on the results and lessons learned from the thesis, a 10 point checklist was developed to provide recommendations and guidance for the design, conduct and analysis of economic evaluations alongside multinational trials. The items in the checklist are summarized in Table 8.1 below.

Item 1: Participating countries: It is recommended that the names and contribution to sample size of all participating countries in a multinational trial should be stated clearly. This would enable the reader and other potential users judge the validity of the results as well as determine how applicable they are to their particular settings. In addition to this, differences and similarities in the way the health system is structured across participating countries and how it affects choice of comparators, and the way care is delivered should be considered.

Item 2: Study perspective: The reason for adopting a particular perspective for the analysis should be clearly stated. Different perspectives govern decision making across countries. For example, the NHS/PSS perspective is recommended in the UK whilst the

societal perspective is seen as more important in countries like the Netherlands. To increase the generalisability of the findings to a number of countries, it is suggested that different perspectives should always be explored in sensitivity analysis.

Item 3: Resource use: It is important to identify differences and similarities in resource use items between participating countries. This will ensure that all the important resource use items across countries are identified. Focus groups, interviews and other qualitative techniques can be used for this purpose. If resource use across countries varies considerably, it is advisable to focus on collecting items that are frequently used and would have a substantial impact on costs.

Item 4: Unit costs: Sources of unit costs for each participating country should be stated and if the one country costing approach is going to be used, this should be made explicit and justified. If assumptions were made when estimating unit costs, these should be made clear. Publishing the source of unit costs in participating countries would also enable other researchers identify potential sources of unit cost data which can be applied to their respective studies.

Item 5: Outcomes: Outcomes should be chosen in such a way as to ensure that the findings can be made generalisable to all countries participating in the trial. If the EQ-5D is used as an outcome measure, the study should provide a justification about why a particular EQ-5D tariff/value set has been used and the effects of alternative tariffs should be explored within sensitivity analysis.

Item 6: Economic evaluation technique: The economic evaluation technique used (CUA, CEA etc) should be clearly stated and justified. Different countries have specific

guidance with respect to the economic evaluation techniques that should be used. Thus, considering alternative economic evaluation techniques would ensure that the results are more applicable to a wide range of countries. If possible alternative approaches should be considered in sensitivity analysis in order to increase generalisability of the findings. If a single economic evaluation technique is used for the analysis, this should be justified. For example, if the analysis is aimed at informing decisions in a particular country.

Item 7: Method of analysis (pooling and splitting): It is suggested that the method of analysis is stated clearly e.g. fully pooled one country costing, fully pooled multicountry costing etc. Whichever approach is adopted should be justified. If the pooling/splitting is at different levels e.g. pooling only outcome data and splitting resource use, this should also be made very clear.

Item 8: Hierarchical nature of the data should be acknowledged: The hierarchical nature of the data from multinational trials should be acknowledged and the appropriate analytical approach such as multilevel modelling should be used to analyse the data. Between-location variability in cost-effectiveness between countries may result due to the correlation in costs/consequences between participants who are located in particular countries and failing to account for the multilevel structure of the data may lead to misleading findings.

Item 9: Country-specific results: Where appropriate, it is advisable to present country-specific results. If country-specific results are presented, the reasons for and the methods for their estimation should be clearly stated. Decision makers may be interested in results

that are relevant to their own country and may want to know how applicable the results are to their settings.

Item 10: Generalisability of findings: The extent to which the results from the economic analysis of a multinational study can be generalisable and transferable to countries that participated in the trial as well as countries that did not participate in the trial should be discussed. There should be a clear indication of whom and what country the results of the analysis apply to. Discussing how generalisable the findings of a study are may prove useful to potential users of the results.

Table 8.1 Checklist to aid the design and conduct of economic evaluations alongside multinational trials

Item	Recommendation	Justification	Reference
1. Participating countries	(a) The names and contribution to sample size of all participating countries in a multinational trial should be stated clearly (b) Differences and similarities in the way the health system is structured across participating countries and how it affects choice of comparators, and the way care is delivered should be considered	This would enable the reader and other potential users judge the validity of the results as well as determine how applicable they are to their particular settings	Oppong PhD thesis Oppong et al. 2015b Hughes et al. 2016
2. Study perspective	The reason for adopting a particular perspective or multiple perspectives should be clearly stated. If possible, alternative perspectives should be explored within sensitivity analysis.	Different perspectives govern decision making across countries. For example, the NHS/PSS	Oppong PhD thesis Oppong et al.

Item	Recommendation	Justification	Reference
		perspective is recommended in the UK whilst the societal perspective is seen as more important in countries like the Netherlands.	2015b
3. Resource use	Identify differences and similarities in resource use items between participating countries. If resource use across countries varies considerably, focus on collecting items that would have a considerable impact on costs	This will ensure that all the important resource use items across countries are identified.	Hughes et al. 2016
4. Unit costs	Sources of unit costs for each participating country should be stated and if the one country costing approach is going to be used, this should be justified. If assumptions were	Publishing the source of unit costs in participating countries would increase	Bachert et al. 2007

Item	Recommendation	Justification	Reference
	made when estimating unit costs, these should be made clear.	transparency and also enable other researchers identify potential sources of unit cost data which can be applied to their respective studies.	Oppong et al. 2015b Oppong PhD thesis
5. Outcomes	(a) Outcomes should be chosen in such a way as to ensure that the findings can be made generalisable to all countries participating in the trial. (b) If the EQ-5D is used as an outcome measure, the study should provide a justification about why a particular EQ-5D tariff/value set has been used and the effects of alternative tariffs should be explored within sensitivity analysis.	Within multinational trial settings there is little consensus with respect to which EQ-5D tariff is the most appropriate. Comparing alternative tariffs within sensitivity analysis would increase	Oppong PhD thesis Oppong et al. 2013b Bernert et al. 2009

Item	Recommendation	Justification	Reference
		the generalisability of the study findings.	Knies et al. 2009
6. Economic evaluation technique	The economic evaluation technique used (CUA, CEA etc) should be clearly stated. If possible alternative approaches should be considered in sensitivity analysis.	Different countries have specific guidance with respect to the economic evaluation techniques that should be used. Thus, considering alternative economic evaluation techniques would ensure that the results are more applicable to a wide range of countries.	Oppong PhD thesis
7. Method of analysis	It is suggested that the method of analysis is stated clearly	Since there is little	Reed et al. 2005;

Item	Recommendation	Justification	Reference
(pooling and splitting)	e.g. fully pooled one country costing, fully pooled multicountry costing etc. Whichever approach is adopted should be justified. If the pooling/splitting is at different levels e.g. pooling only outcome data and splitting resource use, this should also be made very clear.	consensus with respect to methods and also when to pool or split data, being explicit about the methods used would help improve transparency.	Hughes et al. 2016 Oppong PhD thesis
8. Hierarchical nature of the data	The hierarchical nature of the data should be acknowledged and the appropriate analytical approach such as multilevel modelling should be used to analyse the data.	Between-location variability in cost-effectiveness between countries may result due to the correlation in costs/consequences between participants who are located in particular	Manca, 2005; Drummond et al. 2009 Oppong PhD thesis

Item	Recommendation	Justification	Reference
		countries and failing to account for the multilevel structure of the data may lead to misleading findings.	
9. Country-specific results	Where appropriate, it is advisable to present country-specific results. If country-specific results are presented, the methods for estimating this should be clearly stated	Decision makers may be interested in results that are relevant to their own country and may want to know how applicable the results are to their settings.	Drummond et al. 2009 Oppong PhD thesis
10. Generalisability of findings	The extent to which the results from the economic analysis of a multinational study can be generalisable to countries that participated in the trial as well as countries	Discussing how generalisable the findings of a study are may prove	Drummond et al. 2009 Oppong PhD

Item	Recommendation	Justification	Reference
	that did not participate in the trial should be discussed. There should be a clear indication of who/what country the results of the analysis apply to.	useful to potential users of the results.	thesis

Other general recommendations and suggestions for future research

One particular issue which has not been explored is that of sample size requirements in relation to economic evaluation alongside multinational trials. It has been suggested in the literature that sample size requirements may be different for economic evaluations as opposed to clinical studies (Glick et al. 2007). In previous chapters, it was pointed out that one of the main reasons for conducting multinational trials is to increase sample size. Therefore if sample size requirements for economic evaluation are shown to be smaller than those required for the clinical trial, then economic evaluations can be conducted using a smaller sample from countries that may be considered to be similar. However, it should be noted that initial work seems to suggest that sample size requirements for economic evaluations alongside trials may be larger in some cases due to greater uncertainty in economic evaluations (Glick et al. 2007). Exploring the issue of sample size requirements was beyond the scope of this thesis and it is recommended that future research should look at the issue of sample size in relation to economic evaluations alongside multinational trials, since results from such a study may help with decisions about pooling or splitting the data for economic analysis.

Another potential solution to some of the problems of conducting economic evaluations alongside multinational trials is to consider decision analytic modelling. With this approach, there may be the possibility of potentially adapting the results to fit a particular country's needs. Although the literature currently recommends that decision modelling could be used in cases where the country of interest did not participate in the trial (Drummond et al. 2009), there is the potential to also explore the possibility of modelling to inform decisions on the cost-effectiveness of a health technology in a country that

participated in the trial as well. Although, this was not explored in this thesis, it is recommended that future research should look at the role of modelling in solving some of the issues that are related to the generalisability and transferability of results from economic evaluation alongside multinational trials. In addition to modelling, another suggestion would be to undertake a qualitative study that would actually present the results from the various pooled and split approaches to decision makers to elicit their views and find out which approach would be more acceptable to them. A study of this type could be carried out in numerous countries including those involved in the trial as well as those that were not involved in the trial in order to establish common views of decision makers about the results that are obtained from the various approaches. For example, the results that were obtained from the various pooled and split approaches in this thesis could be presented to decision makers in a number of countries. Although conducting this work may require a substantial amount of funding, it is believed that the results from such an exercise might reveal important factors that decision makers are actually looking out for and may in the long run help address some of the problems that are associated with economic evaluations alongside multinational trials.

It is evident that some of the problems associated with the conduct of multinational trials can be minimised at the trial design stages. However, health economists are normally called in after trials have already been designed and as a result they are not able to influence decisions at early stages (Torgerson et al. 1995). Even when health economists are involved in the design stages, it is still difficult for them to properly influence the design of studies. It is therefore recommended that economists should be involved in making recommendations at the early stages of trials. In particular, they could be

involved in the selection of countries to be included in the trial. However, it is unlikely that economists would have a final say with respect to the countries included in the study. If a transferability index table is developed, countries could be selected based on this index. Similar work has been done in terms of generating a generalisability index for the selection of centres within a single country (Gheorghe et al. 2013) and this work could be extended to selecting countries in multinational settings. A potential problem with this suggestion is that several factors such as willingness to participate, political and economic factors determine whether a country is selected for a trial and such an index may have limited applicability in practice.

The recruitment of patients in the case studies that were used in this thesis was limited to European countries and it is suggested that future research should also explore the implications of conducting an economic evaluation alongside a trial recruiting participants from countries with very different characteristics such as developing versus developed countries to explore further challenges and compare results from studies such as those obtained in this thesis. Data collection and availability is known to be more challenging in developing countries (Knapp et al. 2008) and such a study would help shed more light on some of the issues associated with multinational trials.

From the literature review that was conducted in Chapter 3, most studies did not explicitly state who the results of the analysis was aimed at i.e. whether it was aimed at informing decisions in a particular country or informing decisions across a number of countries. A suggestion would be for researchers/health economists to be clear about their audience. This would not only help them make the right choices in terms of designing studies and analysis of data, but also ensure that the results obtained from multinational

trials become more meaningful and useful. In cases where researchers are seeking to use the results of the multinational trial to inform decisions in just one country, then protocols can be developed to represent practice in a particular country. In cases where this is not possible, then the use of statistical methods such as those suggested by Koopmanschap et al. (2001) to adjust resource use data to match that of the country of interest could be used. Future research should therefore focus on the development of methods that could be used for this purpose. If the aim of the study is to inform cost-effectiveness in a number of countries, then the data collection tools should also be designed in such a ways as to pick out important resource use items in each country in order to make the results more generalisable and transferable to as many countries as possible. If important resource use items are missed, decision makers could easily question the validity of the results. To help develop adequate data collection tools, it is suggested that researchers should identify common resource use items in all countries and also identify those that are specific to countries as well. This may be achieved through focus groups and consultations. Once this is done, the data collection tools can then be pilot tested in participating countries to ensure that resource data collection is feasible.

Earlier in Chapter 2, multilevel modelling was introduced as an approach which should be used in the analysis of multinational trials (Manca et al. 2005) and from the literature review that was conducted in Chapter 3, it was revealed that most studies did not acknowledge the hierarchical nature of the data or use multilevel modelling. Thus, in practice statistical approaches such as multilevel modelling are not often used for the analysis of data. The reasons for this were not really clear, but it is quite possible that most researchers find it difficult to implement these models or may not fully understand

them. Thus, efforts should be made to educate researchers about these methods particularly when it comes to using multilevel modelling.

The case studies used in this thesis assessed interventions in the area of antibiotic resistance and the results obtained may have implications for policies aimed at dealing with antimicrobial resistance in specific countries. The importance of accounting for the cost of resistance in economic evaluation studies assessing interventions in the area of antibiotic resistance has been emphasized in recent studies (Cals et al. 2011; Oppong et al. 2016). However, there are challenges associated with estimating this cost which has led to its exclusion from most economic evaluation studies. It has also been noted that this cost may be too small to warrant inclusion in economic evaluation studies (Coast et al. 1996; Cals et al. 2011; Oppong et al. 2016). Accounting for the cost of resistance in the context of multinational trials may lead to additional challenges since the rates of resistance and the resultant costs would vary by country (CDDEP accessed on 21st October, 2016). This PhD work did not include the costs of resistance due to the difficulty and uncertainty associated with the estimation of this cost. However, research into the inclusion of costs associated with resistance is ongoing (Oppong et al. 2016).

In terms of the specific interventions that were considered in this PhD work, the results could have different implications for policy across counties. Different countries have different recommendations for first choice antibiotic treatments for respiratory tract infections (Mcquiston Haslund et al., 2013; EMA accessed on 21st October, 2016). For example, in countries such as the UK, amoxicillin is used on a much wider scale than other countries where other types of antibiotics are used. Thus, the findings obtained in Case study 1 may be less applicable in countries where amoxicillin is not the main

antibiotic that is prescribed for LRTI. In addition, if an antibiotic is not fully authorized in a country, generalisability of the findings may be hampered. With Case study 2, physicians in countries such as Sweden use point of care CRP routinely and this test has been recommended for use in treating respiratory infections such as pneumonia in the UK and Netherlands (NHG 2013; NICE, 2014). Thus, the results may be less applicable to decision makers in countries where these tests are not used.

8.6 Conclusions

This PhD thesis has assessed the implications of conducting economic evaluations alongside multinational studies by looking at issues related to resource use, costing, outcomes and economic evaluation and has pointed out various issues which are often ignored. Some of the real challenges such as the lack of availability of unit cost data have been highlighted and potential solutions suggested. Specifically, the study has shown that the analytic approach adopted (pooling or splitting) could lead to very different results not only in terms of cost-effectiveness, but also in terms of interpreting resource use, costs and outcomes data, a result which could ultimately lead to the inefficient allocation of scarce health care resources. Several potential solutions to these problems have been highlighted and it is hoped that some of the recommendations here will be adopted. It is also hoped that the results presented in this thesis would help stimulate additional research into the economic evaluation alongside multinational trials in order to improve upon current methods and also ensure that the results from these studies are more applicable to decision making. Although there are still unanswered questions, such as

those relating to estimating country-specific cost-effectiveness, it is also important for the analysist to be aware of the advantages and disadvantages of each of the analytical approaches and be transparent about the particular approach that has been adopted. In addition to this, a 10 point checklist which could be used by all stakeholders involved in the design, conduct and analysis of economic evaluation alongside multinational trials was developed. This thesis has therefore added to knowledge in this area by showing some of the consequences associated with the analytical approach that is adopted by a researcher. Overall, it is believed that the work presented in this PhD thesis has given researchers, policy makers and all stakeholders' additional insight into the economic analysis of multinational studies.

Appendix 1: Data extraction form

The data extraction form used in the systematic review conducted in Chapter 3 is presented in A1.1 below.

A1.1 Data extraction form

Author/Year	
Study aims	
Number of countries included (Country EE was carried out)	
Type of economic analysis	
Health outcomes	
EQ-5D Value set used	
Study perspective	
Analytic approach to economic evaluation used	
Country specific results presented	
Adjustments made to account for variation in country	
Discussed challenges associated with Multinational studies	

Appendix 2: Distribution of countries across studies

A2.1 below presents the distribution of countries in each of the 44 studies that were identified in the systematic review carried out in Chapter 3.

A2.1 Distribution of countries across studies

	Author	Specific region	Classification based on region	Countries Involved	Country where Economic Evaluation was carried out	Classification of studies based on GDP
1	Bachert et al. 2007	Europe	1	Austria, Denmark, Germany, Italy, Netherlands, Spain, Sweden, UK		1
2	Gomes et al. 2010	Europe, Australia, Asia	5	Austria, Australia, Belarus, China, Croatia, Czech Republic, Estonia, Germany, Hungary, Greece, Italy, Latvia, Poland, Portugal, Saudi Arabia, Slovakia, Spain, Sweden, Turkey, UK, Georgia, Israel, Netherlands, Ukraine	UK	3

	Author	Specific region	Classification based on region	Countries Involved	Country where Economic Evaluation was carried out	Classification of studies based on GDP
3	Lamy et al. 2003	North America, Europe, South America	2	North America, Europe, South America (Canada, USA, Argentina, Brazil, Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Mexico, Netherlands, Norway, Spain, Sweden, Switzerland UK)		2
4	Aspelin et al. 2005	Europe	1	Denmark, France, Germany, Spain, Sweden		1
5	Lindgren et al. 2005	Europe, North America, South America	2	Europe, North America, South America (Austria, Australia, Belgium, Canada, Chile, Denmark, France, Germany, Greece, Hungary, Ireland, Italy, Lithuania, Netherlands, Norway, Poland, Portugal, Russia, Serbia, South Africa, Spain, Sweden, Switzerland, Turkey , UK)		2
6	Reed et al. 2004	Europe	1	South America, North America, Europe, New Zealand (Argentina, Austria, Australia, Belgium, Brazil, Canada, Chile, France, Germany, Italy, New		2

	Author	Specific region	Classification based on region	Countries Involved	Country where Economic Evaluation was carried out	Classification of studies based on GDP
				Zealand, Peru, Sweden, Switzerland, UK, USA, Uruguay)		
7	Mittmann et al. 2009	North America and Australia	6	Canada and Australia		1
9	Jowett et al. 2009	Europe, Asia and Australia	5	(Denmark, France, UK, Israel, Slovenia, Portugal, Italy, Austria, Spain, Belgium, Australia, Poland, Germany)	UK	1
10	Fernandez et al. 2005	Europe	1	UK, Germany, France, Spain, Denmark, Finland	UK	1
11	Canonica et al. 2007	Europe	1	(Austria, Denmark, Germany, Italy, Netherlands, Spain, Sweden, UK)		1
12	Buxton et al. 2004	Europe, North America, South America, Asia, Australia, Africa	14	Europe, North America, South America, Asia, Australia, Africa (Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Hungary, Norway, Poland, Sweden, UK, Greece, Israel, Italy, Malta, Spain, Canada, USA, Argentina, Australia, China, Indonesia, Korea, Malaysia, Mexico,	USA	3

	Author	Specific region	Classification based on region	Countries Involved	Country where Economic Evaluation was carried out	Classification of studies based on GDP
				Philippines, Singapore, South Africa, Taiwan, Thailand, Portugal		
13	Briggs et al. 2010	North America, Europe, Asia, Africa	7	North America, Europe, Asia, Africa (USA, China, Hong Kong, Malaysia, Philippines, Singapore, Taiwan, Thailand, Bulgaria, Croatia, Czech Republic, Hungary, Estonia, Latvia, Lithuania, Poland, Romania, Russia, Slovakia, Ukraine, Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Italy, Netherlands, Norway, Spain, Sweden, UK, Australia, New Zealand, Argentina, Brazil, Chile, Mexico, South Africa, Canada		3
14	Briggs et al. 2006	South America, Asia, Africa, Europe North America	11	Argentina, Brazil, Chile, Mexico, Australia, China, South Korea, Malaysia, New Zealand, Philippines, South Africa, Taiwan, Thailand, Austria, Belgium, Croatia, Czech	UK	3

	Author	Specific region	Classification based on region	Countries Involved	Country where Economic Evaluation was carried out	Classification of studies based on GDP
				republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Italy, Latvia, Lithuania, Netherlands, Norway, Poland, Portugal, Romania, Russia, Slovakia, Slovenia, Spain, Sweden, UK, Canada, USA, Puerto rico		
15	Bracco et al. 2007	Europe	1		Sweden	1
16	Martin et al. 2003	Europe	1	Germany, Netherlands, UK, Ireland, Belgium, Luxembourg, Italy, South Africa, France, Greece, Switzerland, Poland, Portugal, Hungary, Czech Republic		2
17	Lofdal et al. 2005	Europe, Asia, Africa South America	8	(Belgium, Brazil, France, China, Greece, Hungary, Malaysia, Norway, Poland, Portugal, South Africa, Sweden, Thailand, UK, Taiwan)		2
18	Willan et al. 2006	Europe and North America	3	(Austria, Belgium, Canada, France, Germany, Italy, the Netherlands, Norway, Portugal, Spain,	Canada	2

	Author	Specific region	Classification based on region	Countries Involved	Country where Economic Evaluation was carried out	Classification of studies based on GDP
				Turkey, and the United Kingdom)		
19	Lorgelly et al. 2010	Europe and Africa	4	(Spain, Denmark, Czech Republic, France, UK, Romania, South Africa, Hungary, Bulgaria, Finland, Norway, Sweden, Portugal, Switzerland, Poland, Russia, Ireland, The Netherlands, Belgium, Slovakia, Germany)	UK	2
20	Knapp et al. 2008	Europe	1	Germany, Italy, Spain, Denmark, France, Greece, Denmark, France, Ireland, Netherlands, Portugal, UK		1
21	Radeva et al. 2005	Europe, North America, South America	2	Europe, North and South America (Canada, Argentina, Austria, Belgium, France, Germany, Switzerland, Denmark, Spain, UK, Italy, Norway, Poland, USA)		2
22	Weintraub et al. 2005	Europe, North America, South America,	14	(Argentina, Australia, Austria, Belgium, Brazil, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece,		2

	Author	Specific region	Classification based on region	Countries Involved	Country where Economic Evaluation was carried out	Classification of studies based on GDP
		Australia Asia, Africa		Hungary, Ireland, Israel, Italy, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, South Africa, Spain, Sweden, Switzerland, United states, United Kingdom		
23	Annemans et al. 2003	Europe	1	Spain, Belgium, UK, Netherlands		1
24	Canoui-Poitrine et al. 2009	Europe, Asia and Australia	5	(Australia, Switzerland, Germany, Spain, Denmark, France, Hungary, Israel, Italy, Netherlands, Poland, Belgium, Ireland, Latvia, Portugal)		2
25	Drummond et al. 2003	Europe, Africa, Asia	13	(Germany, Greece, Israel, South Africa, France, UK, Switzerland, Spain, Belgium, Russia)		2
26	Janzon et al. 2003	Europe	1	Scandinavia		1
27	Manca et al. 2003	Europe	1	UK and Ireland		1
28	Marcoff et al. 2009	North America,	14	(Argentina, Australia, Austria, Belgium, Bulgaria, Belarus,		3

	Author	Specific region	Classification based on region	Countries Involved	Country where Economic Evaluation was carried out	Classification of studies based on GDP
		South America, Europe, Asia, Australia, Africa		Brazil, Canada, Switzerland, Chile, China, Germany, Spain, Estonia, Finland, France, UK, Greece, Hong Kong, Croatia, Hungary, India, Ireland, Israel, Italy, Jordan, South Korea, Lebanon, Lithuania, Latvia, Mexico, Malaysia, Netherlands, Norway, New Zealand, Poland, Portugal, Romania, Russia, Singapore, Slovakia, Sweden, Thailand, Turkey, Ukraine, Uruguay, USA, South Africa)		
29	Price et al. 2002	Europe and North America	3	(UK and USA)		1
30	Sullivan et al. 2003	Europe, North America, South America, Asia, Australia, Africa	14	(Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Hungary, Norway, Poland, Sweden, UK, Greece, Israel, Italy, Malta, Spain, Canada, USA, Argentina, Australia, China, Indonesia, Korea, Malaysia, Mexico, Philippines, Singapore, South		3

	Author	Specific region	Classification based on region	Countries Involved	Country where Economic Evaluation was carried out	Classification of studies based on GDP
				Africa, Taiwan, Thailand, Portugal		
31	Glasziou et al. 2010	Asia, Australia, Europe, North America	9	(Australia, New Zealand, China, Canada, Czech Republic, Estonia, France, Germany, Hungary, India, Ireland, Italy, Lithuania, Malaysia, Netherlands, Philippines, Poland, Slovakia, UK, Russia		3
32	Reed et al. 2004	Australia, Europe, North America, Africa	10	(Australia, Belgium, Czech Republic, Denmark, Finland, France, Germany, Hungary, Italy, Norway, Netherlands, South Africa, Spain, Sweden, UK, USA)		2
33	Rutten van Molken et al. 2007	Europe, Australia, North America and Africa	10	(Poland, France, South Africa, Spain, Hungary, Russia, UK, Canada, Austria, Switzerland, Australia, Netherlands, Italy, Portugal)		2
34	Simon et al. 2006	Europe, North America, South America, Africa, Asia,	14	(Australia, Canada, Denmark, Israel, Italy, Singapore, Netherlands, United Arab Emirates, UK, USA, Albania, Argentina, Brazil, Columbia,		4

