

**Health economic aspects in the management
of Bipolar disorder**

Anees Ahmed Abdul Pari

Oriel College

University of

Oxford

Michaelmas Term

2016

D.Phil. Thesis

Acknowledgments

The work presented in this thesis was performed at the Oxford University, within the Health Economics Research Centre, Oxford, UK. I realised during my work that there was a growing interest from researchers and policy makers for health economic research of mental health interventions which only reassured me of the importance and impact this thesis may have in my chosen topic area. The past few years have been a truly inspiring and rewarding journey which resulted in this thesis that finally came to realisation Firstly, thanks to OXTEXT Research Program and participants of this study and many more people who are living with Bipolar disorder. I sincerely hope that the outcomes of this research would contribute in some way to attract the attention of policy makers and improve the functional and social life of individuals with mental illness.

I am indebted to many people who have helped me in various ways. I would like to thank my academic supervisors Professor Judit Simon and Dr Jane Wolstenholme, and clinical supervisor Professor John Geddes for their time and input in my work, for constructive feedback and challenging my thinking and working process which helped me to improve as an academic researcher.

Words cannot express my thanks and appreciation for my Public Health Training Program Director, Professor Premila Webster who has always been supportive and gave me confidence that I can work on my training and DPhil simultaneously. She has encouraged me to work harder and always wishes me well.

The OXTEXT Research Program not only sponsored my DPhil but also provided the data for my work. I want to thank the Principal Investigator, Professor Guy Goodwin for granting the access to data used in my research. I want to thank those who gave their time to fine-tune my research question and bringing my thesis back to track including Professor Sasha Shepperd, Professor Alastair Gray, and Dr. Kate Saunders. I also owe my most sincere gratitude to the reviewers of this thesis, Professor Raymond Fitzpatrick and Professor Sarah Byford; you are inspiring and I wish there were more people like you who could selflessly encourage and motivate younger generation of researchers.

I would like to thank my colleagues at the Health Economics Research Centre and training scheme for being wonderful friends who always willing to listen and help. I want to thank them not only for providing an invaluable feedback on my chapters but also for moral and emotional support. I would like to mention few here in particular: Filipa, Larry (and Roopes), Francesco, Seamus, James, Ines, Apo, Jose, Oliver, Liz, Richeal, Ramon, Iryna, Helen, Bobby, Mara, Mandy, Stefan (and Kluzeks), Anant, Vivek, Anup, Gracia, Stella, Uy, Manisha, Sarah, Mohit (and Sharmas), Sam, Jo J, Lucy, Pavlina, Bucchi, Venkat, Rupam, Malini, Mani, BBK, Barbara, and Luke.

I want to thank my Mom and Dad; They brought me into this world, raised me, supported me, and loved me. They taught the importance of good education, hard work, and always encouraged me to strive to reach my goals. Without their guidance, support and love I could not have even started this journey and achieve what I have made so far. Special thanks go out to my family including my sweet sister Fathima, Sakul, Afzal, Sahana, brother Ashik, Sakeena, and Basia's family. I would also like to extend my heartfelt thanks to all my closest friends who have grown young with me and are part of my life; you know who you are and I wish I had the space to thank personally

Finally, this journey could not have been completed without the love, care and support of my wife Basia who also showed understanding and patience during this busy and stressful period in our life. You believed in me and encouraged me as we shared many of the uncertainties, challenges and sacrifices necessary for completing this DPhil. Thank you, you are simply the best! I look forward to spending relaxing time together. Thanks again for bearing my tantrums, and moodiness 😊.

I want to dedicate this work to our precious daughter Aaliya; whose arrival motivated me to complete this thesis – daddy loves you and from now on we will have lots of fun together.

Publications related to this thesis

- a) Pari AA, Simon J, Wolstenholme J, Geddes J, Goodwin G (2015) Economic burden of bipolar disorder in the United Kingdom: a patient-level exploratory analysis. *Bipolar Disorders*, 17:138.
- b) Pari, Abdul, et al. "Economic evaluations in bipolar disorder: a systematic review and critical appraisal." *Bipolar disorders* (2014).

Health economic aspects in the management of Bipolar disorder

Anees Ahmed Abdul Pari, Oriel College.