	Author	Specific region	Classification based on region	Countries Involved	Country where Economic Evaluation was carried out	Classification of studies based on GDP
		Australia		Cuba, Egypt, Jordan, Malaysia, Mexico, Sri Lanka, South Africa, Thailand, Venezuela, Bangladesh, Ghana, India, Malawi, Nigeria, Pakistan, Sierra Leone, Uganda, Yemen, Zimbabwe)		
35	Wade et al. 2008	Europe and North America	3	(Belgium, Canada, Czech Republic, France, Germany, Italy, Spain, Sweden, UK)		1
36	Dukhovny et al. 2011	North America, Europe, Australia	12	(Canada, Australia, USA, UK, Netherlands, Israel, Germany, Sweden, Switzerland)		1
37	Brown et al. 2003					
38	Edbrooke et al. 2011	Europe and Asia	15	Denmark, France, Israel, Italy, Netherlands, Spain, UK		
39	Gary et al. 2004	Europe and Africa	4	UK, South Africa		
40	Bakhai et al. 2003		No information in both economics and clinical paper			
41	Kolm et al.	Europe, North	14	Argentina, Australia, Austria,		

	Author	Specific region	Classification based on region	Countries Involved	Country where Economic Evaluation was carried out	Classification of studies based on GDP
	2007*	America, South America, Africa, Asia, Australia		Belgium, Brazil, Canada, Denmark, Finland, France, Germany, Hungary, Ireland, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, South Africa, Spain, Sweden, Switzerland, USA, UK		
42	Nasser et al. 2008	Europe	1	UK, Germany, Netherlands, Denmark, Sweden, Spain, Austria, Italy		
43	Welsch et al. 2009	Europe, Asia, South America, Africa, North America, Australia	14	Russia, Poland, Spain, Ukraine, India, Netherlands, UK, Italy, Romania, Bulgaria, Hungary, New Zealand, Chile, Brazil, China, Argentina, Slovakia, South Africa, Australia, Mexico, Korea, Greece, Germany, Canada, Thailand, Belgium, Estonia, Malaysia, Portugal, Austria, Sweden, Switzerland, Belarus, USA, Lithuania, Norway, Latvia, Lebanon, Singapore, Finland, Uruguay, Ireland, Hong Kong, Jordan, France, Croatia, Turkey		

	Author	Specific region	Classification based on region	Countries Involved	Country where Economic Evaluation was carried out	Classification of studies based on GDP
44	Lamy et al. 2004*	Europe, North America, South America, Africa, Asia, Australia	14	Argentina, Australia, Austria, Belgium, Brazil, Canada, Denmark, Finland, France, Germany, Hungary, Ireland, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, South Africa, Spain, Sweden, Switzerland, USA, UK		

*Only mentioned countries that recruited more than 10 participants

Appendix 3: Summary of countries based on World Bank classifications

A3.1 Summary of countries based on World Bank classifications

Country	no of appearances	percentage	Income classification (World bank July 2012)
UK	37	82.22%	High income
Germany	31	68.89%	High income
France	30	66.67%	High income
Spain	30	66.67%	High income
Italy	27	60.00%	High income
Netherlands	25	55.56%	High income
Belgium	24	53.33%	High income
Denmark	24	53.33%	High income
Sweden	22	48.89%	High income
Australia	21	46.67%	High income
Austria	21	46.67%	High income
Canada	20	44.44%	High income
Poland	19	42.22%	High income
Portugal	19	42.22%	High income
Hungary	18	40.00%	High income
South Africa	18	40.00%	Upper middle income
USA	16	35.56%	High income
Norway	15	33.33%	High income
Switzerland	15	33.33%	High income
Finland	14	31.11%	High income
Argentina	13	28.89%	Upper middle income
Greece	13	28.89%	High income
Ireland	13	28.89%	High income
Czech Republic	12	26.67%	High income
Brazil	11	24.44%	Upper middle income
Israel	11	24.44%	High income
Mexico	11	24.44%	Upper middle income
China	9	20.00%	Upper middle income

Country	no of appearances	percentage	Income classification (World bank July 2012)
Malaysia	9	20.00%	Upper middle income
New Zealand	9	20.00%	High income
Russia	9	20.00%	Upper middle income
Thailand	8	17.78%	Upper middle income
Slovakia	7	15.56%	High income
Chile	6	13.33%	Upper middle income
Estonia	6	13.33%	High income
Latvia	6	13.33%	Upper middle income
Lithuania	6	13.33%	Upper middle income
Singapore	6	13.33%	High income
Croatia	5	11.11%	High income
Romania	5	11.11%	Upper middle income
Taiwan	5	11.11%	Upper middle income
Turkey	5	11.11%	Upper middle income
Bulgaria	4	8.89%	Upper middle income
India	4	8.89%	Lower middle income
Turkey	4	8.89%	Upper middle income
Ukraine	4	8.89%	Lower middle income
Belarus	3	6.67%	Upper middle income
Hong Kong	3	6.67%	High income
Jordan	3	6.67%	Upper middle income
Korea	3	6.67%	High income
Philippines	3	6.67%	Lower middle income
Uruguay	3	6.67%	Upper middle income
Iceland	2	4.44%	High income
Indonesia	2	4.44%	Lower middle income
Lebanon	2	4.44%	Upper middle income
Malta	2	4.44%	High income
Slovenia	2	4.44%	High income
South Korea	2	4.44%	High income
Albania	1	2.22%	Lower middle income
Bangladesh	1	2.22%	Low income
Cuba	1	2.22%	Upper middle income
Egypt	1	2.22%	Lower middle income
Georgia	1	2.22%	Lower middle income
Ghana	1	2.22%	Lower middle income
Luxemburg	1	2.22%	High income
Malawi	1	2.22%	Low income
Nigeria	1	2.22%	Lower middle income

Country	no of appearances	percentage	Income classification (World bank July 2012)
Pakistan	1	2.22%	Lower middle income
Peru	1	2.22%	Upper middle income
Puerto rico	1	2.22%	High income
Saudi Arabia	1	2.22%	High income
Serbia	1	2.22%	Upper middle income
Sierra Leone	1	2.22%	Low income
Sri Lanka	1	2.22%	Lower middle income
Uganda	1	2.22%	Low income
United Arab emirates	1	2.22%	High income
Venezuela	1	2.22%	Upper middle income
Yemen	1	2.22%	Lower middle income
Zimbabwe	1	2.22%	Low income

Appendix 4: Transferability checklist

A4.1 Summary of the transferability checklist

HT1	Is the intervention described in sufficient detail?
HT2	Is (are) the comparator(s) described in sufficient details?
SE2	Is (are) the country(ies) in which the economic study took place clearly specified?
P1	Did the authors correctly state which perspective they adopted for the economic analysis?
SP1	Is the target population of the health technology clearly stated by the authors, or when it is not done can it be inferred by reading the article?
SP3	Does the article provide sufficient detail about the study sample(s)?
E5	Have the principal estimates of effectiveness measures been reported?
E6	Are the side-effects or adverse effects addressed in the analysis?
E7	Does the article provide the results of a statistical analysis of the effectiveness results?
B5	Is the level of reporting of benefit data adequate (incremental analysis, statistical analyses)?
C1	Are the cost components/items used in the economic analysis presented?
C5	Are unit prices for resources given?
C6	Are costs and quantities reported separately?
C7	Is the price year given?
C9	Is the currency unit reported?
S1	Are quantitative and/or descriptive analysis conducted to explore variability from place to place?
O1	Did the authors discuss caveats regarding the generalisability of their results?

A4.2 Assessment of studies based on the transferability checklist

Study	Score
Gomes et al. 2010	81.25
Bachert et al 2007	62.5
*Willan et al 2006	81.25
Lofdal et al 2005	68.75
Martin et al 2003	68.75
Bracco et al 2007	56.25
Briggs et al 2006	65.625
Briggs et al 2010	71.875
Buxton et al 2004	71.875
Canonica et al 2007	75
Fernandez et al 2005	81.25
Jowett et al 2009	81.25
Knapp et al 2008	71.875
Lorgelly et al 2010	75
Radeva et al 2005	68.75
Weintraub et al 2005	81.25
Annemans et al 2003	70
Canoui-Piotrine et al 2009	73.33333
Drummond et al 2003	90.625
Glasziou et al 2010	75
Janzon et al 2003	78.125
Manca et al 2003	81.25
Marcoff et al 2009	78.125
Price et al 2002	81.25
Reed et al 2004	71.875
Rutten Von Molken et al 2007	81.25
Simon et al 2006	87.5
Wade et al 2008	90.625
Dukhovny et al 2011	81.25
Lubell et al 2009	87.5
Mittman et al 2009	81.25
Reed et al. 2004	75
Lindgren et al. 2005	81.25
Wollenberg et al. 2008	81.25
Brown et al. 2003	64.20
Aspelin et al 2005	84.375
Lamy et al 2003	78.125

Study	Score
Sullivan et al. 2003	81.25
Edbrooke et al 2011	53.125
Garry et al. 2004	81.25
Bakhai et al. 2003	65.625
Kolm 2007	75
Nasser et al. 2008	65.625
Welsch et al 2009	68.75
Lamy et al 2004	78.125

Appendix 5: List of participating institutions and countries in the GRACE network

Universiteit Antwerpen (Belgium)

Leids Universitair Medisch Centrum (Netherlands)

Universitat Ulm (Germany)

University of Oxford (UK)

Swedish institute for Infectious Disease Control (Sweden)

Cardiff University (UK)

Universitair Medisch Centrum Utrecht (Netherlands)

Hospital Clinic Provincial de Barcelona (Spain)

University of Southampton (UK)

University of Birmingham (UK)

Universita degli Studi di Milano (Italy)

Karolinska Institutet (Sweden)

Medical University of Lodz (Poland)

Diakoniekrankenhaus Rotenburg (Germany)

Rigshospitalet (Denmark)

Appendix 6: Study protocols for the economic analysis alongside the GRACE project

A6.1 Protocol for observational study: Estimating the resource use and costs of treating LRTI alongside the observational cough study in workpackage 8 (Observational study)

Background

This protocol forms one of a series developed for the work in WP11. The document describes the approach that will be used for the economic analysis that will take place alongside the observational cough study in workpackage 8. The study conducted by WP8 involves 14 primary care networks in 12 EU countries. The countries involved operate different health systems making comparisons potentially difficult and complex.

Aims

The aims of the study is to estimate the cost burden of treating LRTI in the countries involved in the study. For estimating cost burden of treating LRTI, specific objectives are:

- To obtain resource use data (resources that go into treatment such as time, medication) and valuation data (the respective values of using resources in

treatment) associated with treating LRTI in the countries that are involved in the GRACE study;

- To calculate a cost per patient treated in each country from the perspectives of both the health budget, and patients and their families.
- Where data are available on the incidence of LRTI, to estimate a total cost burden of LRTI for that country.

Approach

Perspective

Data on resource use will be collected from a societal perspective, comprising resources used by the health system and by patients and their families. In different settings, it should be noted, the relative contribution of each towards total costs will be expected to differ.

Strategy for valuing resources and costs

Given the cross-country nature of the study, there are clear difficulties associated with the comparison of strategies using these data. Issues of pooling cost data that occur in all multi-centre trials are hugely increased where the centres are situated in different countries with vastly different means and methods of financing their health systems. Here, it has been decided that no attempt will be made to determine the most cost-effective strategies at a European level. Instead comparisons of alternative strategies will be carried out at the national level for each of the countries included in WP8. Given this

strategy, for each country, the aim will be to obtain the most accurate and reliable data source. A secondary, rather than primary aim, will be consistency across countries.

Collection of resource use data

Health professionals

The information collected here will include number of visits to the nurse or the doctor and this will be obtained from the patient diary. (Question 3 under general questions part 2). Information on the duration of visits will not be available.

Investigations

Information about investigations will be obtained from the CRF

Medication and prescription costs

Information on the type and volume of medication that patients are given/prescribed by their physician will be collected from the CRF questionnaire (Question 2 under section G management). Information will be categorised to reduce the amount of information that networks have to provide. Information on over-the-counter drugs purchased by patients will be taken from the patient diary (Question 17 under general questions part 1).

Referrals and Procedures to Specialist Services

Information on the number and type of referrals will be obtained from the CRF questionnaire (Question 1 under section G management).

Child/dependant care

Information about whether the patient has dependents is available from the patient diary (Question 4 under general questions part 1), but not information on whether the patient has had to pay for child/dependant care during treatment. Given the nature of this type of illness, the baseline assumption will be that patients do not have to pay for care of dependents during their attendance at consultations. Using information about the number of patients who have dependents, however, it would be possible to adjust this assumption to allow for the cost of caring for dependents in sensitivity analysis.

Days off work and loss of earnings:

Information about time spent off work will be obtained from the CRF questionnaire (Questions 6 and 7).

Additional Information

Information on variation of consultations by age, sex and other socio-economic factors will be obtained from the patient diary (General questions part 1).

Valuation of resource use data

Given the differences in health systems, the source of payment for resource use may be different in different countries. This may result in some countries having extensive costs from a health system perspective and minimal costs from a patient perspective or vice versa. For some aspects of health care, different types of the same good may result in costs being incurred to different groups. For example, medication costs in the UK will fall upon different groups depending upon the type of medication: the costs of an antibiotic to the patient will be the prescription costs whereas the cost of analgesics will generally be directly incurred by the patient as they purchase the drug over-the-counter. In other countries systems will be very different. In the sections below, the broad sources of costs are identified for different types of service. At the analysis stage these costs will be allocated to either the health system perspective or the patient and family perspective.

Health professionals

Wherever possible, the costs associated with physician services will be calculated according to whether the service is provided by a nurse or a doctor and by type of consultation (telephone, home visits, surgery appointment). The source of costs for physicians is given below for each of the 14 primary care networks. Sources will be adapted in the light of further information about more accurate or reliable data.

Cardiff/Southampton

PSSRU Unit Costs of Health and Social Care Netten and Curtis (2005).

Utrecht

Central tariff (CTG) OECD database or ILO database

Barcelona/Mataro

OECD database or ILO database

Rotenburg

OECD database or ILO database

Balatonfured

OECD database or ILO database

Antwerp

Nacional illness and Invalidity Insurance Institute website

OECD database or ILO database

Lodz

Central Office of the National Health Fund

OECD database or ILO database

Milano

OECD database or ILO database

Jonkoping

OECD database or ILO database; Lonestat Excel Sheet (SV: Data)

Tromso

OECD database or ILO database; Statistics Norway (SV forsepel)

Helsinki

OECD database or ILO database

Bratislava

OECD database or ILO database

Investigations

Cardiff/Southampton

Utrecht

Central tariff and registration tariff

Barcelona/Mataro

Rotenburg

Balatonfured

Antwerp

Nacional illness and Invalidity Insurance Institute website

Lodz

Milano

Jonkoping

Tromso

Helsinki

Bratislava

Medication and Prescription costs

For drugs sold over-the-counter, prices will most easily be obtained through contacting retail outlets directly. Where possible, such retail outlets will be identified through the internet and e-mailed or telephoned for prices. Where this is not possible, national networks will be asked to visit retail outlets to obtain information about prices. Costs of prescribed drugs will, where possible be taken from published sources. Country specific sources are detailed below:

Cardiff/Southampton:

British National Formulary (BNF) data: prices from local pharmacies (internet search)

Utrecht

Prices from local pharmacies (internet search)

Barcelona/Mataro

Database of consumer prices available from the Consejo General de Colegios Officiales de Farmaceuticals; Foreign proprietary compendia via The Royal Pharmaceautical Society Library; prices from local pharmacies (internet search)

Rotenburg

The Federal associations of sickness funds would be contacted for information on the costs of drugs; Foreign proprietary compendia via The Royal Pharmaceautical Society Library; prices from local pharmacies (internet search)

Balatonfured

Prices from local pharmacies (internet search)

Antwerp

INAMI/RIZIV would be contacted for information on the costs of drugs in Belgium; prices from local pharmacies (internet search)

Lodz

Prices from local pharmacies (internet search)

Milano

Foreign proprietary compendia via The Royal Pharmaceautical Society Library; prices from local pharmacies (internet search)

Jonkoping

The National cooperation of Swedish pharmacies and the Pharmaceutical benefits board will be contacted for information drug prices; prices from local pharmacies (internet search)

Tromso

The Norwegian Medicines Control Authority will be contacted for information on drug prices and the Norwegian medicines agency price database would be queried; foreign proprietary compendia via The Royal Pharmaceautical Society Library, prices from local pharmacies (internet search)

Helsinki

The prices of drugs would be derived from the Finish Pharmaceutical pricing board, the Ministry of social affairs and health and the National Agency for Medicines; prices from local pharmacies (internet search)

Bratislava

Prices from local pharmacies (internet search)

Price prescription analysis would be used to calculate the average prescription cost for each drug.

Referrals and Procedures to Specialist Services

Costs of referrals and visits to services outside of primary care, including both specialist services and other services such as Walk in Centres in the UK will be obtained from a combination of sources

Cardiff/Southampton

The source of the costs associated with referrals will be NHS reference costs, PSSRU and published literature.

Utrecht

Barcelona/Mataro

Published Literature

Rotenburg

Balatonfured

Antwerp

OECD health data 2006 and statistics Netherlands for other costs associated with specialists.

Lodz

Milano

The costs associated with referrals would be derived from the Italian Health Service SSN

Jonkoping

Tromso

Helsinki

OECD health data 2006

Bratislava

Child/dependant care

The baseline assumption will be that these costs are not incurred.

Days off work and loss of earnings

Loss of earnings will be calculated by using the average salary for the country/region and the occupation of the patient or using the regions deprivation score (Index of Multiple Deprivation) if it is available. Only minimal data are being collected, so although it would be ideal to allow for complexity in the measurement of time of work.

The source of the average salaries/wages in the various networks are as follows:

Cardiff/Southampton

Possible sources of information include the websites www.statistics.gov.uk and www.eurostatistics.gov.uk

Utrecht

The wages would be derived from eurostatistics website www.eurostatistics.gov.uk

Barcelona/Mataro

The wages would be derived from eurostatistics website www.eurostatistics.gov.uk

Rotenburg

The source of information on wages would be the eurostatistics website www.eurostatistics.gov.uk

Balatonfured

The source of information on wages would be the eurostatistics website www.eurostatistics.gov.uk

Antwerp

The source of information on wages would be the eurostatistics website www.eurostatistics.gov.uk

Lodz

The source of information on wages would be the eurostatistics website

www.eurostatistics.gov.uk

Milano

The source of information on wages would be the eurostatistics website

www.eurostatistics.gov.uk

Jonkoping

The source of information on wages would be the eurostatistics website

www.eurostatistics.gov.uk

Tromso

The source of information on wages would be the eurostatistics website

www.eurostatistics.gov.uk

Helsinki

The source of information on wages would be the eurostatistics website

www.eurostatistics.gov.uk

Bratislava

The source of information on wages would be the eurostatistics website

www.eurostatistics.gov.uk

Data analysis

To estimate the cost burden of LRTI for each country, an average cost per patient will be estimated given the data above. Where possible this will be combined with national data on the number of patients attending with this condition, to enable an estimate of the total

cost burden to be made. This cost will be presented in both the country's own currency units and, where appropriate, Euros calculated using Purchasing Power Parities. In presenting these findings, the caveats associated with this approach will be made explicit, in particular, noting that a country whose health system does not have sufficient funding to treat such patients and whose patients cannot afford to purchase care personally, may appear to have a low cost burden but that burden may be felt in other ways (for example loss of productive time among patients who never visit their doctor).

Imputation of data

Given that data are being collected directly from patients, it is likely that there will be missing data for particular items of the questionnaire. The extent to which these data are missing at random will be considered and where data are considered to be missing at random they will be imputed using Stata.

Pooling of data

Attempts to pool data across the different countries are likely to result in problems because differences in both clinical and economic variables across locations will impact on resource use, unit cost and outcome (Manca et al. 2005). One possible option would be to use multi-level modelling, but the use of scenarios rather than attempting to pool data should avoid some of these problems.

Controlling for resistance

In WP8 there will be no information available about whether the patients are suffering from a bacterial infection, or whether any such infection is sensitive to, or resistant to, particular antimicrobials. In some senses this is unproblematic as this is the nature of empirical treatment currently. It means, however, that estimating costs associated with resistant bacteria is not possible as part of WP8.

Discounting

Discounting will not be required given the short time period of the study.

List of Primary Care Networks

Primary Care Network	Country	Currency
Cardiff	UK	Pounds
Southampton	UK	Pounds
Utrecht	Netherlands	Euros
Barcelona	Spain	Euros
Mataro	Spain	Euros
Rotenburg	Germany	Euros
Balatonfured	Hungary	Forint
Antwerp	Belgium	Euro

Lodz	Poland	Zloty
Milano	Italy	Euro
Jonkoping	Sweden	Swedish Krona
Tromso	Norway	Norwegian Krone
Helsinki	Finland	Euro
Bratislava	Slovakia	Slovak Koruna

Health systems for countries involved in WP8

UK

The UK health system is predominantly financed by taxation. All residents in the UK are entitled to free healthcare at the point of delivery. Patients pay for services like dental care and eye tests and they also pay prescription charges. Hospital doctors are paid by salary and family doctors (GPs) through a capitation system.

Sweden

Sweden has a tax-based health care system which covers the entire population. The health system is organised at three levels. The national level, the regional level and the local level. The main mode of paying physicians is through salaries. Patients make out of pocket payments for prescriptions. All residents of Sweden, regardless of nationality are entitled to healthcare at a subsidized rate.

Germany

The Health System In Germany: Germany operates a social insurance system which is compulsory for about 90% of the population. Those earning above the income limit can decide whether to choose private insurance or not. The pharmaceutical sector consists of public pharmacies (supplying over the counter drugs to patients) and hospital pharmacies. The main source of paying physicians is through capitation.

Spain

The Spanish health system is based on the principle of universal coverage and all the residents of Spain regardless of nationality have access to free healthcare. Healthcare has been decentralised and autonomous districts are responsible for providing healthcare to the local population.

Hungary

The Hungarian health system is organised on a county basis and is a mixture of tax and social insurance. The health insurance fund (HIF) as well as budgetary assistance are used to provide healthcare finance.

Belgium

In Belgium, healthcare is publicly funded and privately provided. Health insurance is compulsory and it is mainly funded through employer and employee income contributions. In Belgium, only doctors and other practitioners are allowed to prescribe

drugs. Patients pay the share of reimbursable pharmaceuticals. Physicians are paid fee for service.

Poland

Poland operates a mixed healthcare system (public and private). However, social health insurance is the major source of healthcare finance.

Italy

Italy operates a national health service which provides coverage to all legal residents of the country. Healthcare provision is the responsibility of the national and local government. Healthcare is financed by taxes (National and Regional) as well as copayments by patients. There is also a small private tax system. In Italy, drugs are supplied by the private and public sector pharmacies. Patients make out of pocket payments for drugs. Physicians are paid a combination of salaries and capitation.

Norway

The Norwegian health system is financed mainly through taxation. Patients pay for some of their prescriptions and physicians are paid salaries.

Finland

The Finish health system is mainly tax financed. Both National and municipal governments are responsible for providing healthcare. In Finland, physicians are paid salaries.

Slovakia

Slovakia operates a compulsory social health insurance system. Residents of the country are entitled to free healthcare.

A6.2 Protocol for work alongside Case study 1: Exploring the cost-effectiveness of antibiotics compared with placebo: A randomised placebo controlled double-blind trial

Aims of the economic analysis

The aim of this study is to access the benefits of antibiotic treatment in order to ascertain whether antibiotic treatment should be reduced or not and to provide information to decision makers about the most efficient way of detecting and treating bacterial infections in Europe (in terms of the costs and benefits of treating LRTI). Specific objectives are:

To estimate the cost-effectiveness of antibiotic treatment for acute cough/LRTI in Europe.

To determine which sub-groups of patients benefit from antibiotic therapy.

Description of work

Background

Over the years, economic evaluation has increasingly been carried out alongside clinical trials. With the ever increasing cost of healthcare interventions and drugs, as well as the

scarcity of resources to meet this need, there is the need for economics in decision making. (Lindfors et al., 2007; Drummond 1995; Ramsey et al., 2001). There are also benefits of carrying out economic evaluation alongside clinical trials. Most often, patient specific data on costs and outcomes are readily available in clinical trials. In addition, when economic data is collected alongside clinical trials costs are cut. This is due to the fact that there would not be the need to conduct a stand-alone economic study (Drummond et al., 2005; O'Sullivan et al., 2005).

This protocol forms one of a series developed for the work in WP11. The document describes the approach that will be used for the economic evaluation that will take place alongside the randomised clinical trial (intervention study) in workpackage 10.

Rationale

Acute cough/LRTI (Lower respiratory tract infections) are a class of diseases that account for a high rate of morbidity in Europe and around the world. According to a world health organisation report, 17.4% of all deaths and 13.3% of all Disability Adjusted Life Years (DALYs) are caused by the top five respiratory diseases (WHO 2002; AHRQ, 2003). This class of diseases are normally treated with antibiotics, which has the side effect of resistance development (Coast and Smith, 2001). With the high rates of antibiotic resistance and high cost of antibiotic treatment, there is the need for urgent action. A report for nineteen European countries suggests that the correlation between penicillin use and resistance to the drug is 0.84 (Goossens et al., 2005). A complete eradication of antibiotics would have dire consequences as there could be an outbreak of an epidemic which could lead to several deaths and wipe out entire populations. One alternative that has been recommended is reducing the amount of antibiotics that are prescribed to curb

this problem (Coast et al., 2001; Chung et al., 2007). The question which arises here would be how should these antibiotic prescriptions be reduced?

The costs that are associated with antibiotic use are also very high. In USA, the annual cost of treating resistant infections was about $7billion in 1997 (Cars and Nordberg, 2004). Inappropriate prescribing which potentially leads to antimicrobial resistance also cost the US economy about $18 million annually Ciesla et al., (2004). The cost of resistant infections in the USA and UK has been estimated at $ 6.7 billion and $1.7 billion respectively (*Antibiotics and drugs resistance 2007 accessed 21/08/2007*). From an economic perspective, the costs associated with this problem are important and need to be assessed. Efforts are now being made to reduce the amounts of antibiotics that are now being used with the aim of reducing cost and antibiotic resistance. To this end, GRACE WP 10 has organised a randomised clinical trial to access the benefits of antibiotic use.