D.Phil., Michaelmas Term 2016

Bipolar disorder (BD) is one of the leading causes of disability worldwide and has a detrimental impact on health-related quality of life (HRQoL), and personal and social functioning. Despite this, there is insufficient knowledge of the costs, HRQoL implications relevant to BD, and the cost effectiveness of current treatments for BD in the UK. The aim of this thesis is to inform decisions about local and national service provision by applying a variety of health economic tools to build an economic case for BD.

First, economic evaluations of BD management strategies are systematically reviewed. A cost-of-illness study is then conducted to estimate the societal burden of BD in the UK and explore the factors that drive variations in these costs. The appropriateness of applying the EQ-5D-3L outcome measure in BD is assessed and the feasibility of mapping disease-specific measures to the EQ-5D-3L is explored. Finally, a cost-utility analysis (CUA) is conducted to bring together evidence on the costs and outcomes associated with alternative psychological interventions in BD management.

This thesis makes important contributions to multiple research domains, informing the allocation of scarce healthcare resources in this context. There is a sheer dearth of evidence on cost-effectiveness strategies for the long-term management of BD in the UK, especially the evidence for psychological therapies is limited. The annual societal costs associated with BD in the UK are estimated to be £5.14 billion, demonstrating the significant economic burden associated with this disease. The EQ-5D-3L instrument is found to be useful in measuring HRQoL in BD patients who predominantly experience depressive symptoms, but is not sensitive enough to detect changes in individuals with mania. More psychometric evidence is therefore required before this instrument can be widely applied in economic evaluations of BD-related interventions. Finally, the CUA indicates that a novel structured psychoeducation intervention in individuals on remote mood monitoring in the UK is not cost-effective.

Table of Contents

Acknowledgments	2
Publications related to this thesis.....	4
Abstract	5
Table of Contents	6
List of tables	12
List of Figures	14
List of abbreviations.....	16
Chapter 1: Introduction.....	17
1.1 Introduction to the thesis	17
1.2 Applied study: the OXTEXT research programme	22
1.3 Aims and objective of the thesis.....	26
1.4 Outline of the thesis.....	27
1.5 Contribution to the literature.....	30
1.6 Conclusion	33
Chapter 2: Background on epidemiology and burden of BD	34
2.1 Introduction	34
2.2 Disease area.....	34
2.3 Aetiology	37
2.4 Epidemiology	38
2.5 Psychiatric co-morbidities	39
2.6 Medical co-morbidities	40
2.7 Course of illness	41
2.8 Mortality	43
2.9 Existing management strategies	44
2.10 Monitoring strategies.....	47
2.11 Disease burden, and functional impairment.....	48
2.12 Economic cost and quality of life	52
2.13 Summary and conclusion.....	54
Chapter 3: Health Economics in mental health	56
3.1 Introduction	56
3.2 Background to economics and BD	56

3.3 Cost-of-illness (COI) studies	59
3.3.1 Nomenclature of COI studies	60
3.3.2 Limitations of cost-of-illness studies	62
3.4 Main elements of an Economic Evaluation	62
3.4.1 Comparators	63
3.4.2 Perspective	63
3.4.3 Costs	65
3.4.4 Measurement of Health Outcomes	67
3.4.5 HRQoL Measures Specific to BD	70
3.4.6 Utility measurement	71
3.4.7 Quality Adjusted Life Years (QALYs)	71
3.4.8 Incremental Cost Effectiveness Ratio (ICER)	73
3.4.9 Timeframe and discounting	75
3.4.10 Uncertainty analysis	75
3.5 Types of economic evaluation	78
3.5.1 Cost-consequences analysis	78
3.5.2 Cost-minimisation analysis	79
3.5.3 Cost-effectiveness analysis	79
3.5.4 Cost-utility analysis	80
3.5.5 Cost-benefit analysis	80
3.6 Frameworks for conducting economic evaluation	81
3.6.1 Trial-based economic evaluations	81
3.6.2 Decision analytical models	82
3.7 Discussion and conclusion	83
Chapter 4: Economic evaluations in bipolar disorder: A systematic review and critical appraisal	85
4.1 Introduction	85
4.2. Methodology	87
4.2.1. Search strategy	87
4.2.2 Study selection criteria	88
4.2.3 Data extraction and quality assessment	89
4.3 Results	90
4.4 Trial-based economic evaluations	106
4.4.1 Study design	106