Overview of the trial design

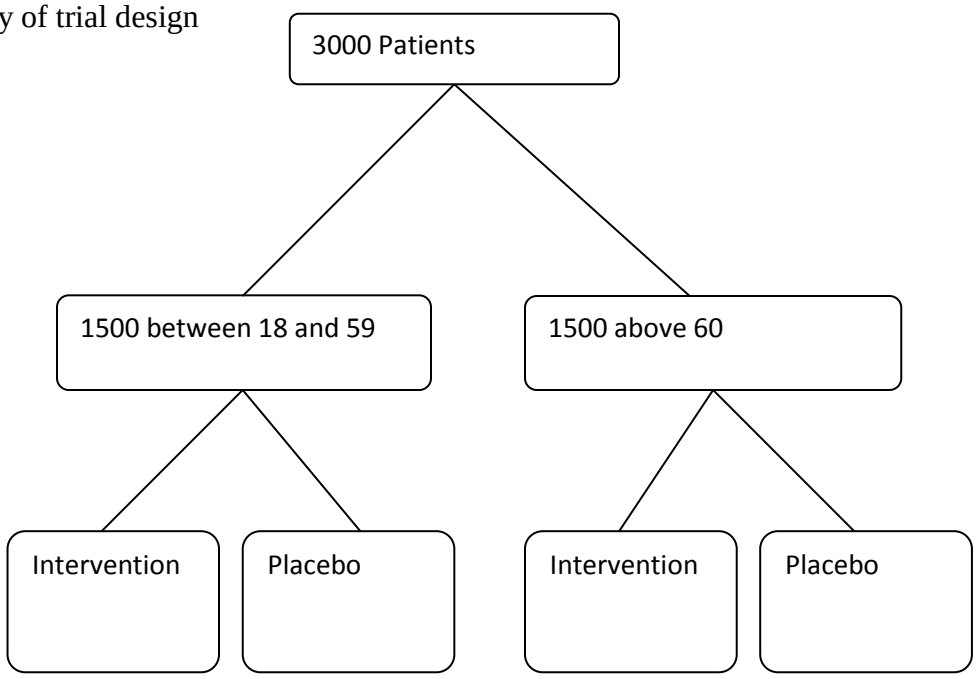
The main aim of workpackage 10 is to determine the effectiveness of antibiotic therapy for the treatment of community acquired lower respiratory tract infections in order to determine which patients actually benefit from antibiotic treatment and which patients do not.

Workpackage 10 consists of two different studies which would be conducted separately. These are: (I) a randomised placebo-controlled double-blind trial with patients as the unit of randomisation comparing antibiotic treatment with placebo in 3000 patients (Case study 1) and (II) a randomised controlled trial with primary care clinicians practises as the unit of randomisation this would be aimed at studying prescribing decisions (Case

study 2). In this study, antibiotic use in patients with LRTI will be compared with usual care. This study protocol concerns the first randomised controlled trial.

Case study 1 would be carried out in 16 primary care networks across 12 countries in Europe. Patients who will be included in the study are adults consulting with acute cough as well as those whose GPs suspect the presence of LRTI. The summary of the trial design is presented in the diagram below.

Summary of trial design



This trial would recruit 3000 patients in 2 age groups consisting of patients aged between 18 and 59 and group 2 would consist of patients aged 60 years and above. Within each group, patients would then be randomised to receive either placebo or the intervention which is amoxicillin. The study will consider two primary outcomes: (1) Deterioration of

illness and (2) Symptom severity and duration. Other outcome measures would be EQ-5D.

Approach

Perspective

Resource use data will be collected from a societal perspective. This comprises of resources used by patients and their families and resources used by the health system. In different settings, it should be noted that the relative contribution of each towards total costs will be expected to differ.

Type of analysis

The economic analysis will be in the form of a cost-utility analysis. QALYs would be estimated from EQ-5D using the European Harmonised Value set in each country.

Data collection

Resource use information will be collected from all 16 primary care networks across the 10 countries taking part in the GRACE study. The main sources of resource use information will be the WP10 CRF questionnaire and the WP 10 patient diary and the WP 10 GRACE symptoms diary. Data on resource use will be collected from within the trial itself and data on costs will be collected from a wide range of sources (published sources, national sources etc).

Valuation of resources and costs

Given the cross-country nature of the study, there are clear difficulties associated with the comparison of strategies using these data. Issues of pooling cost data that occur in all multi-centre trials are hugely increased where the centres are situated in different countries with vastly different means and methods of financing their health systems (Cook et al., 2003).

There have been several approaches that have been adopted to deal with data from multinational economic studies. One approach is to ignore the fact that there would be differences in practices and costs etc. It is also possible to use trial wide clinical data but use price weights from just one country (Glick, 2007).

Another approach is to carry out statistical tests to find out if there is an interaction between country (centre) and treatment effect. If the results on the resource use data are negative, then there is a case for pooling them and attaching prices from various countries. If it turns out to be positive, then there is no case for pooling. Cook et al., (2003) used a similar approach to analyse information from a multinational trial. This approach has been criticized because a negative result suggests that there is no variability in cost-effectiveness by country which is unlikely.

The most recent method is the use of multi-level modelling which accounts for the hierarchical nature of data in multinational trials. This hierarchical nature of data simply means that patients are clustered within centres and centres are clustered within countries. Multilevel modelling allows for the cost-effectiveness across countries to be estimated.

This changes the estimated standard error compared to a method that ignores clustering (Drummond et.,al 2005; Manca et.,al 2005).

Collection of resource use data

Health professionals

The information collected here will include number of visits to the nurse or the doctor and this will be obtained from the WP 10 symptoms diary. Information on the duration of visits will not be available.

Investigations

Information about investigations will be obtained from the GRACE WP 10 CRF questionnaire

Medication and prescription costs

Information on the type and volume of medication that patients are given/prescribed by their physician will be collected from the WP 10 CRF questionnaire. Information on over-the-counter drugs purchased by patients will be taken from the WP 10 symptoms diary.

Referrals and Procedures to Specialist Services

Information on the number and type of referrals will be obtained from the WP 10 CRF questionnaire.

Child/dependant care

Information about whether the patient has dependants is available from the patient diary. Given the nature of this type of illness, the baseline assumption will be that patients do not have to pay for care of dependants during their attendance at consultations. Using information about the number of patients who have dependants, however, it would be possible to adjust this assumption to allow for the cost of caring for dependants in sensitivity analysis.

Days off work and loss of earnings:

Information about time spent off work will be obtained from the WP 10 CRF questionnaire.

Valuation of resource use data

Since health systems differ, the source of payment for resource use may be different in different countries. As a result, some countries may have higher costs from a health system perspective and minimal costs from a patient perspective or vice versa. For some aspects of health care, different types of the same good may result in costs being incurred to different groups. In the sections below, the broad sources of costs are identified for different types of service.

Acquisition of valuation (unit cost) data

Unit cost data for all aspects of resource use will be obtained using a combination of methods. Firstly, a spreadsheet has been developed and circulated to all network co-

ordinators and network facilitators to identify sources of data and, where possible, unit costs. Published material and other sources of costs will be sought. Since the analysis would be done from a societal perspective, both direct and indirect cost will be collected.

Health professionals

Wherever possible, the costs associated with physician services will be calculated according to whether the service is provided by a nurse or a doctor and by type of consultation (telephone, home visits, and surgery appointment). Sources will be adapted in the light of further information about more accurate or reliable data.

Cardiff/Southampton

PSSRU Unit Costs of Health and Social Care Netten and Curtis (2005).

Southampton University Hospital Trust pay scales

Utrecht

Central tariff (CTG) http://www.ctg-zaio.nl/index.php

OECD database or ILO database

Barcelona/Mataro

OECD database or ILO database

Rotenburg

OECD database or ILO database

Balatonfured

OECD database or ILO database

Antwerp

Nacional illness and Invalidity Insurance Institute website

http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf

OECD database or ILO database

http://www.bcfi.be/GGR/PrijsTbl/PTN_IAAACA.cfm#subMPG2019

Lodz

Central Office of the National Health Fund http://www.nfz.gov.pl/

OECD database or ILO database

http://www.proximum.pl/html/cennik.htm

http://www.sanitas.lublin.pl/cennik.php

Milano

OECD database or ILO database

Jonkoping

OECD database or ILO database; Lonestat Excel Sheet (SV: Data)

Tromso

OECD database or ILO database; Statistics Norway (SV forsepel)

Helsinki

OECD database or ILO database; *Hujanen 2003. Guidelines for Health Care Unit Cost in Finland 2001. Stakes Aiheita 1/2003 Helsinki (in Finnish)

Bratislava

OECD database or ILO database

Investigations

Costs of investigations will be elicited from various National and International sources. In addition, attempts will be made to contact the various primary care networks for information on costs of investigations.

Cardiff/Southampton

Southampton University Hospital Trust, Costing Department

Utrecht

Central tariff and registration tariff

Barcelona/Mataro

Source : Cost catalogue 'Hospital Clinic Barcelona' 2004:

Source : Cost catalogue ICS(Catalan Health Institute) 2005:

Rotenburg

Balatonfured

Antwerp

Nacional illness and Invalidity Insurance Institute website

http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf

Lodz

Milano

Jonkoping

Tromso

Helsinki: *Hujanen 2003. Guidelines for Health Care Unit Cost in Finland 2001. Stakes Aiheita 1/2003 Helsinki (in Finnish)

Medication and Prescription costs

Since the trial would be considering amoxicillin this information on the cost of this drug and other prescriptions would be sought from the different countries. For drugs sold over-the-counter, prices will most easily be obtained through contacting retail outlets directly.

Cardiff/Southampton:

British National Formulary (BNF) data: prices from local pharmacies (internet search)

Utrecht

Prices from local pharmacies (internet search)

Barcelona/Mataro

Database of consumer prices available from the Consejo General de Colegios Officiales de Farmaceuticals; Foreign proprietary compendia via The Royal Pharmaceautical Society Library; prices from local pharmacies (internet search)

Rotenburg

The Federal associations of sickness funds would be contacted for information on the costs of drugs; foreign proprietary compendia via The Royal Pharmaceautical Society Library; prices from local pharmacies (internet search)

Balatonfured

Prices from local pharmacies (internet search)

Antwerp

INAMI/RIZIV would be contacted for information on the costs of drugs in Belgium; prices from local pharmacies (internet search)

Lodz

Prices from local pharmacies (internet search)

Milano

Foreign proprietary compendia via The Royal Pharmaceautical Society Library; prices from local pharmacies (internet search)

Jonkoping

The National cooperation of Swedish pharmacies and the Pharmaceutical benefits board will be contacted for information drug prices; prices from local pharmacies (internet search)

Tromso

The Norwegian Medicines Control Authority will be contacted for information on drug prices and the Norwegian medicines agency price database would be queried; foreign proprietary compendia via The Royal Pharmaceautical Society Library, prices from local pharmacies (internet search)

Helsinki

The prices of drugs would be derived from the Finish Pharmaceutical pricing board, the Ministry of social affairs and health and the National Agency for Medicines; prices from local pharmacies (internet search)

Bratislava

Prices from local pharmacies (internet search)

Referrals and Procedures to Specialist Services

Costs of referrals and visits to services outside of primary care, including both specialist services and other services will be obtained from a combination of sources that would be identified.

Cardiff/Southampton

The source of the costs associated with referrals will be NHS reference costs, PSSRU and published literature.

Utrecht

Barcelona/Mataro

Published Literature

Rotenburg

Balatonfured

Antwerp

OECD health data 2006 and statistics Netherlands for other costs associated with specialists.

Lodz

Milano

The costs associated with referrals would be derived from the Italian Health Service SSN

Jonkoping

Tromso

Helsinki

OECD health data 2006

Bratislava

Child/dependant care

The baseline assumption will be that these costs are not incurred.

Days off work and loss of earnings

Loss of earnings will be calculated by using the average salary for the country/region and the occupation of the patient or using the regions deprivation score (Index of Multiple Deprivation) if it is available. Only minimal data are being collected, so although it would be ideal to allow for complexity in the measurement of time of work (see for example, Posnett and Jan (1996) it will not be possible to adjust the data to allow for these complexities.

The source of the average salaries/wages in the various networks are as follows:

Cardiff/Southampton

Possible sources of information include the websites www.statistics.gov.uk and www.eurostatistics.gov.uk

Utrecht

The wages would be derived from eurostatistics website www.eurostatistics.gov.uk

Barcelona/Mataro

The wages would be derived from eurostatistics website www.eurostatistics.gov.uk

Rotenburg

The source of information on wages would be the eurostatistics website www.eurostatistics.gov.uk

Balatonfured

The source of information on wages would be the eurostatistics website www.eurostatistics.gov.uk

Antwerp

The source of information on wages would be the eurostatistics website www.eurostatistics.gov.uk

Lodz

The source of information on wages would be the eurostatistics website www.eurostatistics.gov.uk

Milano

The source of information on wages would be the eurostatistics website www.eurostatistics.gov.uk

Jonkoping

The source of information on wages would be the eurostatistics website www.eurostatistics.gov.uk

Tromso

The source of information on wages would be the eurostatistics website www.eurostatistics.gov.uk

Helsinki

The source of information on wages would be the eurostatistics website www.eurostatistics.gov.uk

Bratislava

The source of information on wages would be the eurostatistics website www.eurostatistics.gov.uk

Outcomes

Two primary outcomes will be considered in this study. These are QALYs derived from the patients' responses to the EQ-5D questions in the patient diary, and the time period between commencement of treatment and full recovery. Further details about each outcome are given below.

QALYs

QALYs will be derived from patients' responses to the five dimensions of the EQ-5D which are asked at weekly intervals up to 4 weeks in the patient diary. EQ-5D is a generic measure of health related quality of life measure. This measure covers the five dimensions of mobility, self-care, usual activity, pain/discomfort and anxiety/depression. Values are anchored at zero (representing death) and one (representing perfect health) and can also be negative (Kind et al., 2007; Mathews and May, 2007). Quality of life values are multiplied by the duration of time spent in each state (in this case 1/52 or one week). It will be assumed that the value of the patient's usual health state is equivalent to their health state at full recovery. The loss in QALYs given their treatment will be assumed to be this value minus the value at earlier time points. Country specific

valuations are available for some countries included in GRACE, but not all. Given the aim of using the best data available for each individual country, values for each country will be selected on the basis of the following order:

The first choice of data will be individual country data based on the time-trade off technique

The second choice of data will be individual country data based on the visual analogue scale technique

The third choice of data will be the European harmonised data based on the visual analogue scale.

Data analysis

A comparison will be made between antibiotics and placebo. This would be done for all the countries that are participating in this study. Cost will be presented in both the country's own currency units and, where appropriate, Euros calculated using Purchasing Power Parities (OECD). Health burden will be presented in terms of lost QALYs. In presenting these findings, the caveats associated with this approach will be made explicit, in particular, noting that a country whose health system does not have sufficient funding to treat such patients and whose patients cannot afford to purchase care personally, may appear to have a low cost and health burden but that burden may be felt in other ways (for example loss of productive time among patients who never visit their doctor).

Sensitivity analysis

The robustness of the results of the clinical trial would also be explored through sensitivity analysis. This will be used to account for particular structural uncertainties and uncertainties associated with the collection of data. Candidates for inclusion in sensitivity analysis will include the potential cost of caring for dependants which will be assumed to be zero in the baseline calculation. Cost-effectiveness acceptability curves would also be used to explore the decision uncertainties that may arise from the analysis.

Imputation of data

Efforts would be made to minimize the problem of missing data. However, given that data are being collected directly from patients, there would almost certainly be missing data to be dealt with. The extent to which these data are missing at random will be considered and where data are considered to be missing at random they will be imputed using Stata (MVIS).

Discounting

Discounting will not be required given the short time period of the study.

Limitations of the Study

For economic evaluation alongside clinical trials, there is the need for a comparator in order to reap the benefits of the study. However, since this trial would compare the intervention (amoxicillin) with placebo, problems may arise when economic evaluation is carried out alongside it. One major problem that may arise is that the incremental impact of the therapy would not be adequately captured. Although this is not potentially harmful

to the clinical study it is not appropriate for the economics because a comparison of the therapy with current practise is what is needed for the economics (Drummond et al., 2005; Lindfors, 2007).

Initially, data on costs and resource use for this study would not be pooled. However, if it becomes necessary, we would pool data. But this would lead to other problems. Most often it has been argued that when data from various countries or centres are pooled, the results do not truly represent the results that would have been obtained in individual countries if the analysis were done separately. This has led to the problem of lack of transferability and generalisation of the results from such trials. This problem is caused by differences in practise patterns in various countries; differences in unit costs between the different countries.

A6.3 Protocol for work alongside Case study 2: Exploring the cost-effectiveness of a web based intervention aimed at reducing antibiotic prescribing

BACKGROUND

The document describes the approach that will be used for the economic evaluation that will take place alongside the randomised clinical trial (intervention study) in workpackage 10b.

Aims of the economic analysis

The main aim of the economic analysis is to determine whether a web based behavioural intervention aimed at reducing antibiotic prescribing is cost-effective. This would ultimately provide information to decision makers about the most effective and cost-effective way of reducing antibiotic prescribing in patients with acute cough/lower respiratory tract infections (LTRI).

Behavioural approaches to reduction in antibiotic prescribing

In recent years there has been a focus on the use of behavioural interventions to influence clinician behaviour such as clinician prescribing behaviour (Simpson et al., 2009). This is due to the fact that other interventions such as verbal persuasion have not been as

effective in dealing with the problem (Martens et al 2006). Results from a few studies have shown that training GPs in a standardised way can provide a reduction in antibiotic prescribing (Cals et al 2009). In addition, another study showed that a web based training programme for GPs can provide a reduction in antibiotic prescribing similar to the standardized methods of training. However, there is the need for the findings from these studies to be replicated in a wider setting.

CRP Near patient testing

The use of rapid point of care tests has been promoted to improve targeting of antibiotics only to those patients who will benefit. A rapid test for C-reactive protein (C-RP) is widely used in some parts of Europe such as in Nordic countries to guide antibiotic management for acute cough/LRTI (Cals et al., 2009). The advantage of near patient testing is that the results are available in minutes (Hansson et al., 1995). However, evidence is mixed about the diagnostic value of C-RP in distinguishing bacterial from viral infection (aetiology) as well as indicating outcome and likelihood of benefit from antibiotic treatment (prognosis). Some studies have shown that C-RP has value as a diagnostic tool and has sufficient sensitivity and specificity in determining whether an infection is bacterial or viral (Dahler-Eriksen et al., 1999). While others conclude that it should not be recommended as a guide for antibiotic therapy (Diederichsen et al 2000). A recent systematic review concluded that the diagnostic value of C-RP testing has not been rigorously studied (Van der Meer et al 2005). Hence, there is the need for more evidence to determine whether C-RP is an effective diagnostic and prognostic tool in primary care.

Overview of the trial design

The main aim of this trial is to determine the effectiveness of a behavioural intervention on Clinician prescribing behaviour. This trial would take place in eight primary care networks located in six different European countries. The countries that are included in this study are: UK (England, and Wales), Netherlands, Belgium, Spain and Poland.

In this study, practices would be the unit of randomisation. The practices selected for participation would be randomised into the following groups:

(1) Normal care

(2) Web based training in communication skills

In each of the groups above, (normal care and web based training) half of the practices would be randomised to receive web based training in the use of CRP (see figure 1).

FIGURE 1: Summary of the trial design

NORMAL CARE	WEB BASED TRAINING
NORMAL CARE + CRP	WEB BASED TRAINING + CRP

Perspective

Resource use data will be collected from a societal perspective. This comprises of resources used by patients and their families and resources used by the health system. In different settings, it should be noted that the relative contribution of each towards total costs will be expected to differ.

Type of analysis

The economic analysis will be in the form of a cost-effectiveness analysis and cost-utility analysis.

DATA COLLECTION

In all the eight primary care networks across the six countries, resource use data would be collected from within the trial itself. The main source of this data would be the WP10b CRF questionnaire, and the WP10b patient diary. However, data on costs would not be limited to the trial itself but these data would come from a number of sources including internet searches, published material and personal communication.

Collection of resource use data

Resource use data would be collected under a number of headings: these include health professionals, medical investigations, referrals and productivity losses.

Health professionals

The main source of this information would be the WP10b symptom diary. This would contain information on the number of visits to the nurse or the doctor.

Medical Investigations

Information about investigations will be obtained from the GRACE WP10b CRF questionnaire.

Medication and prescription costs

Information on the type and volume of medication that prescribed to patients by their physician will be collected from the WP 10b CRF questionnaire. The WP10b symptoms diary would also provide information on over-the-counter drugs purchased by patients.

Referrals to specialist services

Information on the number and type of referrals will be obtained from the WP10b CRF questionnaire.

Days off work and lost productivity

Information about time spent off work will be obtained from the WP 10 CRF questionnaire.

Acquisition of unit cost data

For all aspects of resource use data, information on unit costs would be collected using a number of different methods and approaches. Firstly, published sources of unit cost data would be used. Other methods that would be employed include: (a) Developing a spreadsheet which would be sent to all national network coordinators and national network facilitators in all participating countries. This spreadsheet would seek to elicit information on sources of cost data in the individual countries that have been identified. (b) Contacting researchers/health economists working in the countries participating in the study directly and (c) Internet searches to identify National and International databases with information on unit costs. Unit cost data would be collected under the following headings:

Health Professionals

Wherever possible, the costs associated with physician services will be calculated according to whether the service is provided by a nurse or a doctor and by type and length of consultation (telephone, home visits, and surgery appointment). Sources will be adapted in the light of further information about more accurate or reliable data.

Medical Investigations

The costs associated with medical investigations such as chest X-rays will be elicited from various National and International sources. Where these data are not available, attempts will be made to contact the various national network coordinators and national network facilitators in the specific primary care networks for information on costs of the medical investigations.

Medication and prescription costs

Information on drug costs would be obtained from a number of sources. The main source of this data would be from national databases. Other sources would include published sources. For drugs sold over-the-counter, prices will be obtained from a published study (Oppong et al 2011).

Referrals and procedures to specialists

Costs of referrals and visits to services outside of primary care, including both specialist services and other services will be obtained from a combination of sources that would be identified. This would include published sources and direct contact with researchers.

Days off work and loss of productivity

We would seek to estimate productivity losses. This would be done by considering the average wages in each of the countries and multiplying this by the number of days off work. The average wages would be used due to the lack of detailed information on occupation.

OUTCOMES

The main outcome from this study would be antibiotic prescribing. Secondary outcomes would include Quality adjusted life years (QALYs) derived from patients response to the EQ-5D questionnaire.

QALYs

Quality adjusted life years (QALYs) will be derived from patients' responses to the five dimensions of the EQ-5D which are asked at weekly intervals up to 4 weeks in the patient diary. EQ-5D is one of the most popular generic measures of health related quality of life measure. The EQ-5D covers the five dimensions of mobility, self-care, usual activity, pain/discomfort and anxiety/depression. Values are anchored at zero (representing death) and one (representing perfect health) and can also be negative (Kind et al., 2007; Mathews and May, 2007). Quality of life values are multiplied by the duration of time spent in each state (in this case 1/52 or one week). QALY gains in the treatment and control arms would be compared. Country specific valuations are available for some countries included in GRACE, but not all. Given the aim of using the best data available for each individual country, and being able to generalize across all participating countries, the European harmonised value set would be used to value EQ-5D data in each country.

DATA ANALYSIS

In each of the four trial arms in the figure above, costs (Euros) and outcomes in terms of reduction in antibiotic prescribing would be estimated. In all countries, a comparison in terms of costs and outcomes (antibiotic prescribing and QALYs) would be made between each of the trial arms (normal care vs GP training; normal care vs GP training + crp; normal care vs normal care+crp; normal care+crp vs gp training+crp; normal care+crp vs gp training; gp training vs gp training+crp). In order to determine the incremental impact of the interventions, a cost per unit reduction in antibiotic prescribing and a cost per QALY gained would be estimated in order to determine the most cost-effective option. Costs will be presented in both the country's own currency units and, where appropriate, Euros calculated using Purchasing Power Parities (OECD).

SENSITIVITY ANALYSIS

Sensitivity analysis would focus on exploring the effects of uncertainty in unit cost and resource use focusing particularly on issues such as implementation costs associated with the intervention.

LIMITATIONS OF THE STUDY

The major limitation of this study would be the fact that it would be problematic to adapt findings of multinational studies to individual countries. A number of reasons including Differences in health care systems and differences in demography and patient characteristics can be cited for this (Reinhold et al 2010).

Appendix 7: Approaches to obtaining unit cost in participating countries

A7.1 Example of letter sent out to UK HESG members

Hi all,

We have become involved in an EU study looking at costs of respiratory infection in a number of countries in the EU. We are particularly interested in primary care costs, medication costs and hospitalisation costs. I wondered if anyone in HESG would be able to point us towards good national sources of cost data in any of the following EU countries: Belgium, Finland, Germany, Hungary, Italy, Norway, Poland, Slovakia, Spain, Sweden

Any help would be much appreciated.

A7.2 Example of cost questionnaire sent out to National Network Facilitators and National Network Coordinators

SOURCES OF COST QUESTIONNAIRE

Country

Network

Instructions: Please fill in the yellow boxes below as far as possible. Please give us as much information as possible about how we can access this information. For example, including contact name, address, telephone numbers, emails and/or website address. If you know the cost of any items from these national sources of data, please can you give us the relevant figures, including the currency in which they are given, and the year to which these figures apply.

Section A: This section is about the cost of running the clinic

General Practitioner

	National data available					
	Yes	No	National source of cost	Cost	Currency	Year of cost
Cost per consultation						

National data available

Nurses		Yes	No	National Source of cost	Cost	Currency	Year of cost
	Cost per consultation						

National data available

Non clinical staff		Yes	No	National source of cost	Cost	Currency	Year of cost
	Administrators per consultation						
	Clerical staff per consultation						
	Administrators annual salary						
	Clerical staff annual salary						

Other Staff please indicate the title of staff

National data available

Title	Yes	No	National source of cost	Cost	Currency	Year of cost

If available, please indicate the cost per consultation of the following items

National data available

Yes	No	National source of cost	Cost	Currency	Year of cost

Rent										

Other overheads please specify what is included

Item	National source of cost

Section B: This section is about the cost of performing investigations

	National data available					
	Yes	No	National source of cost	Cost	Currency	Year of cost
Full blood count						

C-reactive protein						
Erythrocyte sedimentation rate						
Procalcitonin						
Urea or creatinine						
Electrolysis						
Chest X-ray						
Blood for serology						
Sputum for serology						
Nose and throat swabs						
Spirometry						
Ausculation:bronchophony						
B-glucose						
BGA						
Cold agglutination (neg)						
ECG						
EKG						
Cukor						
Flu jab						
Hb (heamoglobin)						
INR						
Mycoplasma test						
Spirometry						
Strep test						
Urinary specimen test						
Throat X-ray						

Section C: In this section please indicate other national sources of cost that you are aware of

Item	National source of cost

(I) Example of response to cost questionnaire from Belgium

SOURCES OF COST QUESTIONNAIRE

Country	Belgium
Network	Antwerp

Instructions: Please fill in the yellow boxes below as far as possible. Please give us as much information as possible about how we can access this information. For example, including contact name, address, telephone numbers, emails and/or website address. If you know the cost of any items from these national sources of data, please can you give us the relevant figures, including the currency in which they are given, and the year to which these figures apply.