4.4.2 Data: Costs	107
4.4.3 Data: Outcomes.....	108
4.5 Analysis and interpretation of results	108
4.5.1 Overall quality scores	109
4.6 Model-based economic evaluations.....	111
4.6.1 Model structure	112
4.6.2 Data	116
4.6.3 Model uncertainty	118
4.7 Overall quality of models.....	119
4.8 Discussion.....	120
4.8.1 Trial-based economic evaluations	123
4.8.2 Model-based economic evaluations.....	125
4.9 Conclusions	127
Chapter 5: Societal costs of Bipolar disorder in the UK.....	129
5.1 Introduction	129
5.2 Current evidence.....	130
5.3 Methodology	133
5.3.1 Patient Cohort.....	133
5.3.2 Data collection	134
5.3.3 Cost estimation	135
5.3.4 Statistical analysis.....	138
5.3.5 Prevalence estimates	140
5.3.6 Cost of illness estimates:	142
5.3.7 Sensitivity analyses	142
5.4 Results	143
5.4.1 Study population	143
5.4.2 Costs of BD.....	145
5.4.3 Prevalence estimates	152
5.5 Economic burden of BD in the UK	156
5.6 Sensitivity analysis.....	157
5.6.1 Univariate.....	157
5.6.2 Complete case analysis.....	158
5.6.3 Probability Sensitivity Analysis (PSA)	158

5.7 Determinants of total BD costs.....	159
5.8 Discussion.....	161
5.8.1 Comparison with other studies estimating the burden of BD in the UK:..	162
5.8.2 The relationship between clinical factors and economic costs.....	166
5.8.3 Strengths and limitations.....	167
5.9 Conclusion	170
Chapter 6: Using preference based measures in bipolar disorder: is the EQ-5D-3L appropriate?	172
6.1 Introduction	172
6.2 EQ-5D-3L and mental health	173
6.3 Methods	176
6.3.1 Participants and study design	176
6.3.2 Outcome instruments.....	178
6.4 Statistical methods.....	179
6.4.1 Appropriateness of the EQ-5D-3L.....	179
6.4.2 Reliability	180
6.4.3 Validity	182
6.4.4 Responsiveness	184
6.5 Feasibility of mapping EQ-5D-3L from disease-specific instruments in BD ..	185
6.5.1 Exploratory factor analyses (EFA) and confirmatory factor analysis (CFA)	186
6.6 Results	193
6.6.1 Descriptive statistics of the EQ-5D item distribution	194
6.6.2 Test-retest reliability.....	196
6.6.3 Convergent validity	198
6.6.4 Known-group validity	199
6.6.5 Responsiveness	204
6.6.6 Exploratory factor analysis (EFA)	207
6.6.7 Confirmatory factor analysis (CFA).....	209
6.7 Discussion.....	211
6.7.1 Appropriateness of the self-rated EQ-5D-3L in people with Bipolar disorders	212
6.7.2 Reliability	213

6.7.3 Construct validity as reflected by convergent and Known-group validity (KGV).....	214
6.7.4 Responsiveness	215
6.7.5 Feasibility of ‘mapping’ disease-specific instruments used in Bipolar disorder onto EQ-5D-3L.....	219
6.7.6 Evidence from other studies that explored appropriateness of the EQ-5D-3L in BD.....	219
6.7.7 Strengths and limitations:.....	220
6.8 Conclusion	223
Chapter 7: An economic evaluation of structured psychoeducation program in BD patients using electronic mood monitoring	226
7.1 Introduction	226
7.2 The OXTEXT-6 study.....	229
7.2.1 Trial design	229
7.3 Interventions	232
7.4 Outcome measures.....	233
7.4.1 Clinical Outcomes.....	233
7.4.2 Economic outcomes	234
7.5 Economic evaluation.....	235
7.6 Statistical methods.....	238
7.6.1 Missing data.....	239
7.6.2 Sensitivity analyses	240
7.7 Results.....	241
7.7.1 Resource use and costs	243
7.7.2 Outcomes	249
7.7.3 Cost-utility analysis	251
7.7.4 Sensitivity Analyses	255
7.8 Discussion.....	256
7.8.1 Strengths and Limitations	258
7.8.2 Existing literature	262
7.9 Conclusion	263
Chapter 8: Discussion and conclusions.....	265
8.1 Introduction	265
8.2 Overview of thesis findings	265