Most costs and reimbursement rates can be found at this page on the website of the National Illness and Invalidity Insurance Institute: http://www.inami.fgov.be/insurer/nl/rate/index.htm
Some of the costs are (partly) reimbursed to the patients by the National Illness and Invalidity Insurance Institute (RIZIV).
All costs pertain to ambulatory care (cost can be different for people in hospital).

Section A: This section is about the cost of running the clinic

General Practitioner

	National data available							
	Yes	No	National source of cost		Cost		Currency	Year of cost
Cost per consultation	√		http://www.inami.fgov.be/insurer/nl/rate/pdf/last/doctors/raad20070201nl.pdf		18,10 for a certified GP (20,79 for a certified and accredited GP)		€	2007

Nurses		National data available							
		Yes	No		National Source of cost		Cost	Currency	Year of cost
	Cost per consultation		*						

Non clinical staff		National data available						
		Yes	No	National source of cost		Cost	Currency	Year of cost
	Administrators per consultation		*					
	Clerical staff per consultation		*					
	Administrators annual salary		*					
	Clerical staff annual salary		*					

Other Staff please indicate the title of staff		National data available					
	Yes	No	National source of cost		Cost	Currency	Year of cost
Title		*					

* In general Flemish GPs have no support from nurses or other staff.

If available, please indicate the cost per consultation of the following items

	National data available						
	Yes	No	National source of cost	Cost	Currency	Year of cost	
Rent		*					

* Not sure what you mean to cost here? If you mean rent of equipment and rooms. In general Flemish GPs own their own equipped practice.

Other overheads please specify what is included

Item	National source of cost
*	

* Not sure what you mean to cost here? An example might help.

Section B: This section is about the cost of performing investigations

	National data available					
	Yes	No	National source of cost	Cost	Currency	Year of cost
Full blood count	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf	7.72	€	2007
C-reactive protein	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf	3.64	€	2007
Erythrocyte sedimentation rate	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf	1.16	€	2007
Procalcitonin	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf	17.46	€	2007
Urea or creatinine	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf	1,45 OR 2,70	€	2007
Electrolysis					€	
Chest X-ray	√		2)	13,79*	€	2007

Blood for serology	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf	0	€	2007
Sputum for serology	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf	8.73	€	2007
Nose and throat swabs	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf	7,72 OR 40,73	€	2007
Spirometry	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/arts520070601nl.pdf	21.19	€	2007
Ausculation: bronchophony				0	€	2007
B-glucose	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf	1.45	€	2007
BGA					€	
Cold agglutination (neg)	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf	1.16	€	2007
ECG	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/arts520070601nl.pdf	16.09	€	2007
EKG	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/arts520070601nl.pdf	16.09	€	2007

Cukor						€	
Flu jab	√		http://www.bcfi.be/GGR/PrijsTbl/PTN_IAAACA.cfm#subMPG2019	http://www.bcfi.be/GGR/PrijsTbl/PTN_IAAACA.cfm#subMPG2019	10,38 - 10,84	€	2007
Hb (heamoglobin)	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf		1.16	€	2007
INR	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf		1.75	€	2007
Mycoplasma test	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf		7,27 OR 8,73	€	2007
Spirometry	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/arts520070601nl.pdf		21.19	€	2007
Strep test	√		1 + http://www.riziv.fgov.be/insurer/nl/rate/pdf/last/doctors/bio20070101nl.pdf		2.33	€	2007
Urinary specimen test						€	
Throat X-ray	√		2)		16,09*	€	2007

1) https://www.riziv.fgov.be/webapp/nomen/Search.aspx?lg=N
2) http://www.inami.fgov.be/care/nl/nomenclature/pdf/art17.pdf and http://www.inami.fgov.be/insurer/nl/rate/pdf/last/doctors/rx20070101nl.pdf
* minimum of two takes

Section C: In this section please indicate other national sources of cost that you are aware of

Item	National source of cost
Amoxicilline 1 g	http://www.bcfi.be/GGR/PrijsTbl/PTN_HAAACB.cfm#subMPG1758

(II) Example of response to cost questionnaire from Poland

SOURCES OF COST QUESTIONNAIRE

Country	Poland
Network	Lodz

Instructions: Please fill in the yellow boxes below as far as possible. Please give us as much information as possible about how we can access this information. For example, including contact name, address, telephone numbers, emails and/or website address. If you know the cost of any items from these national sources of data, please can you give us the relevant figures, including the currency in which they are given, and the year to which these figures apply.

Section A: This section is about the cost of running the clinic

General Practitioner

(annual capitation payment)

National data available

	Yes	No	National source of cost	Cost	Currency	Year of cost
Cost per consultation *						
* patient from out of the territory of local branch of NHF	X		NHF**	20	PLN	2007
* foreign patient from EU	X		NHF**	45	PLN	2007

Nurses

(annual capitation payment)

National data available

	Yes	No	National Source of cost	Cost	Currency	Year of cost
* Cost per consultation						

*	from out of the territory of local branch of NHF	X			NHF**		7		PLN		2007
*	from EU	X			NHF**		12		PLN		2007

Non clinical staff

	National data available					
	Yes	No	National source of cost	Cost	Currency	Year of cost
Administrators per consultation		X				
Clerical staff per consultation		X				
Administrators annual salary		X				
Clerical staff annual salary		X				

Other Staff please indicate the title of staff

	National data available					
Title	Yes	No	National source of cost	Cost	Currency	Year of cost
		X				

If available, please indicate the cost per consultation of the following items

	National data available					
	Yes	No	National source of cost	Cost	Currency	Year of cost
Rent		X				

Other overheads please specify what is included

Item	National source of cost

Section B: This section is about the cost of performing investigations*

* only individual contracts with labs

	National data available					
	Yes	No	National source of cost	Cost	Currency	Year of cost
Full blood count		X				
C-reactive protein		X				
Erythrocyte sedimentation rate		X				
Procalcitonin		X				
Urea or creatinine		X				
Electrolysis		X				
Chest X-ray		X				
Blood for serology		X				
Sputum for serology		X				
Nose and throat swabs		X				
Spirometry		X				

Ausculation:bronchophony		X				
B-glucose		X				
BGA		X				
Cold agglutination (neg)		X				
ECG		X				
EKG		X				
Cukor		X				
Flu jab		X				
Hb (heamoglobin)		X				
INR		X				
Mycoplasma test		X				
Spirometry		X				
Strep test		X				
Urinary specimen test		X				
Throat X-ray		X				

Section C: In this section please indicate other national sources of cost that you are aware of

Item	National source of cost

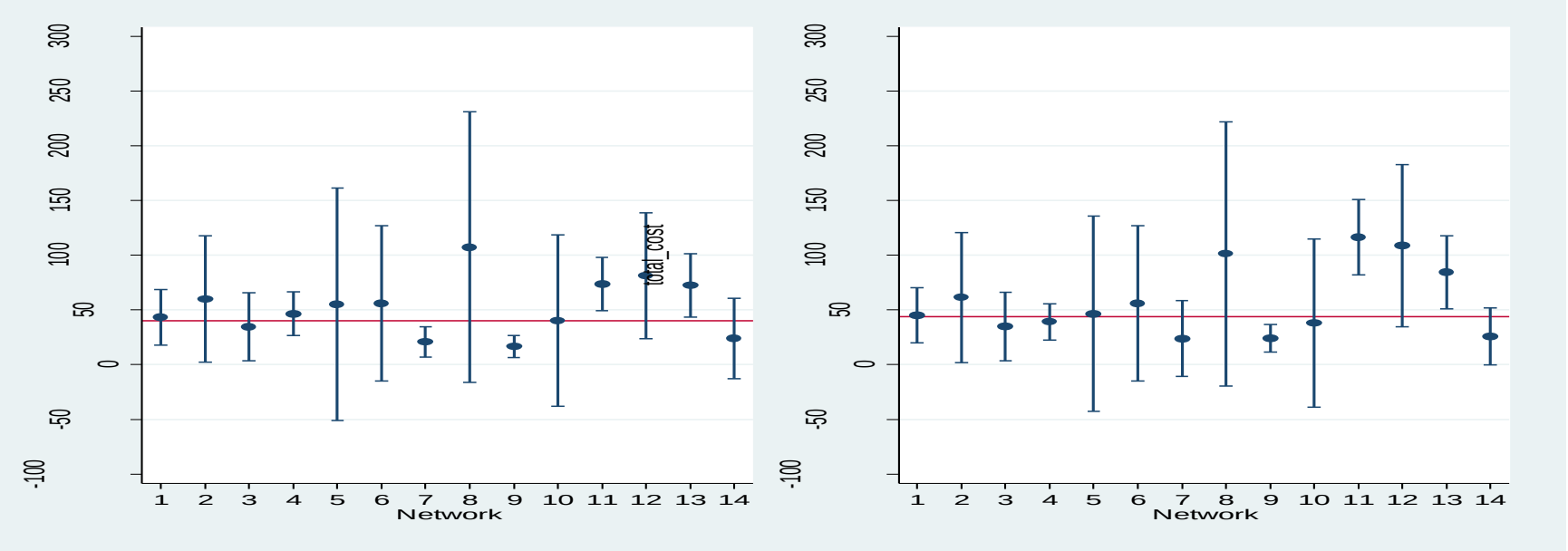
Appendix 8: Conversion rates derived from the market basket approach

A8.1 Conversion rates used in the market basket approach

Country	Conversion index
Belgium	1.04
France	0.97
Germany	1.07
Italy	0.82
Netherlands	1.15
Poland	0.57
Slovakia	0.66
Slovenia	0.78
Spain	0.83
Sweden	1.14
UK	1

Appendix 9: Comparison of mean costs obtained from nominal exchange rates and purchasing power parities

A9.1 Mean costs from exchange rates and PPPs



(a) PPP exchange rates

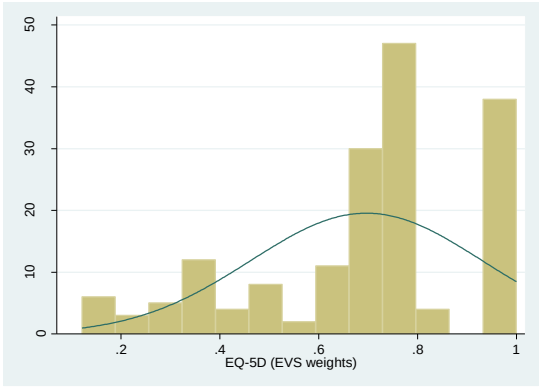
(b) Exchange rates

1: Cardiff 2: Southampton 3: Utrecht 4: Barcelona 5: Mataro 6: Rotenburg 7: Balatonfured 8: Antwerp 9: Lodz 10: Milan 11: Jonkoping 12: Tromso 13: Helsinki 14: Bratislava

Appendix 10: Distribution of EQ-5D scores across countries with different value sets

Figure A10.1: DISTRIBUTION OF EQ-5D SCORES FROM CVS AND EVS (WALES)

(a)



(b)

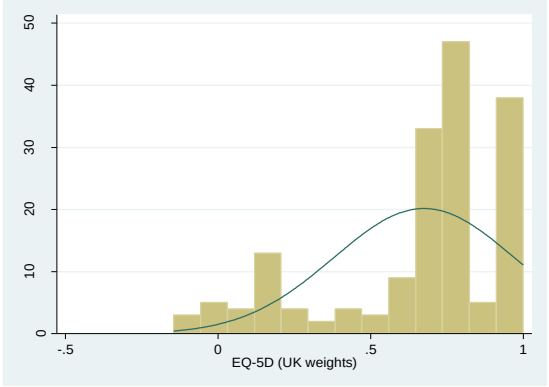


Figure A10.2: DISTRIBUTION OF EQ-5D SCORES FROM CVS AND EVS (ENGLAND)

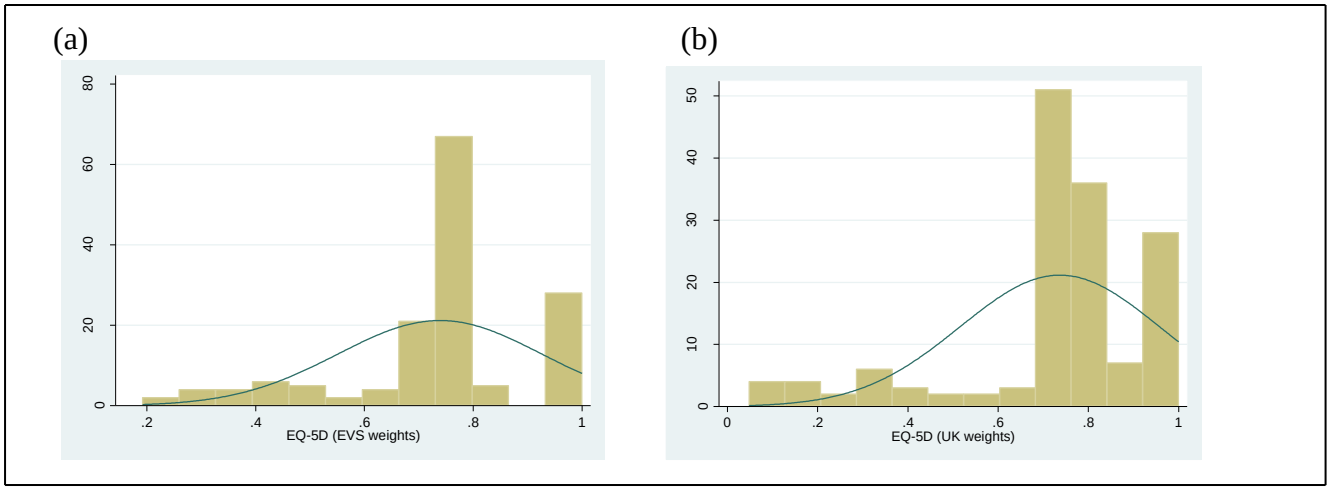


Figure A10.3: DISTRIBUTION OF EQ-5D SCORES FROM CVS EVS AND UKVS (NETHERLANDS)

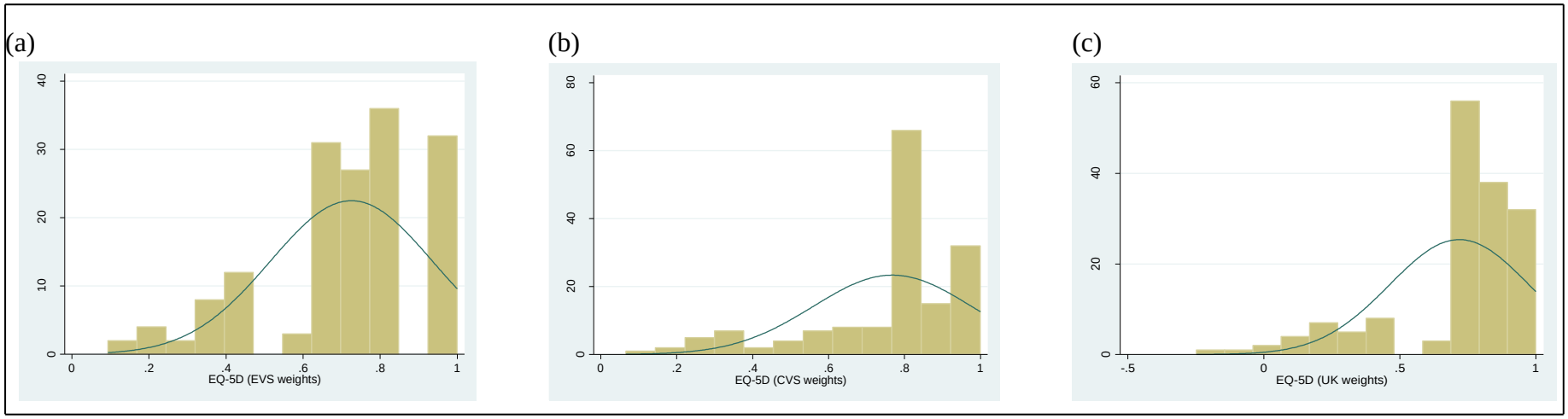


Figure A10.4: DISTRIBUTION OF EQ-5D SCORES FROM CVS EVS AND UKVS (SPAIN)

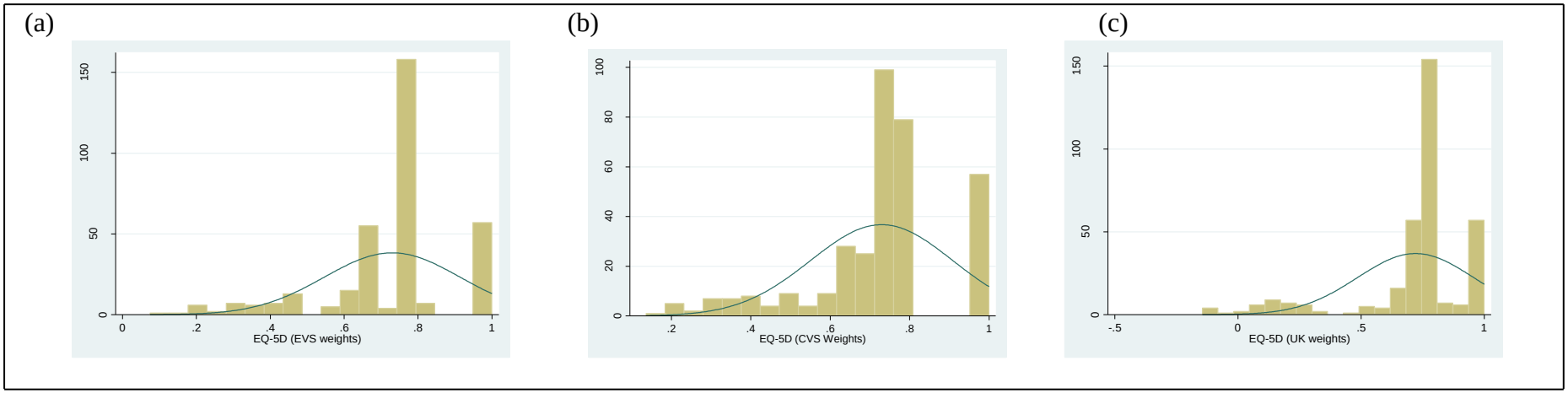


Figure A10.5: DISTRIBUTION OF EQ-5D SCORES FROM CVS EVS AND UKVS (GERMANY)

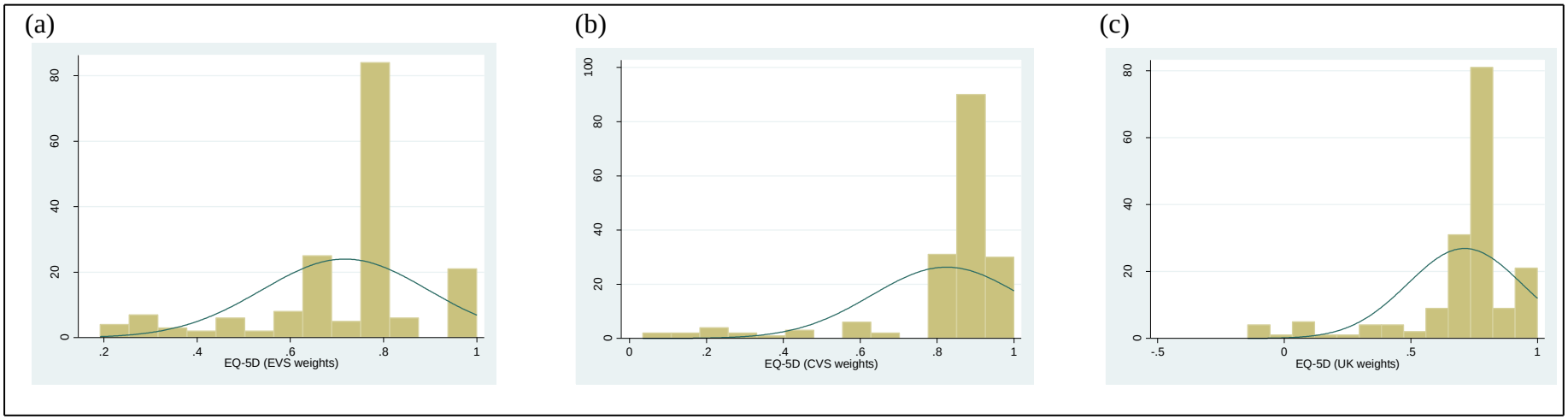


Figure A10.6: DISTRIBUTION OF EQ-5D SCORES FROM CVS EVS AND UKVS (BELGIUM)

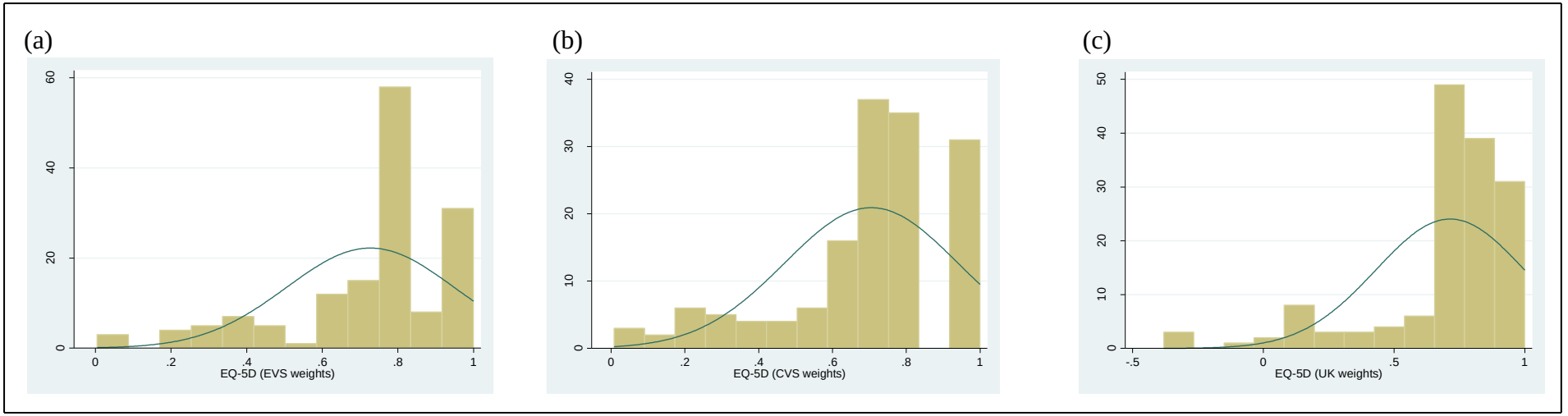
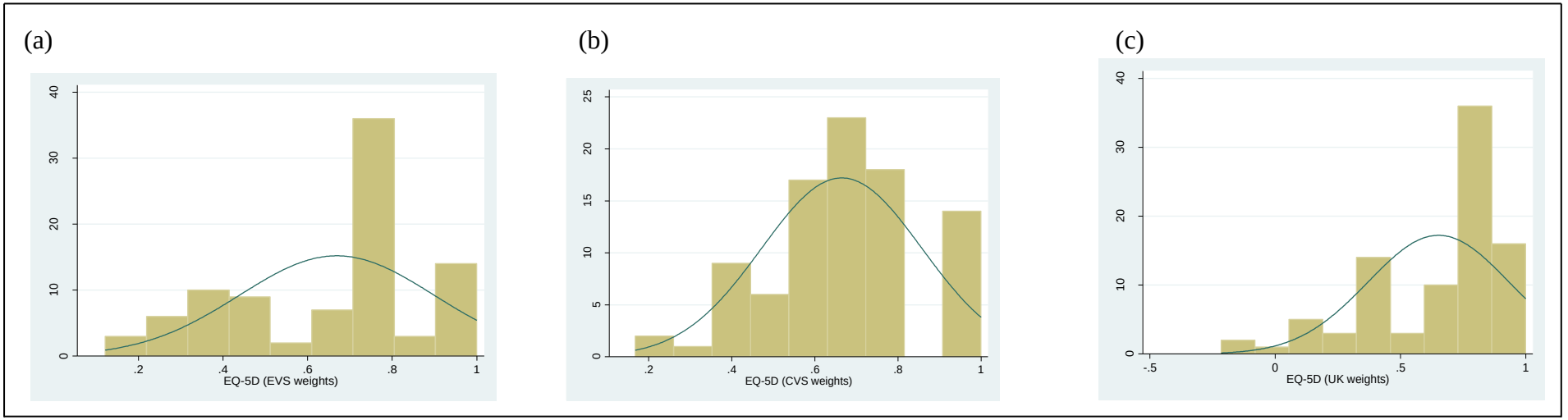


Figure A10.7: DISTRIBUTION OF EQ-5D SCORES FROM CVS EVS AND UKVS (FINLAND)



Appendix 11: Results from the partially split analysis

Case study 1 (Partially-split analysis)

For Case study 1, the partially split one country costing approach (PSOC) considered health outcomes from all participating countries (n=2060) but only considered resource use and costs from the UK (n=329). The partially split multicountry costing approach (PSMC) on the other hand considered outcomes from all participating countries (n=2060) but resource use and costs from the UK, Germany, France, Netherlands and Belgium (n=904).

Results Case study 1 (Partially split analysis)

The results obtained from the partially split analysis alongside Case study 1 are presented in A10.1 below. In terms of resource use, the difference between the intervention and control groups were in the same direction in most cases. The only exception was with GP visits where the intervention group recorded more visits than the control with the PSOC approach whilst the opposite was true with the PSMC approach. In terms of costs, the intervention was associated with a higher cost with both the PSOC and PSMC approaches. However, the difference was not statistically significant (A10.1). Health outcomes were the same for all approaches since all participants in the trial were considered. In terms of cost-effectiveness, both the PSOC and PSMC approaches showed that the intervention (amoxicillin) was cost-effective given the NICE threshold of between £20,000 and £30,000 per QALY gained. The ICERs for the PSOC and PSMC approaches were £10,378 per QALY gained and £5,649 per QALY gained respectively (A10.2).