8.3 Challenges of this research.....	274
8.4 Contribution to the literature.....	278
8.5 Policy context and implications	282
8.6 Future research.....	286
8.7 Concluding remarks	288
References:.....	290
Appendix 1: Search strategies used to identify references which summarise the economic evaluations of management strategies in Bipolar disorder in Chapter 4	307
Appendix 2: Data extraction template for collecting data on studies included in Chapter 4.....	326
Appendix 3: Centre for review and dissemination (CRD) checklist for assessing economic evaluations(104, 321)*	329
Appendix 4: Review of guidelines for good practice in decision-analytic modelling in health technology assessment; Philips Checklist (322).....	330
Appendix 5: list of excluded studies	332
Appendix 6: Assessing the quality of Trial-based economic evaluations in Bipolar disorder using centre of review and dissemination's checklist	333
Appendix 7: Assessing the quality of Modelling-based economic evaluations of management strategies in Bipolar disorder using Philips's checklist.....	335
Appendix 8: Correlation matrix between the ASRM (mania), EQ-5D-3L, and VAS	339
Appendix 9: Horn's parallel analysis.....	340
Appendix 10: Characteristics of studies exploring appropriateness of EQ-5D-3L in Bipolar disorder. Partly adapted and updated from Brazier et al 2014.	341
Appendix 11: Sample Questionnaire used in the OXTEXT-6 Trial-based economic evaluation in Bipolar disorder	343
Appendix 12: Services included and unit costs (£) for 2012-2013.....	360
Appendix 13: Scatter plot of incremental costs and QALYs on the cost-effectiveness plane (after multiple imputation)	363
Appendix 14: Cost effectiveness acceptability curve (after multiple imputation), FIMM Vs MIMM	364
Appendix 15. Results of main and sensitivity analysis (in GBP).....	365

Approximate number of words in this thesis: 50,000

List of tables

Table 2.1: The Ten leading GBD causes of years lost to disability in individuals aged 15-34 in the UK, both sexes combined, 2013.....	49
Table 4.1: Summary of methods and conclusion of included trial-based economic evaluations.....	94
Table 4.2: Summary of methods and conclusion of included model-based economic evaluations.....	97
Table 4.3: Permutation /matrix of the cost-effectiveness results based on Nixon et al. (164).....	122
Table 5.1: Existing studies describing the annual Costs of BD, UK.....	131
Table 5.2: Units costs of resources use, 2010-11.....	136
Table 5.3: Demographic and clinical characteristics of participants.....	144
Table 5.4: Average annual costs of BD per patient based on imputed data.....	146
Table 5.5: Population-based prevalence studies.....	154
Table 5.6: Cost of illness of BD in the UK (2010-11, £).....	156
Table 5.7: Predictors of BD associated costs according to the MI-GEE analysis...	160
Table 6.1: Baseline sociodemographic characteristics of the study.....	193
Table 6.2: The 20 most frequently (> 90% of 1257) reported EQ-5D-3L health states.....	194
Table 6.3: Test-retest Reliability of the EQ VAS score and EQ-5D-3L index scores anchored by no change in the BD health states.....	197
Table 6.4: Correlation matrix between the QIDS-SR16 (depression), EQ-5D-3L, and VAS.....	198
Table 6.5: Known-group validity of the EQ-5D-3L in different levels of depression and mania.....	201
Table 6.6: Responsiveness of the EQ-5D-3L index and the EQ VAS to changes in health status as anchored by change in QIDS-SR16 scores.....	203
Table 6.7: Responsiveness of the EQ-5D-3L index and the EQ VAS to changes in health status as anchored by change in ASRM scores.....	206