Comparison to results obtained from the main analysis

The results obtained from the split approaches showed that amoxicillin is cost-effective compared to the control. This result was similar to those obtained with the FPOC, FSMC and

FPMC. Although none of the partially split approaches showed that amoxicillin is dominant. A result which was obtained with the FPOC approach.

Case study 2

For Case study 2, similar to what was done in the main analysis, the only approach that was considered was the partially split one country costing approach (PSOC). This approach considered health outcomes data from all participating countries (n=2,624) but considered resource use and cost data from the UK (n=462). This was mainly due to the relatively small number of countries that were included in the analysis (five countries) and there was no real basis for splitting the five countries into groups.

Results Case study 2

In terms of resource use, the communication skills group was associated with more primary care visits whilst the combined intervention (CRP comm) was associated with more secondary care use (A10.3). Overall, total costs were higher in the combined intervention group (CRP comm) and lowest in the usual care group. With the exception of the combined intervention (CRP comm), there was no significant difference between all other interventions and usual care (A10.4). The cost-effectiveness showed that communication skills was the most cost-effective intervention whilst usual care was the most cost-effective with the cost-utility analysis (A10.6). This result was similar to that obtained with the FPOC approach.

A11.1 Resource use, cost and health outcomes (Case study 1)

		Intervention	Control	Difference (95% CI)[a]
RESOURCE USE				
GP visits	PSOC	1.16 (0.47)	1.12 (0.41)	0.04 (-0.06, 0.12)
	PSMC	1.19 (0.53)	1.23 (0.60)	-0.04 (-0.11, 0.03)
Nurse visits	PSOC	0.01 (0.11)	0.02 (0.15)	-0.01 (-0.04, 0.02)
	PSMC	0.01 (0.08)	0.02 (0.12)	-0.01 (-0.02, 0.004)
Specialist visits	PSOC	0	0	0
	PSMC	0.01 (0.12)	0.02 (0.23)	-0.01 (-0.04, 0.009)
Out of hours visits	PSOC	0.006 (0.08)	0	0.006 (0, 0.02)
	PSMC	0.007 (0.08)	0.002 (0.05)	0.004 (-0.004, 0.014)
Hospital emergency visits	PSOC	0.006 (0.08)	0	0.006 (0, 0.02)
	PSMC	0.002 (0.05)	0.002 (0.05)	0
Prescribed medication	PSOC	0.35 (0.56)	0.24 (0.49)	0.11 (-0.002, 0.22)
	PSMC	0.90 (0.91)	0.85 (0.91)	0.04 (-0.08, 0.16)
Over the counter medication	PSOC	0.88 (1.15)	0.93 (1.25)	-0.04 (-0.30, 0.21)
	PSMC	0.92 (1.12)	0.93 (1.15)	-0.01 (-0.16, 0.13)
COSTS				
Staff costs	PSOC	41.12 (28.27)	38.75 (15.39)	2.36 (-2.29, 7.54)
	PSMC	27.78 (34.93)	28.27 (36.97)	-0.49 (-5.39, 3.80)
Prescribed drugs costs	PSOC	2.52 (6.95)	1.48 (5.88)	1.04 (-0.35, 2.50)
	PSMC	5.43 (10.19)	4.79 (8.96)	0.64 (-0.60, 1.94)
Over the counter drug costs	PSOC	0.83 (1.19)	0.83 (1.13)	-0.008 (-0.26, 0.24)
	PSMC	1.51 (3.46)	1.56 (3.99)	-0.05 (-0.52, 0.40)
Intervention/other drug costs	PSOC	7.03 (12.31)	7.83 (12.04)	-0.79 (-3.19, 2.09)
	PSMC	6.31 (10.55)	7.60 (12.42)	-1.29 (-2.73, 0.29)
Other healthcare costs	PSOC	0.58 (7.42)	0	0.58 (0, 2.22)
	PSMC	0.42 (6.42)	0.47 (9.53)	-0.05 (-1.29, 0.80)
Intervention cost [b]	PSOC	1.20 (0)	0	1.2
	PSMC	3.17 (2.23)	0	3.17 (2.98, 3.38)
Total costs	PSOC	53.27 (32.39)	48.90 (21.32)	4.37 (-0.95, 11.16)
	PSMC	44.62 (39.09)	42.69 (46.83)	1.93 (-4.11, 7.47)
HEALTH OUTCOMES				
EQ-5D baseline	PSOC	0·760 (0·185)	0·752 (0·192)	0·008 (-0·007, 0·024)
	PSMC	0·760 (0·185)	0·752 (0·192)	0·008 (-0·007, 0·024)
EQ-5D week 1	PSOC	0·840 (0·173)	0·824 (0·176)	0·016 (0·002, 0·033)
	PSMC	0·840 (0·173)	0·824 (0·176)	0·016 (0·002, 0·033)
EQ-5D week 2	PSOC	0·908 (0·134)	0·900 (0·134)	0·008 (-0·004, 0·018)
	PSMC	0·908 (0·134)	0·900 (0·134)	0·008 (-0·004, 0·018)
EQ-5D week 3	PSOC	0·929 (0·122)	0·925 (0·122)	0·004 (-0·006, 0·015)
	PSMC	0·929 (0·122)	0·925 (0·122)	0·004 (-0·006, 0·015)
EQ-5D week 4	PSOC	0·936 (0.107)	0·936 (0·109)	0·0001 (-0·010, 0·008)
	PSMC	0·936 (0.107)	0·936 (0·109)	0·0001 (-0·010, 0·008)

[a] Bootstrapped CI [b] Trial intervention costs

A11.2 Cost-effectiveness (Case study 1)

Difference in costs (CI) {CI width}	**PSOC**	£3.84 (-1.98, 9.67) {11.65}
	PSMC	£2.09 (-3.32, 7.49) {10.81}
Difference in QALYs (CI) {CI width}	**PSOC/PSMC**	0.00037 (-0.0002, 0.0009) {0.0011}
ICER	**PSOC**	£10,378 per QALY gained
	PSMC	£5,649 per QALY gained

A11.3 Resource use (Case study 2)

		Usual care	**CRP no Comm**	**Comm no CRP**	**CRP comm**
		PRIMARY CARE VISITS			
GP visits	**PSOC**	0.134 (0.399)	0.117 (0.399)	0.171 (0.549)	0.140 (0.396)
Nurse Visits	**PSOC**	0.010 (0.102)	0.009 (0.095)	0.079 (0.341)	0.026 (0.161)
Out hours GP visits	**PSOC**	0.010 (0.102)	0.009 (0.095)	0	0.018 (0.187)
		SECONDARY CARE VISTIS			
Hospital emergency visits	**PSOC**	0	0	0	0.009 (0.094)
Walk in centre visits	**PSOC**	0	0	0.007 (0.085)	0.009 (0.094)
Specialist visits	**PSOC**	0	0	0.021 (0.118)	0.009 (0.094)
Admissions	**PSOC**	0.010 (0.102)	0	0	0.061 (0.656)
		PRESCRIPTIONS n (%)			
Antibiotic prescription	**PSOC**	72 (74.23%)	47 (42.345)	66 (47.14%)	37 (32.46%)
Over the counter medication	**FSOC**				
CRP test	**FSOC**	2 (2.06%)	91 (81.89%)	1 (0.71%)	55 (48.25%)

A11.4 Cost (Case study 2)

		Usual care	CRP no Comm	Comm no CRP	CRP comm
PRIMARY CARE VISITS					
GP visits	**PSOC**	5.36 (15.95)	4.68 (15.94)	6.85 (21.96)	5.61 (15.86)
Nurse Visits	**PSOC**	0.18 (1.79)	0.16 (1.68)	1.39 (6.02)	0.46 (2.84)
Out hours GP visits	**PSOC**	2.91 (28.63)	2.54 (26.77)	0	4.95 (52.82)
SECONDARY CARE VISTIS					
Hospital emergency visits	**PSOC**	0	0	0	0.98 (10.49)
Walk in centre visits	**PSOC**	0	0	0.29 (3.47)	0.36 (3.84)
Specialist visits	**PSOC**	0	0	3.36 (29.58)	1.38 (14.70)
Admissions	**PSOC**	4.26 (41.93)	0	0	25.36 (270.77)
OTHER COSTS					
Prescription	**PSOC**	1.24 (2.94)	1.54 (3.38)	1.68 (3.86)	1.58 (3.76)
OTC medication	**PSOC**	1.21 (1.25)	1.74 (3.59)	1.40 (1.74)	1.78 (3.56)
CRP test	**PSOC**	0.12 (0.82)	4.72 (2.22)	0.04 (0.48)	2.78 (2.89)
Trial intervention cost	**PSOC**	0	22.35 (0)	10.61 (0)	26.33 (0)
TOTAL COSTS					
Total cost[a]	**PSOC**	15.68 (53.03)	38.45 (39.00)	26.77 (46.23)	72.28 (334.11)
Total cost[b]		17.12	27.00	36.10	73.07
Difference (95% CI)[c]			18.98 (-26.70, 64.65)	9.88 (-33.50, 53.26)	55.95 (10.60, 101.30)
CI Width			91.35	88.76	90.70

[a] Unadjusted costs [b] Adjusted costs (3-level model) excludes cost of resistance [c] Difference with reference to usual care

A11.5 Mean EQ-5D scores over 4 weeks (Case study 2)

		Usual care	CRP no Comm	Comm no CRP	CRP comm
			EQ-5D		
Baseline	PSOC	0.720 (0.205)	0.692 (0.218)	0.711 (0.199)	0.715 (0.233)
Week 1	PSOC	0.759 (0.202)	0.761 (0.220)	0.772 (0.219)	0.763 (0.228)
Week 2	PSOC	0.835 (0.203)	0.832 (0.196)	0.830 (0.206)	0.826 (0.221)
Week 3	PSOC	0.851 (0.200)	0.845 (0.210)	0.873 (0.186)	0.842 (0.223)
Week 4	PSOC	0.869 (0.197)	0.858 (0.205)	0.882 (0.187)	0.849 (0.209)
			QUALITY ADJUSTED LIFE YEARS		
QALYs[a]	PSOC	0.0656 (0.012)	0.0657 (0.013)	0.0643 (0.012)	0.0646 (0.013)
QALYs[b]		0.0649	0.0647	0.0648	0.0643
Difference (95% CI)[c]			-0.0002 (-0.002, 0.001)	-0.0001 (-0.002, 0.001)	-0.0006 (-0.002, 0.001)
CI Width			0.003	0.003	0.003
			ANTIBIOTIC PRESCRIBING		
Antibiotic Prescribing [a]	PSOC	0.596 (0.491)	0.336 (0.473)	0.409 (0.492)	0.341 (0.474)
Antibiotic Prescribing [b]		0.555	0.350	0.397	0.341
Difference (95% CI)[c]			-0.204 (-0.304, -0.103)	-0.157 (-0.260, -0.054)	-0.213 (-0.309, -0.118)
CI Width			0.201	0.206	0.191

[a] Unadjusted [b] Adjusted (3-level model) [c] Difference with reference to usual care

A11.6 Cost-effectiveness analysis (Case study 2)

	Mean cost (£)	QALYs	ICER (£ per QALY)
CRP comm	73.07	0.0643	Dominated by usual care
CRP no comm	36.10	0.0647	Dominated by usual care
Comm no CRP	27.00	0.0648	Dominated by usual care
Usual care	17.12	0.0649	N/A
	COST-EFFECTIVENESS ANALYSIS		**ICER (£ per unit reduction in prescribing)**
CRP comm	73.07	0.341	4107[a]
CRP no comm	36.10	0.350	193.61[b]
Comm no CRP	27.00	0.397	62.53[c]
Usual care	17.12	0.555	N/A

[a] ICER derived from a comparison of CRP comm with CRP no comm

[b] ICER derived from a comparison of CRP no comm with Comm no CRP [c] ICER derived from a comparison of CRP no comm and Usual care

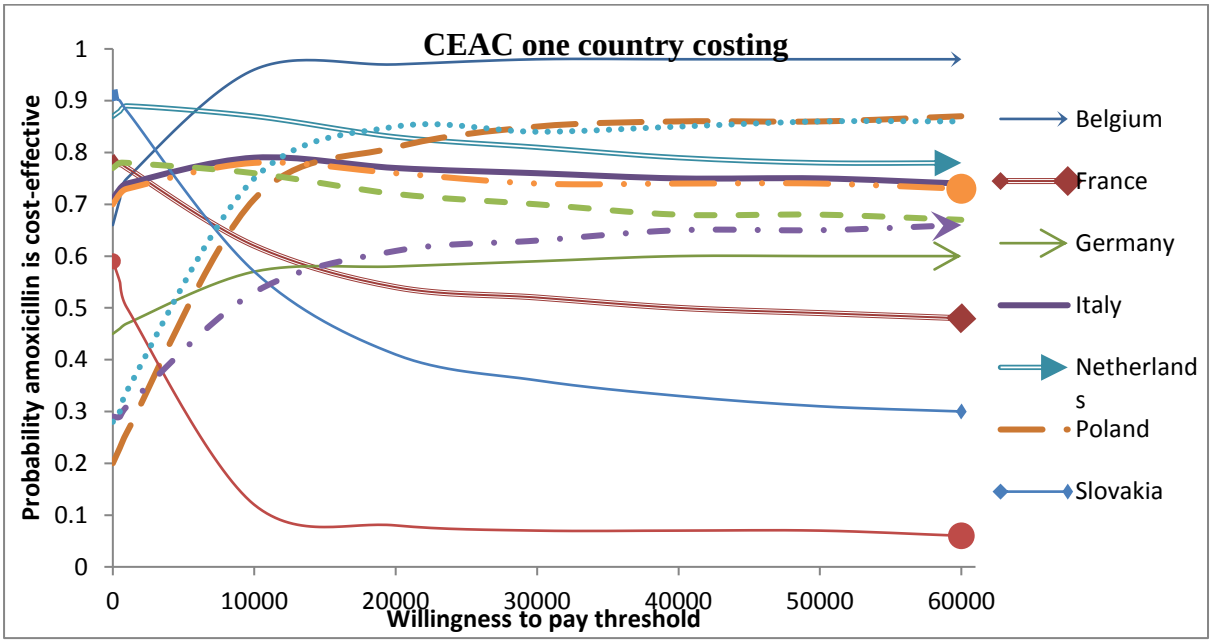
Appendix 12: Winbugs code for estimating country-specific cost-effectiveness

```
model{
for(i in 1:N){
tmp1[i] <- id[i]
tmp2[i] <- eqbase[i]
tmp3[i] <- country[i]
out[i]        ~  dnorm(muout[i],precout[trt[i]])
costs[i]      ~  dnorm(mucosts[i],preccosts[trt[i]])
muout[i]     <- cmeanout[trt[i],country[i]]     + betaEQbase1*(eqbase[i] - mean(eqbase[]))
mucosts[i]  <- cmeancosts[trt[i],country[i]] + lambda[trt[i]]*(out[i] -
cmeanout[trt[i],country[i]]) + betaEQbase2*(eqbase[i] - mean(eqbase[]))
}

for(j in 1:nt){
   for(k in 1:ncountries){
        cmeanout[j,k]      ~ dnorm(meanout[j],preccout[j])         #country specific random
effects for outcome
        cmeancosts[j,k]  ~ dnorm(meancosts[j],precccosts[j])   #country specific random
effects for outcome

#country-specific cost-effectiveness
cinccosts[j,k] <- cmeancosts[j,k] - cmeancosts[1,k]
cincout[j,k]     <- cmeanout[j,k]    - cmeanout[1,k]
cicer[j,k]      <- cinccosts[j,k]/cincout[j,k]
nb[j,k]<-20000*(cincout[j,k])-(cinccosts[j,k])
pce[j,k] <- step(nb[j,k])
}
preccout[j]      <- pow(secout[j],-2)
precccosts[j]  <- pow(seccosts[j],-2)
secout[j]        ~ dunif(0,10)
seccosts[j]    ~ dunif(0,100)
}

for(j in 1:nt){
precout[j]     <- pow(seout[j],-2)                                    #precout = 1/variance
outcome
preccosts[j] <- 1/varcosts[j]                                        #preccosts = 1/variance costs
varcosts[j]    <- pow(secosts[j],2)-(pow(seout[j],2)*pow(lambda[j],2))  #conditional variances
for costs

#prior distributions
secosts[j]      ~  dunif(0,2000)
seout[j]         ~  dunif(0,10)
meancosts[j]  ~ dnorm(0,0.00000001)I(0,)
```

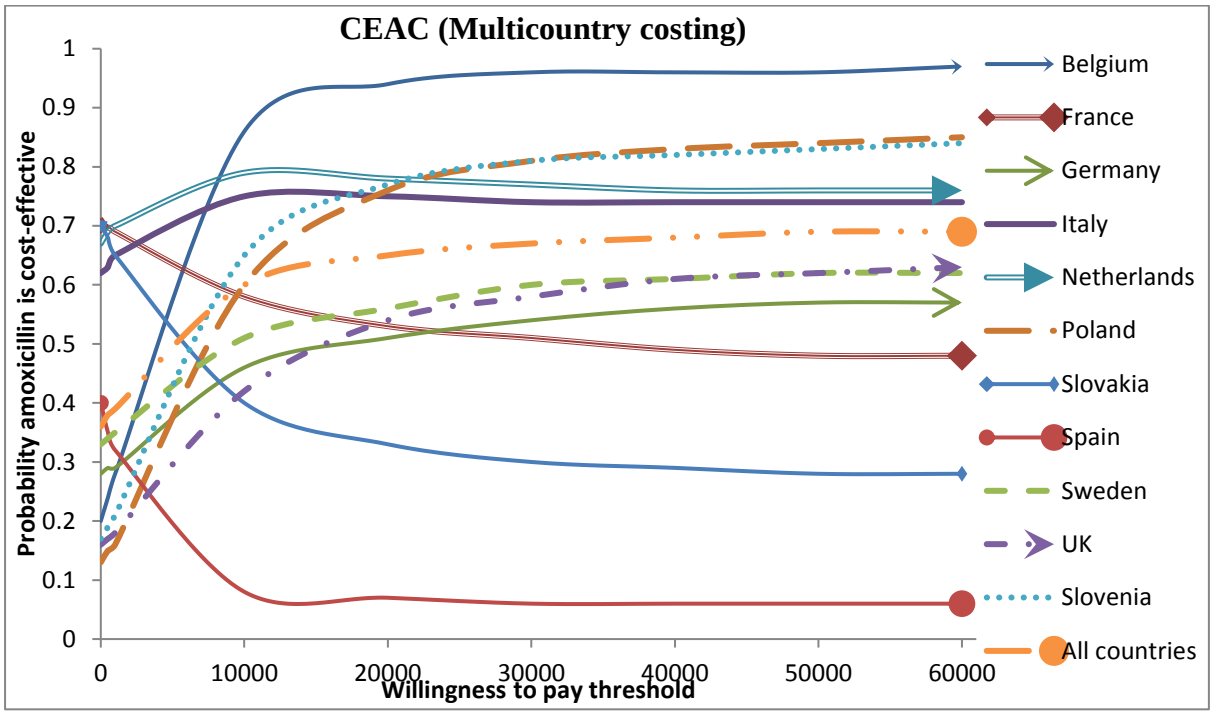
```
meanout[j]     ~ dnorm(0,0.0001)
lambda[j]      ~ dunif(-10,10)

#cost-effectiveness
inccosts[j] <- meancosts[j] - meancosts[1]
incout[j]     <- meanout[j] - meanout[1]
icer[j]        <- inccosts[j]/incout[j]
onb[j] <-20000*(incout[j])-(inccosts[j])
opce[j]<-step(onb[j])

}
betaEQbase1  ~ dnorm(0,0.0001)
betaEQbase2  ~ dnorm(0,0.0001)
}
```

Appendix 13: Country-specific cost-effectiveness acceptability curves

A13.1 Country-specific cost-effectiveness acceptability curves for antibiotics vs control (sensitivity analysis with one country costing)



A13.2 Country-specific cost-effectiveness acceptability curves for antibiotics vs control (sensitivity analysis with multi country costing)



LIST OF REFERENCES

ALBREHT, T., PRIBAKOVIC, B. R., JOSAR, D., POLDRUGOVAC, M., KOSTNAPFEL, T., ZALETEL, M., PANTELI, D. & MARESSO, A. 2016. Slovenia: Health System Review. *Health Systems in Transition,* 18, 1-207.

AL-JANABI, H., FLYNN, T. N. & COAST, J. 2011. QALYs and carers. *Pharmacoeconomics,* 29**,** 1015-1023.

ANELL, A., GLENNGARD, A. H. & MERKUR, S. M. 2012. Sweden: Health system review. *Health Systems in Transition,* 14**,** 1-159.

ANNEMANS, L., MOEREMANS, K., LAMOTTE, M., CONDE, J. G., VAN DEN BERG, H., MYINT, H., PIETERS, R. & UYTTEBROECK, A. 2003. Pan-European multicentre economic evaluation of recombinant urate oxidase (rasburicase) in prevention and treatment of hyperuricaemia and tumour lysis syndrome in haematological cancer patients. *Supportive Care in Cancer,* 11**,** 249-257.

ANTHIERENS, S., TONKIN-CRINE, S., CALS, J. W., COENEN, S., YARDLEY, L., BROOKES-HOWELL, L., FERNANDEZ-VANDELLOS, P., KRAWCZYK, J., GODYCKI-CWIRKO, M. & LLOR, C. 2015. Clinicians' views and experiences of interventions to enhance the quality of antibiotic prescribing for acute respiratory tract infections. *Journal of General Internal Medicine,* 30**,** 408-416.

ANTHIERENS, S., TONKIN-CRINE, S., DOUGLAS, E., FERNANDEZ-VANDELLOS, P., KRAWCZYK, J., LLOR, C., CALS, J. W., FRANCIS, N. A., YARDLEY, L. & COENEN, S. 2012. General practitioners' views on the acceptability and applicability of a web-based intervention to reduce antibiotic prescribing for acute cough in multiple European countries: a qualitative study prior to a randomised trial. *BMC Family Practice,* 13**,** 101.

ANTONANZAS, F., RODRÍGUEZ-IBEAS, R., JUÁREZ, C., HUTTER, F., LORENTE, R. & PINILLOS, M. 2009. Transferability indices for health economic evaluations: methods and applications. *Health Economics,* 18**,** 629-643.

APPLEBY, J., DEVLIN, N. & PARKIN, D. 2007. NICE's cost effectiveness threshold. *British Medical Journal,* 335**,** 358-359.

ARROW, K. J. 1963. Uncertainty and the welfare economics of medical care. *The American Economic Review,* 53, 941-973.

ASPELIN, P., AUBRY, P., FRANSSON, S.-G., STRASSER, R., WILLENBROCK, R. & LUNDKVIST, J. 2005. Cost-effectiveness of iodixanol in patients at high risk of contrast-induced nephropathy. *American Heart Journal,* 149**,** 298-303.

AUGUSTINE, E. F., ADAMS, H. R. & MINK, J. W. 2013. Clinical Trials in Rare Disease Challenges and Opportunities. *Journal of Child Neurology,* 28**,** 1142-1150.

BACHERT, C., VESTENBAEK, U., CHRISTENSEN, J., GRIFFITHS, U. & POULSEN, P. 2007. Cost-effectiveness of grass allergen tablet (GRAZAX®) for the prevention of seasonal grass pollen induced rhinoconjunctivitis–a Northern European perspective. *Clinical & Experimental Allergy,* 37**,** 772-779.

BADIA, X., ROSET, M., HERDMAN, M. & KIND, P. 2001. A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. *Medical Decision Making,* 21**,** 7-16.

BADIA, X., ROSET, M., MONSERRAT, S. & HERDMAN, M. The Spanish VAS tariff based on valuation of EQ-5D health states from the general population. Proceedings of the EuroQol Plenary Meeting, Rotterdam, 1997.

BAKHAI, A., STONE, G. W., GRINES, C. L., MURPHY, S. A., GITHIORA, L., BEREZIN, R. H., COX, D. A., STUCKEY, T., GRIFFIN, J. J. & TCHENG, J. E. 2003. Cost-Effectiveness of Coronary Stenting and Abciximab for Patients With Acute Myocardial Infarction Results From the CADILLAC (Controlled Abciximab and Device Investigation to Lower Late Angioplasty Complications) Trial. *Circulation,* 108**,** 2857-2863.

BALL, P., BAQUERO, F., CARS, O., FILE, T., GARAU, J., KLUGMAN, K., LOW, D., RUBINSTEIN, E. & WISE, R. 2002. Antibiotic therapy of community respiratory tract infections: strategies for optimal outcomes and minimized resistance emergence. *Journal of Antimicrobial Chemotherapy,* 49**,** 31-40.

BANQUEDEFRANCE. Exchange rates Available: https://www.banque-france.fr/en/economics-statistics/rates/exchange-rates.html [Accessed 6th October, 2016].

BARBIERI, M., DRUMMOND, M., RUTTEN, F., COOK, J., GLICK, H. A., LIS, J., REED, S. D., SCULPHER, M. & SEVERENS, J. L. 2010. What do international pharmacoeconomic guidelines say about economic data transferability? *Value in Health,* 13**,** 1028-1037.

BARNSLEY, P., TOWSE, A., KARLSBERG SCHAFFER, S. & SUSSEX, J. 2013. Critique of CHE Research Paper 81: Methods for the Estimation of the NICE Cost Effectiveness Threshold. *Available at* https://www.ohe.org/publications/critique-che-research-paper-81-methods-estimation-nice-cost-effectiveness-threshold [Accessed 12th September 2016]

BARTON, M. G. R., BANKART, J., DAVIS, A. C. & SUMMERFIELD, Q. A. 2004. Comparing utility scores before and after hearing-aid provision. *Applied Health Economics and Health Policy,* 3**,** 103-105.