Table 6.8: Overview of factor loadings and uniqueness based on an exploratory factors.....	208
Table 6.9: Confirmatory factor analysis, goodness-of-fit indices of five models.....	209
Table 7.1: Summary of the features employed in the economic evaluation.....	236
Table 7.2: Demographic and clinical characteristics of participants.....	242
Table 7.3: Resource use by participants during the 12 months follow-up period (Complete case analyses).....	243
Table 7.4: Resource use related to the psychoeducation interventions (FIMM and MIMM).....	244
Table 7.5: Estimated mean costs per patient per year (£, 2012-13 values) based on complete data.....	246
Table 7.6: Summary statistics and differences in outcomes between interventions at each time-point based on complete case analysis (FIMM, n=33; MIMM, n=33).....	250
Table 7.7: Cost-utility analyses for FIMM versus MIMM.....	252

List of Figures

Figure 1.1 Summary of the OXTEXT research programme.....	23
Figure 1.2 True Colours mood monitoring system.....	24
Figure 1.3 Contribution to the literature.....	31
Figure 2.1 Diagnostic criteria for bipolar disorder DSM-IV=Diagnostic and Statistical Manual of Mental Disorders, fourth edition.....	36
Figure 2.2 Life course of Bipolar disorder.....	42
Figure 2.3 Clinical management of Bipolar disorder.....	45
Figure 2.4: Broad categories of Psychological interventions for adults with BD.....	46
Figure 2.5: Bipolar disorder disability-adjusted life years (DALYs) in the UK (2013), by age and gender.....	50
Figure 2.6: Changes in Bipolar disorder DALYs in the UK from 1990 to 2013.....	51
Figure 3.1: Key Mental health – economic questions.....	57
Figure 3.2: Type of Health related quality of life (HrQoL) instruments.....	68
Figure 3.3: Comparison of health profile between intervention and no-intervention.....	72
Figure 3.4: Hierarchy of utility methods as recommended by NICE (26).....	73
Figure 3.5: The cost-effectiveness plane.....	74
Figure 3.6: Cost-effectiveness acceptability curves.....	78
Figure 4.1: Flowchart of search strategy and included studies.....	91
Figure 4.2: Study setting and year of study.....	93
Figure 4.3: Chart showing the fulfilment of quality criteria by trial based studies.....	110
Figure 4.4: Stacked bar showing hierarchy of evidence used in DAMs as adapted from Copper et al. (163).....	117
Figure 4.5: Chart showing the fulfilment of quality criteria by model based studies.....	119
Figure 5.1: Distribution of the total costs and health care costs of BD by major sub-categories.....	148
Figure 5.2: Annual costs of BD to society and the public sector (£, 2010/11 prices, per person with BD).....	150
Figure 5.3: Study flow diagram.....	153
Figure 5.4: Pooled analysis of BD prevalence in population-based prevalence studies.....	155
Figure 5.5: One-way sensitivity analysis of key parameters (in GB£ millions).....	157

Figure 5.6: Probabilistic sensitivity analysis (PSA) results of annual costs per individual with BD.....	159
Figure 5.7: Distribution of annual societal costs of BD in different studies.....	165
Figure 6.1: Unidimensional model with one latent construct (psychological distress).....	189
Figure 6.2: Two first-ordered correlated factors.....	190
Figure 6.3: Three first-ordered correlated factors (depression, HRQoL, mania).....	191
Figure 6.4: Bi-factor orthogonal model with a general psychological distress factor and individual domain factors (i.e. depression, HRQoL, mania).....	192
Figure 6.5: Distribution of responses to items of EQ-5D in patient sample at baseline and 12 months in comparison to the general population.....	196
Figure 6.6: Known-group validity of EQ-5D-3L in different levels of depression.....	200
Figure 6.7: Distribution of responses on individual EQ-5D-3L dimensions for mania sub-categories.....	202
Figure 6.8: Responsiveness as measured by SRM in the EQ-5D-3L index and the EQ-VAS to changes in health status as anchored by change in QIDS-SR16 scores.....	205
Figure 6.9: A bi-factor CFA model with a general psychological distress factor along with three factors representing three included instruments.....	210
Figure 7.1: Illustration of inter-relationship between psychological treatments in BD.....	228
Figure 7.2: Example “prompt” sent by the TC system to study participants.....	230
Figure 7.3: Distribution of costs in treatment arms.....	249
Figure 7.4: (a) Cost effectiveness plane of unadjusted bootstrapped net cost and QALY pairs (complete case analysis); (b) Cost effectiveness plane of adjusted bootstrapped net cost and QALY pairs.....	253
Figure 7.5: Cost effectiveness acceptability curve (complete case analysis).....	254
Figure 8.1: Contribution to the literature.....	279
Figure 8.2: Burden of disease and associated research funding.....	283