BEINART, S. C., KOLM, P., VELEDAR, E., ZHANG, Z., MAHONEY, E. M., BOUIN, O., GABRIEL, S., JACKSON, J., CHEN, R. & CARO, J. 2005. Long-term cost effectiveness of early and sustained dual oral antiplatelet therapy with clopidogrel given for up to one year after percutaneous coronary intervention: results from the clopidogrel for the reduction of events during observation (CREDO) trial. *Journal of the American College of Cardiology,* 46**,** 761-769.

BELL, S.A. & TUDUR SMITH C. 2014. A comparison of interventional trials in rare versus non-rare diseases: an analysis of ClinicalTials.gov *Orphanet Journal of Rare Diseases* 9:170

BERNERT, S., FERNÁNDEZ, A., HARO, J. M., KÖNIG, H. H., ALONSO, J., VILAGUT, G., SEVILLA-DEDIEU, C., DE GRAAF, R., MATSCHINGER, H. & HEIDER, D. 2009. Comparison of Different Valuation Methods for Population Health Status Measured by the EQ-5D in Three European Countries. *Value in Health,* 12**,** 750-758.

BIRCH, S. & GAFNI, A. 2002. On being NICE in the UK: guidelines for technology appraisal for the NHS in England and Wales. *Health Economics,* 11**,** 185-191.

BIRCH, S. & DONALDSON, C. 2003. Valuing the benefits and costs of health care programmes: where's the 'extra'in extra-welfarism? *Social Science & Medicine,* 56, 1121-1133.

BLACK, N. 1996. Why we need observational studies to evaluate the effectiveness of health care. *British Medical Journal,* 312(7040): 1215-1218.

BOEHLER, C. E. & LORD, J. 2015. Mind the gap! A multilevel analysis of factors related to variation in published cost-effectiveness estimates within and between countries. *Medical Decision Making***,** DOI: 10.1177/0272989X15579173

BOLIN, K., SANDIN, R., KOLTOWSKA-HÄGGSTRÖM, M., LOFTUS, J., PRÜTZ, C. & JONSSON, B. 2013. The cost-effectiveness of growth hormone replacement therapy

(Genotropin®) in hypopituitary adults in Sweden. *Cost Effectiveness and Resource Allocation,* 11:24 DOI:10.1186/1478-7547-1124

BOULENGER, S., NIXON, J., DRUMMOND, M., ULMANN, P., RICE, S. & DE POUVOURVILLE, G. 2005. Can economic evaluations be made more transferable? *The European Journal of Health Economics,* 6**,** 334-346.

BRACCO, A., JÖNSSON, B., RICCI, J. F., DRUMMOND, M. & NYHLIN, H. 2007. Economic evaluation of tegaserod vs. placebo in the treatment of patients with irritable bowel syndrome: an analysis of the TENOR study. *Value in Health,* 10**,** 238-246.

BRAZIER, J., HARPER, R., MUNRO, J., WALTERS, S. & SNAITH, M. 1999. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. *Rheumatology,* 38**,** 870-877.

BRAZIER, J., ROBERTS, J., TSUCHIYA, A. & BUSSCHBACH, J. 2004. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Economics,* 13**,** 873-884.

BRIDGES, J. F., ONUKWUGHA, E. & MULLINS, C. D. 2010. Healthcare rationing by proxy. *Pharmacoeconomics,* 28**,** 175-184.

BRIGGS, A.H. 2010. Transportability of Comparative Effectiveness and Cost-Effectiveness between Countries. *Value in Health,* 13**,** S22-S25.

BRIGGS, A.H., BOUSQUET, J., WALLACE, M., BUSSE, W. W., CLARK, T., PEDERSEN, S. & BATEMAN, E. 2006. Cost-effectiveness of asthma control: an economic appraisal of the GOAL study. *Allergy,* 61**,** 531-536.

BRIGGS, A.H., GLICK, H.A., LOZANO-ORTEGA, G., SPENCER, M., CALVERLEY, P., JONES, P., VESTBO, J. & INVESTIGATORS, T. A. R. I. C. H. 2010. Is treatment with ICS and LABA cost-effective for COPD? Multinational economic analysis of the TORCH study. *European Respiratory Journal,* 35**,** 532-539.

BRIGGS, A. H. & O'BRIEN, B. J. 2001. The death of cost-minimization analysis? *Health Economics,* 10**,** 179-184.

BROUWER, W. B., CULYER, A. J., VAN EXEL, N. J. A. & RUTTEN, F. F. 2008. Welfarism vs. extra-welfarism. *Journal of Health Economics,* 27**,** 325-338.

BROWN, R. & ARMSTRONG, P. 2003. Cost effectiveness in Canada of eptifibatide treatment for acute coronary syndrome patients using PURSUIT subgroup analysis. *The Canadian Journal of Cardiology,* 19**,** 161-166.

BRUYNDONCKX, R., HENS, N., AERTS, M., GOOSSENS, H., ABRAHANTES, J. C. & COENEN, S. 2015. Exploring the association between resistance and outpatient antibiotic use expressed as DDDs or packages. *Journal of Antimicrobial Chemotherapy***,** 70(4), 1241-1244.

BURGER, E. A., SY, S., NYGÅRD, M., KRISTIANSEN, I. S. & KIM, J. J. 2014. Prevention of HPV-related cancers in Norway: cost-effectiveness of expanding the HPV vaccination program to include pre-adolescent boys. PLoS One, 9, e89974.

BUTLER, C. C., HOOD, K., VERHEIJ, T., LITTLE, P., MELBYE, H., NUTTALL, J., KELLY, M. J., MÖLSTAD, S., GODYCKI-CWIRKO, M. & ALMIRALL, J. 2009. Variation in antibiotic prescribing and its impact on recovery in patients with acute cough in primary care: prospective study in 13 countries. *British Medical Journal,* 338:b2242

BUXTON, M., SULLIVAN, S., ANDERSSON, L., LAMM, C., LILJAS, B., BUSSE, W., PEDERSEN, S. & WEISS, K. 2004. Country-specific cost-effectiveness of early intervention with budesonide in mild asthma. *European Respiratory Journal,* 24**,** 568-574.

CALS, J.W.L., BUTLER, C.C., HOPSTAKEN, R.M., HOOD, K., DINANT, G.J. 2009 Effect of point of care testing for C reactive protein and training in communication skills on antibiotic use in lower respiratory tract infections: cluster randomised trial. *British Medical Journal* 338 b1374.

CALS, J. W., AMENT, A. J., HOOD, K., BUTLER, C. C., HOPSTAKEN, R. M., WASSINK, G. F. & DINANT, G. J. 2011. C-reactive protein point of care testing and physician communication skills training for lower respiratory tract infections in general practice: economic evaluation of a cluster randomized trial. *Journal of Evaluation in Clinical Practice,* 17**,** 1059-1069.

CANONICA, G., POULSEN, P. & VESTENBAEK, U. 2007. Cost-effectiveness of GRAZAX® for prevention of grass pollen induced rhinoconjunctivitis in Southern Europe. *Respiratory Medicine,* 101**,** 1885-1894.

CANOUI-POITRINE, F., JEANBLANC, G., ALBERTI, C., ARMOOGUM, P., CEBRIAN, A., CARRIÉ, D., HENRY, P., TEIGER, E., SLAMA, M. & SPAULDING, C. 2009. Cost effectiveness of sirolimus-eluting stents compared with bare metal stents in acute myocardial infarction. *Applied Health Economics and Health Policy,* 7**,** 19-29.

CANTOR, S. B. 1994. Cost-effectiveness analysis, extended dominance, and ethics a quantitative assessment. *Medical Decision Making,* 14**,** 259-265.

CDC. 2011. .*World Health Day: Media Fact Sheet, April 7, 2011. 2011.* [Online]. Centers for Disease Control and Prevention. Available: http://www.cdc.gov/media/releases/2011/f0407_antimicrobialresistance.html [Accessed 6th October, 2016].

CDDEP (The centre for disease dynamics, economics and policy) Resistance Map http://resistancemap.cddep.org/Countries.php [Accessed on 21st October, 2016]

CHUNG A., PERERA, P., BRUEGGEMANN, AB., ELAMIN, AE., et al. 2007 Effect of Antibiotic Prescribing on Antibiotic Resistance in Individual Children in Primary Care: Prospective Cohort Study *British Medical Journa*l 335:429

CLAXTON, K., MARTIN, S., SOARES, M., RICE, N., SPACKMAN, E., HINDE, S., DEVLIN, N., SMITH, P. C. & SCULPHER, M. 2015. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technology Assessment* PP1-542. ISSN 2046-4924

CLEEMPUT, I. 2003. *Economic Evaluation in Renal Transplantation: Outcome Assessment & Cost-utility of Non-compliance,* Katholieke Universiteit Leuven.

CLEEMPUT, I. 2010. A social preference valuations set for EQ-5D health states in Flanders, Belgium. *The European Journal of Health Economics,* 11**,** 205-213.

CLEEMPUT, I., KESTELOOT, K., MOONS, P., VANRENTERGHEM, Y., VAN HOOFF, J. P., SQUIFFLET, J. P. & DE GEEST, S. 2004. The Construct and Concurrent Validity of the EQ-5D in a Renal Transplant Population. *Value in Health,* 7**,** 499-509.

CLEEMPUT, I., NEYT, M., THIRY, N., DE LAET, C. & LEYS, M. 2008. Threshold values for cost-effectiveness in health care. *Health Technology Assessment (HTA). Brussels: Belgian Health Care Knowledge Centre (KCE).* https://kce.fgov.be/publication/report/threshold-values-for-cost-effectiveness-in-health-care#.WA6CFoWcGUk [Accessed on 24th October, 2016]

COAST, J. 2004. Is economic evaluation in touch with society's health values? *British Medical Journal,* 329**,** 1233.

COAST, J. & SMITH, R. 2001. Antimicrobial resistance: can economics help? *Eurohealth,* 7**,** 32-33.

COAST, J., SMITH, R., KARCHER, A. M., WILTON, P. & MILLAR, M. 2002. Superbugs II: how should economic evaluation be conducted for interventions which aim to contain antimicrobial resistance? *Health Economics,* 11**,** 637-647.

COAST, J., SMITH, R. D. & LORGELLY, P. 2008. Welfarism, extra-welfarism and capability: the spread of ideas in health economics. *Social Science & Medicine,* 67**,** 1190-1198.
COAST, J., SMITH, R. D. & MILLAR, M. R. 1996. Superbugs: should antimicrobial resistance be included as a cost in economic evaluation? *Health Economics,* 5**,** 217-226.
COAST, J. 2009. Maximisation in extra-welfarism: A critique of the current position in health economics. *Social Science & Medicine,* 69, 786-792.
COLLINS, M. & LATIMER, N. 2013. NICE's end of life decision making scheme: impact on population health. *British Medical Journal,* 346**,** f1363.
CONNER-SPADY, B. L., MARSHALL, D. A., BOHM, E., DUNBAR, M. J., LOUCKS, L., KHUDAIRY, A. A. & NOSEWORTHY, T. W. 2015. Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee replacement. *Quality of Life Research,* 24**,** 1775-84.
COOK, J. R., DRUMMOND, M., GLICK, H.A. & HEYSE, J. F. 2003. Assessing the appropriateness of combining economic data from multinational clinical trials. *Statistics in Medicine,* 22**,** 1955-1976.
COOKSON, R. & CLAXTON, K. 2012. The Humble Economist: Tony Culyer on Health. *Health Care and Social Decision Making. York and London: University of York and Office of Health Economics*.
COSTELLOE, C., METCALFE, C., LOVERING, A., MANT, D. & HAY, A. D. 2010. Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: systematic review and meta-analysis. *British Medical Journal,* 340**,** c2096.
CRAIG, B. M., PICKARD, A. S. & LUBETKIN, E. I. 2014. Health problems are more common, but less severe when measured using newer EQ-5D versions. *Journal of Clinical Epidemiology,* 67**,** 93-9.
CRD. 2009. *Centre for Reviews and Dissemination University of York. Systematic reviews: CRD's guidance for undertaking reviews in health care* [Online]. Available: http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf [Accessed 24th September, 2012].
CULYER, A. J. 1989. The normative economics of health care finance and provision. *Oxford Review of Economic Policy*, 5, 34-58.
CUNNINGHAM, S. 2000. Economics: Economic evaluation of healthcare–is it important to us? *British Dental Journal,* 188**,** 250-254.
Curtis L & Burns A (2015) Unit costs of health and social care 2015 Personal Social Services Research Unit http://www.pssru.ac.uk/project-pages/unit-costs/2015/ [Accessed 19th October, 2016]
CVZ (2008) Guidance for outcomes research 'for the assessment of the cost-effectiveness of in-patient medicines' http://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/publicaties/publications-in-english/2008/0812-guidance-for-outcomes-research/0812-guidance-for-outcomes-research/Guidance%2Bfor%2BOutcomes%2BResearch.pdf [Accessed on 7th October, 2016]
DAHLER-ERIKSEN B.S., LAURITZEN T., LASSEN, J.F., LUND E.D., BRANDSLUND, I. (1999) Near-patient test for C-reactive protein in general practice: assessment of clinical, organizational, and economic outcomes. *Clinical Chemistry* 45[4], 478.
DAKIN, H. & WORDSWORTH, S. 2013. Cost-Minimisation Analysis Versus Cost-Effectiveness Analysis, Revisited. *Health Economics,* 22**,** 22-34.

DE SMET, P. A. 2005. Herbal medicine in Europe--relaxing regulatory standards.
DEPARTMENTOFHEALTH. *UK five year antimicrobial resistance strategy* [Online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/244058/20130902_UK_5_year_AMR_strategy.pdf [Accessed 27th November 2015].
DEVLIN, N. & PARKIN, D. 2007. Guidance to users of EQ-5D value sets. *EQ-5D Value Sets.* Springer.
DEVLIN, N., SHAH, K, FENG, Y, MULHERN, B, VAN HOUT, B 2016. Valuing Health Related Quality of Life: An EQ-5D-5L Value Set for England. *Office of Health Economics Research Paper 16/01*.
Diederichsen HZ, Skamling M, Diederichsen A, Grinsted P, Antonsen S, Petersen PH, et al. 2000 Randomised controlled trial of CRP rapid test as a guide to treatment of respiratory infections in general practice. *Scandinavian Journal of Primary Health Care* 18[1], 39-43. Informa UK Ltd UK.
DOLAN, P. 1997. Modeling valuations for EuroQol health states. *Medical Care,* 35**,** 1095-1108.
DONNER, A. & KLAR, N. 2000. *Design and analysis of cluster randomization trials in health research*. Wiley
DRUMMOND, M., BARBIERI, M., COOK, J., GLICK, H. A., LIS, J., MALIK, F., REED, S. D., RUTTEN, F., SCULPHER, M. & SEVERENS, J. 2009. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. *Value in Health,* 12**,** 409-418.
DRUMMOND, M., MANCA, A. & SCULPHER, M. 2005. Increasing the generalizability of economic evaluations: recommendations for the design, analysis, and reporting of studies. *International journal of technology assessment in health care,* 21**,** 165-171.
DRUMMOND, M. & PANG, F. 2001. Transferability of economic evaluation results. in DRUMMOND M & McGUIRE, A (eds) *Economic evaluation in health care: merging theory with practice***,** Oxford University Press 256-276.
DRUMMOND, M. F. 2015. *Methods for the economic evaluation of health care programmes Fourth Edition,* Oxford university press.
DRUMMOND, M. F., BECKER, D. L., HUX, M., CHANCELLOR, J. V., DUPRAT-LOMON, I., KUBIN, R. & SAGNIER, P.-P. 2003. An economic evaluation of sequential iv/po moxifloxacin therapy compared to iv/po co-amoxiclav with or without clarithromycin in the treatment of community-acquired pneumonia. *CHEST Journal,* 124**,** 526-535.
DRUMMOND, M. F. & JEFFERSON, T. 1996. Guidelines for authors and peer reviewers of economic submissions to the *British Medical Journal,* 313**,** 275-283.
DUKHOVNY, D., LORCH, S. A., SCHMIDT, B., DOYLE, L. W., KOK, J. H., ROBERTS, R. S., KAMHOLZ, K. L., WANG, N., MAO, W. & ZUPANCIC, J. A. 2011. Economic evaluation of caffeine for apnea of prematurity. *Pediatrics,* 127**,** e146-e155.
DURHAM, J., STEELE, J. G., BRECKONS, M., STORY, W. & VALE, L. 2015. DEEP Study: does EQ-5D-5L measure the impacts of persistent oro-facial pain? *Journal of Oral Rehabilitation*. 42(9): 643-50
ECDC 2009. The bacterial challenge: time to react. European Centre for Disease Prevention and Control and European Medicines Agency joint report. 2009. http://ecdc.europa.eu/en/publications/Publications/0909_TER_The_Bacterial_Challenge_Time_to_React.pdf [Accessed 24th October, 2016]
ECRIN (European clinical research infrastructure network) Funding for multinational trials http://www.ecrin.org/knowledge/funding-multinational-clinical-trials [Accessed 30th November, 2016]

EDBROOKE, D. L., MINELLI, C., MILLS, G. H., IAPICHINO, G., PEZZI, A., CORBELLA, D., JACOBS, P., LIPPERT, A., WIIS, J. & PESENTI, A. 2011. Implications of ICU triage decisions on patient mortality: a cost-effectiveness analysis. *Critical Care,* 15**,** R56.

EDEJER, T. T.-T. 2003. *Making choices in health: WHO guide to cost-effectiveness analysis*, World Health Organization. http://www.who.int/choice/publications/p_2003_generalised_cea.pdf [Accessed 24th October, 2016]

ESAC. 2014. *Summary of the latest data on antibiotic consumptionin the European Union November 2014* [Online]. Available: http://ecdc.europa.eu/en/eaad/Documents/antibiotic-consumptio-ESAC-Net-2014-EAAD.pdf [Accessed 9th March, 2016 2016].

EUROSTAT. *Eurostat* [Online]. Available: http://ec.europa.eu/eurostat/tgm/table.do?tab=table&language=en&pcode=tps00001&tableSelection=1&footnotes=yes&labeling=labels&plugin=1 [Accessed 24th October, 2016].

FENG, Y., PARKIN, D. & DEVLIN, N. J. 2014. Assessing the performance of the EQ-VAS in the NHS PROMs programme. *Quality of Life Research,* 23**,** 977-989.

FERRÉ, F., DE BELVIS, A. G., VALERIO, L., LONGHI, S., LAZZARI, A., FATTORE, G., RICCIARDI, W. & MARESSO, A. 2014. Italy: health system review. *Health Systems in Transition* 16(4)

FERNANDEZ, J.-L., MONTGOMERY, S. & FRANCOIS, C. 2005. Evaluation of the cost effectiveness of escitalopram versus venlafaxine XR in major depressive disorder. *Pharmacoeconomics,* 23**,** 155-167.

FRANKEN, M. 2014. Decision making in drug reimbursement. http://www.bmg.eur.nl/fileadmin/ASSETS/bmg/Onderzoek/Promoties/Promoties_2014/Decision_making_in_drug_reimbursement_M_Franken_2014.pdf [Accessed 30th September, 2016]

FREMPONG, S. N., GORANITIS, I. & OPPONG, R. 2015. Economic evaluation alongside factorial trials: a systematic review of empirical studies. *Expert Review of Pharmacoeconomics & Outcomes Research*, 15**,** 801-811.

FRICK, K. D. 2009. Micro-costing quantity data collection methods. *Medical Care*, 47, S76.

FRIEDMAN, M. 1953. The methodology of positive economics. in HAUSMAN D,M (eds) The philosophy of economics an anthology (third edition) Cambridge

FRITH, J. & NEWTON, J. L. 2015. Validation of a questionnaire for orthostatic hypotension for routine clinical use. *Geriatrics and Gerontology International*. 16(7):785-90

GARCIA-GORDILLO, M., DEL POZO-CRUZ, B., ADSUAR, J. C., SANCHEZ-MARTINEZ, F. I. & ABELLAN-PERPINAN, J. M. 2014. Validation and comparison of 15-D and EQ-5D-5L instruments in a Spanish Parkinson's disease population sample. *Quality of Life Research,* 23**,** 1315-26.

GARRY, R., FOUNTAIN, J., BROWN, J., MANCA, A., MASON, S., SCULPHER, M., NAPP, V., BRIDGMAN, S., GRAY, J. & LILFORD, R. 2004. EVALUATE hysterectomy trial: a multicentre randomised trial comparing abdominal, vaginal and laparoscopic methods of hysterectomy, National Coordinating Centre for Health Technology Assessment. 8(26): 1-154

GHEORGHE, A., ROBERTS, T. E., IVES, J. C., FLETCHER, B. R. & CALVERT, M. 2013. Centre selection for clinical trials and the generalisability of results: a mixed methods study. *PLoS One* 8(2): e56560

GLASZIOU, P. P., CLARKE, P., ALEXANDER, J., RAJMOKAN, M., BELLER, E., WOODWARD, M., CHALMERS, J., POULTER, N. & PATEL, A. 2010. Cost-

effectiveness of lowering blood pressure with a fixed combination of perindopril and indapamide in type 2 diabetes mellitus: an ADVANCE trial-based analysis. *The Medical Journal of Ausraliat,* 193**,** 320-4.
GLICK, H. A., DOSHI, J. A., SONNAD, S. S. & POLSKY, D. 2007 Economic evaluation in clinical trials, Oxford University Press

GOA. 2011. *Antibiotic resistance data gaps will remain despite HHS taking steps to improve monitoring United States Government Accountability Office Report to the Committee on Agriculture, House of Representatives GOA-11-406. 2011* [Online]. [Accessed].
GOLDSMITH, K. A., DYER, M. T., SCHOFIELD, P. M., BUXTON, M. J. & SHARPLES, L. D. 2009. Relationship between the EQ-5D index and measures of clinical outcomes in selected studies of cardiovascular interventions. *Health and Quality of Life Outcomes,* 7**,** 96.
GOLDSTEIN, H. 2011. *Multilevel statistical models,* John Wiley & Sons.
GOLICKI, D., NIEWADA, M., BUCZEK, J., KARLINSKA, A., KOBAYASHI, A., JANSSEN, M. F. & PICKARD, A. S. 2015a. Validity of EQ-5D-5L in stroke. *Quality of Life Research,* 24**,** 845-50.
GOLICKI, D., NIEWADA, M., KARLINSKA, A., BUCZEK, J., KOBAYASHI, A., JANSSEN, M. F. & PICKARD, A. S. 2015b. Comparing responsiveness of the EQ-5D-5L, EQ-5D-3L and EQ VAS in stroke patients. *Quality of Life Research,* 24**,** 1555-63.
GOMES, M., SOARES, M., DUMVILLE, J., LEWIS, S., TORGERSON, D., BODENHAM, A., GOUGH, M. & WARLOW, C. 2010. Cost-effectiveness analysis of general anaesthesia versus local anaesthesia for carotid surgery (GALA Trial). *British Journal of Surgery,* 97**,** 1218-1225.
GOOSSENS, H., FERECH, M., VANDER STICHELE, R., ELSEVIERS, M. & GROUP, E. P. 2005. Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. *The Lancet,* 365**,** 579-587.
GRACE. *Genomics to combat Resistance against Antibiotics in Community-acquired LRTI in Europe (GRACE) project* [Online]. Available: www.grace-lrti.org [Accessed 27th November 2015].
GRAY, A. M., CLARKE, P. M., WOLSTENHOLME, J. L. & WORDSWORTH, S. 2010. *Applied methods of cost-effectiveness analysis in healthcare,* Oxford University Press.
GRAY, A. M. & WILKINSON, T. 2016. Economic evaluation of healthcare interventions: old and new directions. *Oxford Review of Economic Policy,* 32**,** 102-121.
GREENE, M. E., RADER, K. A., GARELLICK, G., MALCHAU, H., FREIBERG, A. A. & ROLFSON, O. 2015. The EQ-5D-5L Improves on the EQ-5D-3L for Health-related Quality-of-life Assessment in Patients Undergoing Total Hip Arthroplasty. *Clinical Orthopaedics and Related Research.* 473 (11): 3383-90
GREINER, W., CLAES, C., BUSSCHBACH, J. & VON DER SCHULENBURG, J.-M. G. 2005. Validating the EQ-5D with time trade off for the German population. *The European Journal of Health Economics,* 6**,** 124-130.
GREINER, W., WEIJNEN, T., NIEUWENHUIZEN, M., OPPE, S., BADIA, X., BUSSCHBACH, J., BUXTON, M., DOLAN, P., KIND, P. & KRABBE, P. 2003. A single European currency for EQ-5D health states. *The European Journal of Health Economics, formerly: HEPAC,* 4**,** 222-231.
GRIEVE, R., CAIRNS, J. & THOMPSON, S. G. 2010. Improving costing methods in multicentre economic evaluation: the use of multiple imputation for unit costs. *Health Economics,* 19**,** 939-954.

GRIEVE, R., NIXON, R., THOMPSON, S. G. & CAIRNS, J. 2007. Multilevel models for estimating incremental net benefits in multinational studies. *Health Economics,* 16**,** 815-826.
GRIEVE, R., NIXON, R., THOMPSON, S. G. & NORMAND, C. 2005. Using multilevel models for assessing the variability of multinational resource use and cost data. *Health Economics,* 14**,** 185-196.
GRIFFITHS, U. K., LEGOOD, R. & PITT, C. 2016. Comparison of Economic Evaluation Methods Across Low-income, Middle-income and High-income Countries: What are the Differences and Why? *Health Economics*. 25 Suppl 1:29-41
GRYTTEN, J. & SØRENSEN, R. 2001. Type of contract and supplier-induced demand for primary physicians in Norway. *Journal of Health Economics,* 20**,** 379-393.
GUILLEMIN, F., BOMBARDIER, C. & BEATON, D. 1993. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology,* 46**,** 1417-1432.
GYRD-HANSEN, D. 2003. Willingness to pay for a QALY. *Health Economics,* 12**,** 1049-1060.
HANSSON, L.O., CARLSSON, I., HANSSON, E., HOVELIUS, B., SVENSSON, P., TRYDING, N. (1995) Measurement of C-reactive protein and the erythrocyte sedimentation rate in general practice. *Scandinavian Journal of Primary Health Care* 13[1], 39-45.
HAYWOOD, K. L., GARRATT, A. M., LALL, R., SMITH, J. F. & LAMB, S. E. 2008. EuroQol EQ-5D and condition-specific measures of health outcome in women with urinary incontinence: reliability, validity and responsiveness. *Quality of Life Research,* 17**,** 475-483.
HERDMAN, M., GUDEX, C., LLOYD, A., JANSSEN, M., KIND, P., PARKIN, D., BONSEL, G. & BADIA, X. 2011. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research,* 20**,** 1727-36.
HERRMANN, M. ANTIBIOTIC EFFECTIVENESS: NEW CHALLENGES IN NATURAL RESOURCE MANAGEMENT24. Antibiotic Resistance: Implications for Global Health and Novel Intervention Strategies: Workshop Summary, 2010. National Academies Press, 21. ISBN-13: 978-0-309-15611-0
HEYLAND, D. K., KERNERMAN, P., GAFNI, A. & COOK, D. J. 1996. Economic evaluations in the critical care literature: Do they help us improve the efficiency of our unit? *Critical Care Medicine,* 24**,** 1591-1598.
HICKS, J. R. 1939. The foundations of welfare economics. *The Economic Journal***,** 696-712.
HILL, J. C., THOMAS, E., HILL, S., FOSTER, N. E. & VAN DER WINDT, D. A. 2015. Development and Validation of the Keele Musculoskeletal Patient Reported Outcome Measure (MSK-PROM). *PLoS One,* 10**,** e0124557.
HINZ, A., KOHLMANN, T., STOBEL-RICHTER, Y., ZENGER, M. & BRAHLER, E. 2014. The quality of life questionnaire EQ-5D-5L: psychometric properties and normative values for the general German population. *Quality of Life Research,* 23**,** 443-7.
HOCH, J. S., ROCKX, M. A. & KRAHN, A. D. 2006. Using the net benefit regression framework to construct cost-effectiveness acceptability curves: an example using data from a trial of external loop recorders versus Holter monitoring for ambulatory monitoring of. *BMC Health Services Research,* 6**,** 68.
HOLLOWAY, L., HUMPHREY, L., HERON, L., PILLING, C., KITCHEN, H., HOJBJERRE, L., STRANDBERG-LARSEN, M. & HANSEN, B. B. 2014. Patient-reported outcome measures for systemic lupus erythematosus clinical trials: a review

of content validity, face validity and psychometric performance. *Health and Quality of Life Outcomes,* 12**,** 116.
HSIEHCHEN, D., ESPINOZA, M. & HSIEH, A. 2015. The Cooperative Landscape of Multinational Clinical Trials. *PloS one,* 10, e0130930.
HUANG, I.-C., WILLKE, R. J., ATKINSON, M. J., LENDERKING, W. R., FRANGAKIS, C. & WU, A. W. 2007. US and UK versions of the EQ-5D preference weights: Does choice of preference weights make a difference? *Quality of Life Research,* 16**,** 1065-1072.
HUGHES, D., CHARLES, J., DAWOUD, D., EDWARDS, R. T., HOLMES, E., JONES, C., PARHAM, P., PLUMPTON, C., RIDYARD, C. & LLOYD-WILLIAMS, H. 2016. Conducting Economic Evaluations Alongside Randomised Trials: Current Methodological Issues and Novel Approaches. *PharmacoEconomics***,** 1-15.
HURLEY, J. 2000. An overview of the normative economics of the health sector. *Handbook of Health Economics,* 1**,** 55-118.
HUSTED, J. A., COOK, R. J., FAREWELL, V. T. & GLADMAN, D. D. 2000. Methods for assessing responsiveness: a critical review and recommendations. *Journal of Clinical Epidemiology,* 53**,** 459-468.
HUTUBESSY, R., CHISHOLM, D. & EDEJER, T. T. 2003. Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. *Cost-effectiveness and Resource Allocation,* 1**,** 8.
Institute for Quality and Efficiency in Health Care (2015) General Methods Version 4.1 https://www.iqwig.de/download/IQWiG_General_Methods_Version_%204-1.pdf [Accessed 7th October, 2016]
ILO. *ILO labour statistics* [Online]. Available: http://laborsta.ilo.org [Accessed 27th March 2015].
JANSSEN, B. M., OPPE, M., VERSTEEGH, M. M. & STOLK, E. A. 2013a. Introducing the composite time trade-off: a test of feasibility and face validity. *European Journal of Health Economics,* 14 Suppl 1**,** S5-13.
JANSSEN, M. F., PICKARD, A. S., GOLICKI, D., GUDEX, C., NIEWADA, M., SCALONE, L., SWINBURN, P. & BUSSCHBACH, J. 2013b. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multicountry study. *Quality of Life Research,* 22**,** 1717-27.
JANZON, M., LEVIN, L.-Å. & SWAHN, E. 2003. Cost effectiveness of extended treatment with low molecular weight heparin (dalteparin) in unstable coronary artery disease: results from the FRISC II trial. *Heart,* 89**,** 287-292.
JIA, Y. X., CUI, F. Q., LI, L., ZHANG, D. L., ZHANG, G. M., WANG, F. Z., GONG, X. H., ZHENG, H., WU, Z. H., MIAO, N., SUN, X. J., ZHANG, L., LV, J. J. & YANG, F. 2014. Comparison between the EQ-5D-5L and the EQ-5D-3L in patients with hepatitis B. *Quality of Life Research,* 23**,** 2355-63.
JIT, M., BILCKE, J., MANGEN, M.-J. J., SALO, H., MELLIEZ, H., EDMUNDS, W. J., YAZDAN, Y. & BEUTELS, P. 2009. The cost-effectiveness of rotavirus vaccination: Comparative analyses for five European countries and transferability in Europe. *Vaccine,* 27**,** 6121-6128.
JOHANNESSON, M., JONSSON, B., JONSSON, L., KOBELT, G. & ZETHRAEUS, N. 2009. Why should economic evaluations of medical innovations have a societal perspective? *Available at* . https://papers.ssrn.com/sol3/papers2.cfm?abstract_id=2640123 [Accessed on 25th January, 2016]

JOHNSON, J. A., LUO, N., SHAW, J. W., KIND, P. & COONS, S. J. 2005. Valuations of EQ-5D health states: are the United States and United Kingdom different? *Medical Care,* 43**,** 221-228.
KALDOR, N. 1939. Welfare propositions of economics and interpersonal comparisons of utility. *The Economic Journal***,** 549-552.
KCE (Belgian Health Care Knowledge Centre) (2012) Belgian guidelines for economic evaluations and budget impact analyses: second edition https://kce.fgov.be/sites/default/files/page_documents/KCE_183_economic_evaluations_second_edition_Report.pdf [Accessed 2nd October, 2016]
KEELEY, T., AL-JANABI, H., LORGELLY, P. & COAST, J. 2013. A qualitative assessment of the content validity of the ICECAP-A and EQ-5D-5L and their appropriateness for use in health research. *PLoS One,* 8**,** e85287.
KENNEDY, A. D., LEIGH-BROWN, A. P., TORGERSON, D. J., CAMPBELL, J. & GRANT, A. 2002. Resource use data by patient report or hospital records: Do they agree? *BMC Health Services Research,* 2, 1.
KIM, S. H., KIM, H. J., LEE, S. I. & JO, M. W. 2012. Comparing the psychometric properties of the EQ-5D-3L and EQ-5D-5L in cancer patients in Korea. *Quality of Life Research,* 21**,** 1065-73.
KIM, T. H., JO, M. W., LEE, S. I., KIM, S. H. & CHUNG, S. M. 2013. Psychometric properties of the EQ-5D-5L in the general population of South Korea. *Quality of Life Research,* 22**,** 2245-53.
KIND, P. 2003. Guidelines for value sets in economic and non-economic studies using EQ-5D. *The measurement and valuation of health status using EQ-5D: A European perspective.* Springer.
KIND, P., DOLAN, P., GUDEX, C. & WILLIAMS, A. 1998. Variations in population health status: results from a United Kingdom national questionnaire survey. *British Medical Journal,* 316**,** 736-741.
KING, J. T., TSEVAT, J., LAVE, J. R. & ROBERTS, M. S. 2005. Willingness to pay for a quality-adjusted life year: implications for societal health care resource allocation. *Medical Decision Making,* 25**,** 667-677.
KNAPP, M., WINDMEIJER, F., BROWN, J., KONTODIMAS, S., TZIVELEKIS, S., HARO, J. M., RATCLIFFE, M., HONG, J. & NOVICK, D. 2008. Cost-utility analysis of treatment with olanzapine compared with other antipsychotic treatments in patients with schizophrenia in the pan-European SOHO study. *Pharmacoeconomics,* 26**,** 341-358.
KNIES, S., EVERS, S. M., CANDEL, M. J., SEVERENS, J. L. & AMENT, A. J. 2009. Utilities of the EQ-5D. *Pharmacoeconomics,* 27**,** 767-779.
KOLM, S.-C. 2000. A historical introduction to normative economics. *Social Choice and Welfare,* 17, 707-738.
KOLM, P., YUAN, Y., VELEDAR, E., MEHTA, S. R., O'BRIEN, J. A. & WEINTRAUB, W. S. 2007. Cost-effectiveness of clopidogrel in acute coronary syndromes in Canada: a long-term analysis based on the CURE trial. *Canadian Journal of Cardiology,* 23**,** 1037-1042.
KOOPMANSCHAP, M. A., TOUW, K. C. & RUTTEN, F. F. 2001. Analysis of costs and cost-effectiveness in multinational trials. *Health Policy,* 58**,** 175-186.
KUSPINAR, A. & MAYO, N. E. 2014. A review of the psychometric properties of generic utility measures in multiple sclerosis. *Pharmacoeconomics,* 32**,** 759-73.
LAAKKEIDEN HINTALAUTAKUNTA 2013 Preparing a health economic evaluation to be attached to the application for reimbursement status and wholesale price for a medical product

http://www.ispor.org/peguidelines/source/2013Pricing%20Board_Guidance_applications_english.pdf [Accessed on 25th November, 2016]
LAMERS, L. M., MCDONNELL, J., STALMEIER, P. F., KRABBE, P. F. & BUSSCHBACH, J. J. 2006. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. *Health Economics,* 15**,** 1121-1132.
LAMY, A., JÖNSSON, B., WEINTRAUB, W. S., ZHAO, F., CHROLAVICIUS, S., BAKHAI, A., CULLER, S., GAFNI, A., LINDGREN, P. & MAHONEY, E. 2004. The cost-effectiveness of the use of clopidogrel in acute coronary syndromes in five countries based upon the CURE study. *European Journal of Cardiovascular Prevention & Rehabilitation,* 11**,** 460-465.
LAMY, A., YUSUF, S., POGUE, J. & GAFNI, A. 2003. Cost implications of the use of ramipril in high-risk patients based on the Heart Outcomes Prevention Evaluation (HOPE) study. *Circulation,* 107**,** 960-965.
LAXMINARAYAN, R. & BROWN, G. M. 2001. Economics of antibiotic resistance: a theory of optimal use. *Journal of Environmental Economics and Management,* 42**,** 183-206.
LIN, F. J., PICKARD, A. S., KRISHNAN, J. A., JOO, M. J., AU, D. H., CARSON, S. S., GILLESPIE, S., HENDERSON, A. G., LINDENAUER, P. K., MCBURNIE, M. A., MULARSKI, R. A., NAURECKAS, E. T., VOLLMER, W. M., LEE, T. A. & CONSORTIUM, C. 2014. Measuring health-related quality of life in chronic obstructive pulmonary disease: properties of the EQ-5D-5L and PROMIS-43 short form. *BMC Medical Research Methodology,* 14**,** 78.
LINDGREN, B., SEARS, M., CAMPBELL, M., VILLASANTE, C., HUANG, S., LINDH, A., PETERMANN, W., SVENSSON, K., BERGGREN, F. & PAUWELS, R. 2005. Cost-effectiveness of formoterol and salbutamol as asthma reliever medication in Sweden and in Spain. *International Journal of ClinicalPractice,* 59**,** 62-68.
LINDFORS, A., FELTELIUS,N.n LUNDKVIST, J. (2007) Health Economic Evaluation Alongside Clnical Trials: A Review of Study Protocols at the Swedish Medical Products Agency *International Journal of Technology Assessment in Health Car*e 23(3) 392-396
LITTLE, P. 2005. Delayed prescribing of antibiotics for upper respiratory tract infection: with clear guidance to patients and parents it seems to be safe. *British Medical Journal,* 331**,** 301.
LITTLE, P., STUART, B., MOORE, M., COENEN, S., BUTLER, C. C., GODYCKI-CWIRKO, M., MIERZECKI, A., CHLABICZ, S., TORRES, A. & ALMIRALL, J. 2013a. Amoxicillin for acute lower-respiratory-tract infection in primary care when pneumonia is not suspected: a 12-country, randomised, placebo-controlled trial. *The Lancet Infectious Diseases,* 13**,** 123-129.
LITTLE, P., STUART, B., FRANCIS, N., DOUGLAS, E., TONKIN-CRINE, S., ANTHIERENS, S., CALS, J. W., MELBYE, H., SANTER, M. & MOORE, M. 2013b. Effects of internet-based training on antibiotic prescribing rates for acute respiratory-tract infections: a multinational, cluster, randomised, factorial, controlled trial. *The Lancet,* 382**,** 1175-1182.
LÖFDAHL, C.-G., ERICSSON, Å., SVENSSON, K. & ANDREASSON, E. 2005. Cost effectiveness of budesonide/formoterol in a single inhaler for COPD compared with each monocomponent used alone. *Pharmacoeconomics,* 23**,** 365-375.
LÓPEZ-BASTIDA, J., OLIVA, J., ANTOÑANZAS, F., GARCÍA-ALTÉS, A., GISBERT, R., MAR, J. & PUIG-JUNOY, J. 2010. Spanish recommendations on economic evaluation of health technologies. *The European Journal of Health Economics,* 11**,** 513-520.

LORGELLY, P. K., BRIGGS, A. H., WEDEL, H., DUNSELMAN, P., HJALMARSON, Å., KJEKSHUS, J., WAAGSTEIN, F., WIKSTRAND, J., JANOSI, A. & VELDHUISEN, D. J. 2010. An economic evaluation of rosuvastatin treatment in systolic heart failure: evidence from the CORONA trial. *European Journal ofHeart Failure,* 12**,** 66-74.

LUBELL, Y., YEUNG, S., DONDORP, A., DAY, N., NOSTEN, F., TJITRA, E., ABUL FAIZ, M., BIN YUNUS, E., ANSTEY, N. & MISHRA, S. 2009. Cost-effectiveness of artesunate for the treatment of severe malaria. *Tropical Medicine & International Health,* 14**,** 332-337.

MACFARLANE, J., LEWIS, S., MACFARLANE, R. & HOLMES, W. 1997. Contemporary use of antibiotics in 1089 adults presenting with acute lower respiratory tract illness in general practice in the UK: implications for developing management guidelines. *Respiratory Medicine,* 91**,** 427-434.

MACFARLANE, J., MACFARLANE, R., ROSE, D., COLVILLE, A. & GUION, A. 1993. Prospective study of aetiology and outcome of adult lower-respiratory-tract infections in the community. *The Lancet,* 341**,** 511-514.

MAGNELL, C., BROWN, L., MOSKOWITZ, A. & GELIJNS, A. 2005. Chapter 6 Health care evaluation: issues in transferability. *Health Technologies and Decision Making. Paris: OECD.*

MALHOTRA-KUMAR, S., LAMMENS, C., COENEN, S., VAN HERCK, K. & GOOSSENS, H. 2007. Effect of azithromycin and clarithromycin therapy on pharyngeal carriage of macrolide-resistant streptococci in healthy volunteers: a randomised, double-blind, placebo-controlled study. *The Lancet,* 369**,** 482-490.

MANCA, A., LAMBERT, P. C., SCULPHER, M. & RICE, N. 2007. Cost-effectiveness analysis using data from multinational trials: the use of bivariate hierarchical modeling. *Medical Decision Making*. 27(4): 471-490

MANCA, A., RICE, N., SCULPHER, M. J. & BRIGGS, A. H. 2005. Assessing generalisability by location in trial-based cost-effectiveness analysis: the use of multilevel models. *Health Economics,* 14**,** 471-485.

MANCA, A., SCULPHER, M. J. & GOEREE, R. 2010. The analysis of multinational cost-effectiveness data for reimbursement decisions. *Pharmacoeconomics,* 28**,** 1079-1096.

MANCA, A., SCULPHER, M. J., WARD, K. & HILTON, P. 2003. A cost–utility analysis of tension-free vaginal tape versus colposuspension for primary urodynamic stress incontinence. *BJOG: An International Journal of Obstetrics & Gynaecology,* 110**,** 255-262.

MARCOFF, L., ZHANG, Z., ZHANG, W., EWEN, E., JURKOVITZ, C., LEGUET, P., KOLM, P. & WEINTRAUB, W. S. 2009. Cost Effectiveness of Enoxaparin in Acute ST-Segment Elevation Myocardial Infarction: The ExTRACT–TIMI 25 (Enoxaparin and Thrombolysis Reperfusion for Acute Myocardial Infarction Treatment–Thrombolysis In Myocardial Infarction 25) Study. *Journal of the American College of Cardiology,* 54**,** 1271-1279.

MARINA, N., BIELACK, S., WHELAN, J., SMELAND, S., KRAILO, M., SYDES, M., BUTTERFASS-BAHLOUL, T., CALAMINUS, G. & BERNSTEIN, M. 2009. International collaboration is feasible in trials for rare conditions: the EURAMOS experience. *Pediatric and Adolescent Osteosarcoma.* 152: 339-353 Springer.

MARSCHNER, I. C. 2010. Regional differences in multinational clinical trials: anticipating chance variation. *Clinical Trials*. 7(2): 147-56

MARTENS J.D., WINKENS, R.A.G. (2006) Does a joint development and dissemination of multidisciplinary guidelines improve prescribing behaviour: A pre/post study with

concurrent control group and a randomised trial *BMC Health Services Research* 2(6) PP 145
MARTIN, S. C., GAGNON, D. D., ZHANG, L., BOKEMEYER, C., KOOY, M. V. M. & VAN HOUT, B. 2003. Cost-utility analysis of survival with epoetin-alfa versus placebo in stage IV breast cancer. *Pharmacoeconomics,* 21**,** 1153-1169.
MASON, J. M. & MASON, A. R. 2006. The generalisability of pharmacoeconomic studies. *Pharmacoeconomics,* 24**,** 937-945.
MATHEWS, W.C., May, S., (2007) EuroQol (EQ-5D) Measure of Quality of Life Predicts Mortality, Emergency Department Utilization, and Hospital Discharge Rates in HIV-Infected Adults Under Care *Health and Quality of Life Outcomes* 5:5
MATTHEWS, J., ALTMAN, D. G., CAMPBELL, M. & ROYSTON, P. 1990. Analysis of serial measurements in medical research. *British Medical Journal,* 300**,** 230-235.
MCCABE, C. 2009. What is cost-utility analysis. *London: Hayward Medical Communications*.
MCCABE, C., CLAXTON, K. & CULYER, A. J. 2008. The NICE cost-effectiveness threshold. *Pharmacoeconomics,* 26**,** 733-744.
MCINTOSH, E. 2010. *Applied methods of cost-benefit analysis in health care,* Oxford University Press.
MCQUISTON HASLUND, J., ROSBORG DINESEN, M., STERNHAGEN NIELSEN, A. B., LLOR, C. & BJERRUM, L. 2013. Different recommendations for empiric first-choice antibiotic treatment of uncomplicated urinary tract infections in Europe. *Scandinavian Journal of Primary Health Care,* 31, 235-240.
MEINERT, C.L. 2012. Clinical Trials: Design, Conduct and Analysis, Oxford University Press
MENCACI, C., DI SCIASCIO, G., KATZ, P., RIPELLINO C. 2013 Cost-effectiveness evaluation of escitalopram in major depressive disorder in Italy *ClinicoEconomics and Outcomes Research* 5, 87-99
MITTMANN, N., AU, H.-J., TU, D., O'CALLAGHAN, C. J., ISOGAI, P. K., KARAPETIS, C. S., ZALCBERG, J. R., EVANS, W. K., MOORE, M. J. & SIDDIQUI, J. 2009. Prospective cost-effectiveness analysis of cetuximab in metastatic colorectal cancer: evaluation of National Cancer Institute of Canada Clinical Trials Group CO. 17 trial. *Journal of the National Cancer Institute,* 101**,** 1182-1192.
MOONEY, G. and RUSSELL, E. 2003 Equity in health care: the need for a new economics paradigm? In Scott, A., Maynard, A. and Elliott B. (eds) Advances in Health Economics London: John Wiley and Sons Ltd.
MOORE, M., STUART, B., COENEN, S., BUTLER, C. C., GOOSSENS, H., VERHEIJ, T. J. & LITTLE, P. 2014. Amoxicillin for acute lower respiratory tract infection in primary care: subgroup analysis of potential high-risk groups. *British Journal of General Practice,* 64**,** e75-e80.
MORRIS, S., DEVLIN, N. & PARKIN, D. 2007. *Economic analysis in health care,* John Wiley & Sons.
MORRIS, S., DEVLIN, N. & PARKIN, D. 2012. *Economic analysis in health care (Second Edition),* John Wiley & Sons.

MULDER, I., TIJHUIS, M., SMIT, H. A. & KROMHOUT, D. 2001. Smoking cessation and quality of life: the effect of amount of smoking and time since quitting. *Preventive Medicine,* 33**,** 653-660.
MULLIGAN, J. & FOX-RUSHBY, J. 2005. Chapter 18 Transferring cost-effectiveness data across space and time. *Economic Evaluation, J. Fox-Rushby and J. Cairns, Editors*. McGraw-Hill Education (UK)

NASSER, S., VESTENBAEK, U., BERIOT-MATHIOT, A. & POULSEN, P. 2008. Cost-effectiveness of specific immunotherapy with Grazax in allergic rhinitis co-existing with asthma. *Allergy,* 63**,** 1624-1629.
NHG (Netherlands huisartsen genootschap) (2013) Acute Cough M78 https://guidelines.nhg.org/wp-content/uploads/2014/05/summary-card-m78-acute-cough.pdf
NICE. National Institute for Health and Care Excellence. Available: https://www.nice.org.uk/about/what-we-do [Accessed 3rd October 2015].
NICE. 2012. Methods for the development of NICE public health guidance (third edition) (PMG4) https://www.nice.org.uk/process/pmg4/resources/methods-for-the-development-of-nice-public-health-guidance-third-edition-pdf-2007967445701 [Accessed 14th September 2016]
NICE. 2013. *Guide to the methods of technology appraisal 2013* [Online]. National Institute for Health and Care Excellence. Available: https://www.nice.org.uk/article/pmg9/chapter/the-reference-case [Accessed 25th October, 2016].
NICE (2014) Pneumonia Diagnosis and management of community and hospital-acquired pneumonia in adults. NICE clinical guideline 191 https://www.nice.org.uk/guidance/cg191/resources/guidance-pneumonia-pdf
NIXON, J., RICE, S., DRUMMOND, M., BOULENGER, S., ULMANN, P. & DE POUVOURVILLE, G. 2009. Guidelines for completing the EURONHEED transferability information checklists. *The European Journal of Health Economics,* 10**,** 157-165.
NOMA (Norwegian Medicines Agency) (2012) Guidelines on how to conduct pharmacoeconomic analyses. Oslo. https://legemiddelverket.no/Documents/English/Price%20and%20reimbursement/Application%20for%20reimbursement/Pharmacoeconomic%20guidelines%20-%20Norway.pdf [Accessed on 15th October, 2016]
NORMAN, R., CRONIN, P., VINEY, R., KING, M., STREET, D. & RATCLIFFE, J. 2009. International Comparisons in Valuing EQ-5D Health States: A Review and Analysis. *Value in Health,* 12**,** 1194-1200.
O'LEARY, E., DRUMMOND, F. J., GAVIN, A., KINNEAR, H. & SHARP, L. 2015. Psychometric evaluation of the EORTC QLQ-PR25 questionnaire in assessing health-related quality of life in prostate cancer survivors: a curate's egg. *Quaity ofl Life Research,* 24**,** 2219-30.
ODDERSHEDE, L. & PETERSEN, K. D. 2015. Adjustment of foreign EQ-5D-3L utilities can increase their transferability. *ClinicoEconomics and Outcomes Research: CEOR,* 7**,** 629.
OECD. *OECD Purchasing Power Parities (PPPs) for OECD Countries Since 1980* [Online]. Available: http://www.oecd.org [Accessed 25th October, 2016].
OECD. 2012. *OECD. Organisation for Economic Co-operation and Development Purchasing power parities and exchange rates* [Online]. Available: http://stats.oecd.org/index.aspx?queryid=22519 [Accessed 25th October, 2012].
OHINMAA, A., EIJA, H. & SINTONEN, H.1995 Modelling EuroQol values of Finnish adult population. EuroQol plenary meeting, 1995 67-76.
OPPE, M., DEVLIN, N. J. & SZENDE, A. 2007a. *EQ-5D value sets: inventory, comparative review and user guide*, Springer.
OPPE, M., SZENDE, A. & DE CHARRO, F. 2007b. Comparative review of visual analogue scale value sets. *EQ-5D Value Sets.* Springer.

OPPONG, R., COAST, J., HOOD, K., NUTTALL, J., SMITH, R. D. & BUTLER, C. C. 2011a. Resource use and costs of treating acute cough/lower respiratory tract infections in 13 European countries: results and challenges. *The European Journal of Health Economics,* 12**,** 319-329.

OPPONG, R., JIT, M., SMITH, R. D., BUTLER, C. C., MELBYE, H., MÖLSTAD, S. & COAST, J. 2013a. Cost-effectiveness of point-of-care C-reactive protein testing to inform antibiotic prescribing decisions. *British Journal of General Practice,* 63**,** e465-e471.

OPPONG, R., JOWETT, S., NICHOLLS, E., WHITEHURST, D. G., HILL, S., HAMMOND, A., HAY, E. M. & DZIEDZIC, K. 2015a. Joint protection and hand exercises for hand osteoarthritis: an economic evaluation comparing methods for the analysis of factorial trials. *Rheumatology,* 54**,** 876-883.

OPPONG, R., JOWETT, S. & ROBERTS, T. E. 2015b. Economic evaluation alongside multinational studies: a systematic review of empirical studies. *PloS Oone,* 10**,** e0131949.

OPPONG, R., KAAMBWA, B., NUTTALL, J., HOOD, K., SMITH, R. D. & COAST, J. 2013b. The impact of using different tariffs to value EQ-5D health state descriptions: an example from a study of acute cough/lower respiratory tract infections in seven countries. *The European Journal of Health Economics,* 14**,** 197-209.

OPPONG, R., KAAMBWA, B., NUTTALL, J., HOOD, K., SMITH, R. D., COAST, J. & TEAM, G.-S. 2011b. Assessment of the construct validity of the EQ-5D in patients with acute cough/lower respiratory tract infections. *Applied Research in Quality of Life,* 6**,** 411-423.

OPPONG, R., SMITH RICHARD, LITTLE PAUL, VERHEIJ THEO, BUTLER CHRISTOPHER, GOOSSENS HERMAN, COENEN SAMUEL, MOORE MICHAEL, COAST JOANNA 2016. Cost-effectiveness of amoxicillin for LRTI in primary care: An Economic Evaluation Accounting for the Cost of Resistance *British Journal of General Practice*. 66(650) e633-e639

O'SULLIVAN, A.K., THOMPSON, D and DRUMMOND, M.F. (2005) Collection of Health-Economic Data Alongside Clinical Trials: Is There a Future for Piggyback Evaluations *Value In Health* 8(1) PP 67-79

OTT, R. & LONGNECKER, M. 2008. *An introduction to statistical methods and data analysis,* Cengage Learning.

PANG, F. 2002. Design, analysis and presentation of multinational economic studies. *Pharmacoeconomics,* 20**,** 75-90.

PARK, S. & LAKE, E. T. 2005. Multilevel modeling of a clustered continuous outcome: Nurses' work hours and burnout. *Nursing Research,* 54**,** 406.

PATTANAPHESAJ, J. & THAVORNCHAROENSAP, M. 2015. Measurement properties of the EQ-5D-5L compared to EQ-5D-3L in the Thai diabetes patients. *Health and Quality of Life Outcomes,* 13**,** 14.

PETROU, S. & GRAY, A. 2011. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. *British Medical Journal,* 342: d1548

PIANTADOSI, S. 2013 Clinical trials: a methodologic perspective, John Wiley & Sons.

PICKARD, A. S., DE LEON, M. C., KOHLMANN, T., CELLA, D. & ROSENBLOOM, S. 2007. Psychometric comparison of the standard EQ-5D to a 5 level version in cancer patients. *Medical Care,* 45**,** 259-63.

PITT, C., VASSALL, A., TEERAWATTANANON, Y., GRIFFITHS, U. K., GUINNESS, L., WALKER, D., FOSTER, N. & HANSON, K. 2016. Foreword: Health Economic Evaluations in Low-and Middle-income Countries: Methodological Issues and Challenges for Priority Setting. *Health Economics,* 25**,** 1-5.

PRICE, D., HAUGHNEY, J., DUERDEN, M., NICHOLLS, C. & MOSELEY, C. 2002. The cost effectiveness of chlorofluorocarbon-free beclomethasone dipropionate in the treatment of chronic asthma: a cost model based on a 1-year pragmatic, randomised clinical study. *Pharmacoeconomics,* 20**,** 653-664.

PULLENAYEGUM, E. M., CHAN, K. K. & XIE, F. 2015. Quantifying Parameter Uncertainty in EQ-5D-3L Value Sets and Its Impact on Studies That Use the EQ-5D-3L to Measure Health Utility A Bayesian Approach. *Medical Decision Making***,** 0272989X15591966.

RABE-HESKETH, S. & SKRONDAL, A. 2008. *Multilevel and longitudinal modeling using Stata,* STATA press.

RABIN, R. & CHARRO, F. D. 2001. EQ-SD: a measure of health status from the EuroQol Group. *Annals of Medicine,* 33**,** 337-343.

RADEVA, J. I., REED, S. D., KALÓ, Z., KAUF, T. L., CANTU, E., CRETIN, N. & SCHULMAN, K. A. 2005. Economic evaluation of everolimus vs. azathioprine at one year after de novo heart transplantation. *Clinical Transplantation,* 19**,** 122-129.

RAFTERY, J. 2001. NICE: faster access to modern treatments? Analysis of guidance on health technologies. *British Medical Journal,* 323**,** 1300.

RAFTERY, J. 2006. Review of NICE's recommendations, 1999-2005. *British Medical Journal,* 332**,** 1266.

RAFTERY, J. 2014. NICE's Cost-Effectiveness Range: Should it be Lowered? *PharmacoEconomics,* 32**,** 613-615.

RAFTERY, J. P. 2008. Paying for costly pharmaceuticals: regulation of new drugs in Australia, England and New Zealand. *Medical Journal of Australia,* 188**,** 26.

RAIKOU, M., BRIGGS, A., GRAY, A. & MCGUIRE, A. 2000. Centre-specific or average unit costs in multi-centre studies? Some theory and simulation. *Health Economics,* 9**,** 191-198.

RAMSEY, S.D., McINTOSH, M. and SULLIVAN, S.D. 2001 Design for Conducting Cost-Effectiveness Analysis Alongside Clinical Trials *Annual Review of Public Health* 22PP 129-141

RAMSEY, S. D., WILLKE, R. J., GLICK, H., REED, S. D., AUGUSTOVSKI, F., JONSSON, B., BRIGGS, A.H. & SULLIVAN, S. D. 2015. Cost-effectiveness analysis alongside clinical trials II—an ISPOR Good Research Practices Task Force report. *Value in Health,* 18**,** 161-172.

REED, S. D. 2012. How country-specific should a country-specific cost-effectiveness analysis be? *European Heart Journal***,** ehs204.

REED, S. D., ANSTROM, K. J., BAKHAI, A., BRIGGS, A. H., CALIFF, R. M., COHEN, D. J., DRUMMOND, M. F., GLICK, H. A., GNANASAKTHY, A. & HLATKY, M. A. 2005. Conducting economic evaluations alongside multinational clinical trials: toward a research consensus. *American Heart Journal,* 149**,** 434-443.

REED, S. D., FRIEDMAN, J. Y., VELAZQUEZ, E. J., GNANASAKTHY, A., CALIFF, R. M. & SCHULMAN, K. A. 2004a. Multinational economic evaluation of valsartan in patients with chronic heart failure: results from the Valsartan Heart Failure Trial (Val-HeFT). *American Heart Journal,* 148**,** 122-128.

REED, S. D., LAXMINARAYAN, R., BLACK, D. J. & SULLIVAN, S. D. 2002. Economic issues and antibiotic resistance in the community. *Annals of Pharmacotherapy,* 36**,** 148-154.

REED, S. D., RADEVA, J. I., GLENDENNING, G. A., SAAD, F. & SCHULMAN, K. A. 2004b. Cost-effectiveness of zoledronic acid for the prevention of skeletal complications in patients with prostate cancer. *The Journal of Urology,* 171**,** 1537-1542.

REINHOLD, T., BRÜGGENJÜRGEN, B., SCHLANDER, M., ROSENFELD, S., HESSEL, F. & WILLICH, S. N. 2010. Economic analysis based on multinational studies: methods for adapting findings to national contexts. *Journal of Public Health,* 18**,** 327-335.

RICHTER T.A 2014. Clinical research: A globalized network *PLoS ONE* 9(12): e115063. doi:10.1371/journal.pone.0115063

RIDYARD, C.H., HUGHES D.A. 2010. Methods for the collection of resource use data within clinical trials: A systematic review of studies funded by the UK health technology assessment program *Value in Health* 13(8) 867-872.

RIDYARD, C. H., HUGHES, D. A. & TEAM, D. 2012. Development of a database of instruments for resource-use measurement: purpose, feasibility, and design. *Value in Health*, 15, 650-655.

RIVERO-ARIAS, O. & GRAY, A. 2010. The Multinational Nature of Cost-Effectiveness Analyses Alongside Multinational Clinical Trials. *Value in Health,* 13**,** 34-41.

ROBERTS, T., HENDERSON, J., MUGFORD, M., BRICKER, L., NEILSON, J. & GARCIA, J. 2002. Antenatal ultrasound screening for fetal abnormalities: a systematic review of studies of cost and cost effectiveness. *British Journal of Obstettrics and GynaecologyG: An International Journal of Obstetrics & Gynaecology,* 109**,** 44-56.

RURIK, I. 2007. Primary care diabetes in Hungary. *Primary Care Diabetes,* 1**,** 177-179.

RUTTEN-VAN MÖLKEN, M. P., VAN NOOTEN, F. E., LINDEMANN, M., CAESER, M. & CALVERLEY, P. M. 2007. A 1-year prospective cost-effectiveness analysis of roflumilast for the treatment of patients with severe chronic obstructive pulmonary disease. *Pharmacoeconomics,* 25**,** 695-711.

RYAN, M., WATSON, V. & AMAYA-AMAYA, M. 2003. Methodological issues in the monetary valuation of benefits in healthcare. *Expert Review of Pharmacoeconomics and Outcomes Research* 3(6): 717-27

SAGAN, A., PANTELI, D., BORKOWSKI, W., DMOWSKI, M., DOMANSKI, F., CZYZEWSKI, M., GORYNSKI, P., KARPACKA, D., KIERSZTYN, E. & KOWALSKA, I. 2010. Poland health system review. *Health Systems in Transition,* 13, 1-193.

SAKTHONG, P., CHAROENVISUTHIWONGS, R. & SHABUNTHOM, R. 2008. A comparison of EQ-5D index scores using the UK, US, and Japan preference weights in a Thai sample with type 2 diabetes. *Health and Quality of Life Outcomes,* 6**,** 71.

SAKTHONG, P., SONSA-ARDJIT, N., SUKARNJANASET, P. & MUNPAN, W. 2015. Psychometric properties of the EQ-5D-5L in Thai patients with chronic diseases. *Quality of Life Research*. 24(12): 3015-22

SANGHERA, S., FREW, E., GUPTA, J. K., KAI, J. & ROBERTS, T. E. 2015. Exploring the Use of Cost-Benefit Analysis to Compare Pharmaceutical Treatments for Menorrhagia. *PharmacoEconomics***,** 1-9.

SANTÉ, H. A. D. 2012. Choices in methods for economic evaluation. *Public Health Assessment Haute Autorité de Santé*.

SAUNDERS, C. & MARSDEN, D. 2014. *Pay Inequalities in the European Community: Butterworths European Studies*, Butterworth-Heinemann.

SCALONE, L., CIAMPICHINI, R., FAGIUOLI, S., GARDINI, I., FUSCO, F., GAETA, L., DEL PRETE, A., CESANA, G. & MANTOVANI, L. G. 2013. Comparing the performance of the standard EQ-5D 3L with the new version EQ-5D 5L in patients with chronic hepatic diseases. *Quality of Life Research,* 22**,** 1707-16.

SCHULMAN, K., BURKE, J., DRUMMOND, M., DAVIES, L., CARLSSON, P., GRUGER, J., HARRIS, A., LUCIONI, C., GISBERT, R. & LLANA, T. 1998.

Resource costing for multinational neurologic clinical trials: methods and results. *Health Economics,* 7**,** 629-638.
SCOTT, J. G., COHEN, D., DICICCO-BLOOM, B., ORZANO, A. J., JAEN, C. R. & CRABTREE, B. F. 2001. Antibiotic use in acute respiratory infections and the ways patients pressure physicians for a prescription. *Journal of Family Practice,* 50**,** 853-858.
SEN, A. 1979. Personal utilities and public judgements: or what's wrong with welfare economics. *The Economic Journal,* 537-558.
SEN, A. 1980. *Equality of what?* The Tanner Lecture on Human Values http://tannerlectures.utah.edu/_documents/a-to-z/s/sen80.pdf
SKOUPA, J., ANNEMANS, L., HAJEK P. (2014) Health economic data requirements and availability in the European Union: results of a survey among 10 European countries *Value in Health Regional Issues,* 4C 53-57
SCULPHER, M. 2001. The role and estimation of productivity costs in economic evaluation. *Economic evaluation in health care: Merging theory with practice***,** 94-112.
SHILLCUTT, M. S. D., WALKER, D. G., GOODMAN, C. A. & MILLS, A. J. 2009. Cost effectiveness in low-and middle-income countries. *Pharmacoeconomics,* 27**,** 903-917.
SIGN. 2014. *Community management of lower respiratory tract infections in adults* [Online]. Scottish Intercollegiete Guidance Network. Available: http://www.sign.ac.uk/guidelines/fulltext/59/section1.html [Accessed 21st May, 2013].
SIMON, J., GRAY, A. & DULEY, L. 2006. Cost-effectiveness of prophylactic magnesium sulphate for 9996 women with pre-eclampsia from 33 countries: economic evaluation of the Magpie Trial. *British Journal of Obstetrics and Gynaecology: An International Journal of Obstetrics & Gynaecology,* 113**,** 144-151.
SIMPSON, S.A., BUTLER, C.C., HOOD, K. (2009) Stemming the tide of antibiotic resistance (STAR): A protocol for a trial of a complex intervention addressing the why and how of appropriate prescribing in general practice *BMC Family Practice* 10:20 doi:10.1186/1471-2296-10-20
SMITH, R. & COAST, J. 2013. The true cost of antimicrobial resistance. *British Medical Journal,* 346: f1493
ŠMIT, R. 2012. Cost-effectiveness of tick-borne encephalitis vaccination in Slovenian adults. *Vaccine*, 30, 6301-6306.
SPÄTH, H.-M., CARRERE, M.-O., FERVERS, B. & PHILIP, T. 1999. Analysis of the eligibility of published economic evaluations for transfer to a given health care system: methodological approach and application to the French health care system. *Health Policy,* 49**,** 161-177.
SPSS Inc. Released 2007. SPSS for Windows, Version 16.0. Chicago, SPSS Inc.
STATACorp. STATA multiple imputation reference manual. Release 12.0. 2011
STATACorpLP.Stata/IC10.1 for windows. 2010
STATACorpLP.Stata/IC 11.1 for Windows. 2011
STATACorp. LP. Stata/IC 12.1 for Windows. 2012
STANTON, N., HOOD, K., KELLY, M. J., NUTTALL, J., GILLESPIE, D., VERHEIJ, T., LITTLE, P., GODYCKI-CWIRKO, M., GOOSSENS, H. & BUTLER, C. C. 2010. Are smokers with acute cough in primary care prescribed antibiotics more often, and to what benefit? An observational study in 13 European countries. *European Respiratory Journal,* 35**,** 761-767.
STEENBERGEN, M. R. & JONES, B. S. 2002. Modeling multilevel data structures. *American Journal of Political Science***,** 218-237.
STREINER, D. L., NORMAN, G. R. & CAIRNEY, J. 2003

Health measurement scales: a practical guide to their development and use, Oxford university press.
SUGDEN, R., WILLIAMS, A. 1978 The principles of practical cost-benefit analysis. Oxford university press. Oxford
SULLIVAN, S. D., BUXTON, M., ANDERSSON, L. F., LAMM, C. J., LILJAS, B., CHEN, Y. Z., PAUWELS, R. A. & WEISS, K. B. 2003. Cost-effectiveness analysis of early intervention with budesonide in mild persistent asthma. *Journal of Allergy and Clinical Immunology,* 112**,** 1229-1236.
SWAN, J. S., HUR, C., LEE, P., MOTAZEDI, T. & DONELAN, K. 2013. Responsiveness of the testing morbidities index in colonoscopy. *Value in Health,* 16**,** 1046-53.
SZENDE, Á., MOGYOROSY, Z., MUSZBEK, N., NAGY, J., PALLOS, G. & DÓZSA, C. 2002. Methodological guidelines for conducting economic evaluation of healthcare interventions in Hungary: a Hungarian proposal for methodology standards. *The European Journal of Health Economics,* 3**,** 196-202.
SZENDE, A., OPPE, M. & DE CHARRO, F. 2007. Comparative review of time trade-off value sets. *EQ-5D Value Sets.* Springer.
TAN, S. S., BOUWMANS, C. A., RUTTEN, F. F. & HAKKAART-VAN ROIJEN, L. 2012. Update of the Dutch manual for costing in economic evaluations. *International Journal of Technology Assessment in Health Care,* 28, 152-158.
THE EUROQOLGROUP 1990. EuroQol-a new facility for the measurement of health-related quality of life. *Health Policy,* 16**,** 199-208.
THOMPSON, S. G., NIXON, R. M. & GRIEVE, R. 2006. Addressing the issues that arise in analysing multicentre cost data, with application to a multinational study. *Journal of Health Economics,* 25**,** 1015-1028.
TORGERSON, D., RYAN, M. & RATCLIFFE, J. 1995. Economics in sample size determination for clinical trials. *QJM,* 88**,** 517-521.
TORRANCE, G. W. 1986. Measurement of health state utilities for economic appraisal: a review. *Journal of Health Economics,* 5**,** 1-30.
TORRANCE, G. W., BLAKER, D., DETSKY, A., KENNEDY, W., SCHUBERT, F., MENON, D., TUGWELL, P., KONCHAK, R., HUBBARD, E. & FIRESTONE, T. 1996. Canadian guidelines for economic evaluation of pharmaceuticals. *Pharmacoeconomics,* 9**,** 535-559.
TORTI, F. M., REED, S. D. & SCHULMAN, K. A. 2006. Analytic considerations in economic evaluations of multinational cardiovascular clinical trials. *Value in Health,* 9**,** 281-291.
TRAN, B. X., OHINMAA, A. & NGUYEN, L. T. 2012. Quality of life profile and psychometric properties of the EQ-5D-5L in HIV/AIDS patients. *Health and Quality of Life Outcomes,* 10**,** 132.
TSUCHIYA, A., WILLIAMS, A 2001 Welfare economics and economic evaluation. In: Drummond M, McGuire A, editors. Economic Evaluation in Health Care Merging Theory With Practice Oxford: Oxford University Press.
TSUCHIYA, A., IKEDA, S., IKEGAMI, N., NISHIMURA, S., SAKAI, I., FUKUDA, T., HAMASHIMA, C., HISASHIGE, A. & TAMURA, M. 2002. Estimating an EQ-5D population value set: the case of Japan. *Health Economics,* 11**,** 341-353.
VALLEJO-TORRES, L., GARCÍA-LORENZO, B. & SERRANO-AGUILAR, P. 2016. Estimating a cost-effectiveness threshold for the Spanish NHS. FEDEA.
VAN dER MEER V., NEVEN, A.K., VAN DEN BROEK, P.J., ASSENDELFT, W.J.J. 2005 Diagnostic value of C reactive protein in infections of the lower respiratory tract: systematic review. *British Medical Journal* 331[7507], 26. 2005.

VAN DEN HOUT, W. 2010. The value of productivity: human-capital versus friction-cost method. *Annals of the Rheumatic Diseases,* 69**,** i89-i91.
VAN DUIJN, H. J., KUYVENHOVEN, M. M., BUTLER, C. C., COENEN, S. & VERHEIJ, T. J. 2005. Variation in outpatient antibiotic use in three European countries: exploration of possible determinants. *European Journal of General Practice,* 11**,** 139-140.
VEMER, P. & RUTTEN-VAN MÖLKEN, M. P. 2013. The road not taken: transferability issues in multinational trials. *Pharmacoeconomics,* 31**,** 863-876.
VENMANS, L. M., BONT, J., GORTER, K. J., VERHEIJ, T. J., RUTTEN, G. E. & HAK, E. 2008. Prediction of complicated lower respiratory tract infections in older patients with diabetes. *British Journal of General Practice,* 58**,** 564-568.
VERHEIJ, T., KAPTEIN, A. & MULDER, J. 1989. Acute bronchitis: aetiology, symptoms and treatment. *Family Practice,* 6**,** 66-69.
VERSTEEGH, M. M., VERMEULEN, K. M., EVERS, S. M., DE WIT, G. A., PRENGER, R. & STOLK, E. A. 2016. Dutch Tariff for the Five-Level Version of EQ-5D. *Value in Health*. 19(4): 343-52
VIRK, K. P. 2008. Global clinical trials: the challenge of language. *Drug Information. Assoc,* 44**,** 21-23.
WADE, A. G., FERNÁNDEZ, J.-L., FRANÇOIS, C., HANSEN, K., DANCHENKO, N. & DESPIEGE, N. 2008. Escitalopram and Duloxetine in Major Depressive Disorder. *Pharmacoeconomics,* 26**,** 969-981.
WEINSTEIN, M., SIEGEL, J., GOLD, M., KAMLET, M. & RUSSELL, L. 1996. Cost-effectiveness in health and medicine. *New York: Oxford University,* 55.
WEINSTEIN, M. C. & MANNING, W. G. 1997. Theoretical issues in cost-effectiveness analysis. *Journal of Health Economics,* 16**,** 121-128.
WELSH, R. C., SAURIOL, L., ZHANG, Z., KOLM, P., WEINTRAUB, W. S. & THEROUX, P. 2009. Cost-effectiveness of enoxaparin compared with unfractionated heparin in ST elevation myocardial infarction patients undergoing pharmacological reperfusion: A Canadian analysis of the Enoxaparin and Thrombolysis Reperfusion for Acute Myocardial Infarction Treatment–Thrombolysis in Myocardial Infarction (ExTRACT-TIMI) 25 trial. *Canadian Journal of Cardiology,* 25**,** e399-e405.
WELTE, R., FEENSTRA, T., JAGER, H. & LEIDL, R. 2004. A decision chart for assessing and improving the transferability of economic evaluation results between countries. *Pharmacoeconomics,* 22**,** 857-876.
WELTE, T., TORRES, A. & NATHWANI, D. 2010. Clinical and economic burden of community-acquired pneumonia among adults in Europe. *Thorax,* 67(1): 71-9.
WHITE, I. R., ROYSTON, P. & WOOD, A. M. 2011. Multiple imputation using chained equations: issues and guidance for practice. *Statistics in Medicine,* 30, 377-399.
WHITEHEAD, S. J. & ALI, S. 2010. Health outcomes in economic evaluation: the QALY and utilities. *British Medical Bulletin,* 96**,** 5-21.
WHITEHURST, D. & BRYAN, S. 2013. Trial-based clinical and economic analyses: the unhelpful quest for conformity. *Trials,* 14**,** 421.
WHO. *Clinical Trials* [Online]. World Health Organisation. Available: http://www.who.int/topics/clinical_trials/en/ [Accessed 12th January 2016].
WHO. *WHO-CHOICE. Choosing interventions that are cost-effective* [Online]. Available: http://www.who.int/choice/en/ [Accessed 24th October 2015].
WHO 1996. Investing in health research and development: report of the ad hoc committee on health research relating to future intervention options.

http://www.who.int/tdr/publications/tdr-research-publications/investing-in-health/en/ [Accessed 12th January 2016]
WHO. 2002. *Antimicrobial Resistance* [Online]. World Health Organisation. Available: http://www.who.int/mediacentre/factsheets/fs194/en/ [Accessed 25th October, 2016].
WILD, D., EREMENCO, S., MEAR, I., MARTIN, M., HOUCHIN, C., GAWLICKI, M., HAREENDRAN, A., WIKLUND, I., CHONG, L. Y. & VON MALTZAHN, R. 2009. Multinational Trials—Recommendations on the Translations Required, Approaches to Using the Same Language in Different Countries, and the Approaches to Support Pooling the Data: The ISPOR Patient-Reported Outcomes Translation and Linguistic Validation Good Research Practices Task Force Report. *Value in Health,* 12**,** 430-440.
WILLAN, A. R., GOEREE, R., PULLENAYEGUM, E. M., MCBURNEY, C. & BLACKHOUSE, G. 2006. Economic evaluation of rivastigmine in patients with Parkinson's disease dementia. *Pharmacoeconomics,* 24**,** 93-106.
WILLIAMS, J. V., HARRIS, P. A., TOLLEFSON, S. J., HALBURNT-RUSH, L. L., PINGSTERHAUS, J. M., EDWARDS, K. M., WRIGHT, P. F. & CROWE JR, J. E. 2004. Human metapneumovirus and lower respiratory tract disease in otherwise healthy infants and children. *New England Journal of Medicine,* 350**,** 443-450.
WILLKE, R. J., GLICK, H. A., POLSKY, D. & SCHULMAN, K. 1998. Estimating country-specific cost-effectiveness from multinational clinical trials. *Health Economics,* 7**,** 481-493.
WORDSWORTH, S. & LUDBROOK, A. 2005. Comparing costing results in across country economic evaluations: the use of technology specific purchasing power parities. *Health Economics,* 14**,** 93-99.
WORDSWORTH, S., LUDBROOK, A., CASKEY, F. & MACLEOD, A. 2005. Collecting unit cost data in multicentre studies. *The European Journal of Health Economics,* 6**,** 38-44.
WORLD BANK. *The World Bank Consumer Price Index. 2013* [Online]. Available: http://data.worldbank.org/indicator/FP.CPI.TOTL?page=6 [Accessed 23rd February, 2015].
WORLD BANK 1993. World development report. Washington DC: World Bank. https://openknowledge.worldbank.org/bitstream/handle/10986/5976/9780195208900_fm.pdf [Accessed, 20th December 2016]
YARDLEY, L., DOUGLAS, E., ANTHIERENS, S., TONKIN-CRINE, S., O'REILLY, G., STUART, B., GERAGHTY, A. W., ARDEN-CLOSE, E., VAN DER VELDEN, A. W. & GOOSENS, H. 2013. Evaluation of a web-based intervention to reduce antibiotic prescribing for LRTI in six European countries: quantitative process analysis of the GRACE/INTRO randomised controlled trial. *Implementation Science,* 8**,** 134.
YUSUF, S., COLLINS, R. & PETO, R. 1984. Why do we need some large, simple randomized trials? *Statistics in Medicine,* 3**,** 409-420.
ZARATE, V. 2007. DALYs and QALYs in developing countries. *Health Affairs,* 26**,** 1197-1198.
ZHAO, F.-L., XIE, F., HU, H. & LI, S.-C. 2013. Transferability of indirect cost of chronic disease: a systematic review and meta-analysis. *Pharmacoeconomics,* 31**,** 501-508.
ZWART-VAN RIJKOM, J. E., LEUFKENS, H. G., BUSSCHBACH, J. J., BROEKMANS, A. W. & RUTTEN, F. F. 2000. Differences in attitudes, knowledge and use of economic evaluations in decision-making in the Netherlands. *Pharmacoeconomics,* 18**,** 149-160.