List of abbreviations

AIC	Akaike information criterion
BIC	Bayesian information criterion
CBA	Cost-benefit analysis
CCA	Cost-consequence analysis
CDF	Cumulative distribution function
CEA	Cost-effectiveness analysis
CEAC(s)	Cost-effectiveness acceptability curve
CI	Confidence interval
CMA	Cost-minimisation analysis
CRD	Centre for Reviews and Dissemination
CT	Computerised tomography scan
CUA	Cost-utility analysis
DALY(s)	Disability-adjusted life year(s)
EQ-5D	EuroQol five dimensions questionnaire
EU	European Union
FCA	Friction cost approach
HCA	Human capital approach
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
LY(s)	Life year(s)
NICE	National Institute for Health and Care Excellence
NHS	National Health Service
NHS EED	The NHS economic evaluation database
OOP	Out of pocket
PSA	Probabilistic sensitivity analysis
QALY(s)	Quality-adjusted life year(s)
SD	Standard deviation
SG	Standard gamble
UK	United Kingdom
US(A)	United States of America
WTP	Willingness-to-pay (or willing-to-pay)

1.1 Introduction to the thesis

Bipolar disorder (BD) is a chronic severe mental illness and one of the leading causes of disability for both men and women worldwide. Individuals with BD have a higher risk of suicide, a higher prevalence of mental and physical comorbidities and a higher prevalence of cardiovascular risk factors than the general population. The pathophysiology and aetiology of BD remain poorly understood with an unpredictable prognosis (1). There is no definitive cure, and patient management strategies primarily focus on preventing relapse. However, once medicines are discontinued, there is an increased risk of relapse and subsequent increase in morbidity and mortality. Further, BD is often misdiagnosed as unipolar depression, delaying appropriate diagnosis and management by up to ten years (more in **Chapter 2**) (2).

As the disease worsens over time, BD significantly impairs health related quality of life (HRQoL) and personal and social functioning, and poses an enormous economic burden to society (3-5). Over the last two decades, the management of BD has evolved with new therapies emerging, including second-generation antipsychotics, modalities of care (adjunct psychological therapies) and service delivery (collaborative care) (6). Despite this, there is insufficient knowledge of the costs and HRQoL implications of BD, and the cost effectiveness of current treatments for BD in the UK (7). Moreover, most of the burden of BD falls outside the healthcare system (i.e. National Health Service), unlike other severe mental illnesses such as schizophrenia, where the

relative burden on health care provider is higher than that of BD (8). The relatively limited impact of BD on healthcare services (NHS) has likely contributed to a relatively low priority for research into BD, resulting in inadequate funding (9, 10)

Publicly funded healthcare services such as the UK National Health Service (NHS) offer free healthcare at the point of delivery; everyone can access these services regardless of their ability to pay, and service provision is solely based on clinical need. In the current climate of austerity, rising healthcare costs and competing disease priorities, prioritisation of scarce resources in all areas of healthcare is pivotal for rational healthcare planning and policy-making. To facilitate priority setting and policy development in this challenging area, the discipline of health economics has become an increasingly important decision-making aid to ensure that new interventions not only maximise health outcomes but also offer value for money (11).

Although systematic reviews on the cost-effectiveness of BD management strategies have been previously published, the evidence was too limited with wide variation in the quality of reported studies to draw any conclusions (12, 13). Advances in medicine and technology is driving up the cost of using limited resources. It is, therefore, important to identify cost-effective treatments and management strategies for BD in the UK.

The relative neglect in research funding and weak evidence base in BD could be partly explained by lack of understanding of the scale and cost of BD in the UK (14). The recently published "Priorities for mental health" report commissioned by NHS England

makes a critical contribution in highlighting the economic cost of mental illness and possible priorities for service improvement in the UK (15).

However, most of the reported evidence related to the cost of severe mental illness and cost-effectiveness of interventions was derived from research undertaken on schizophrenia. The economic evidence on costs included costs of schizophrenia on healthcare and society based on studies conducted in the UK (15). The only costs related to BD reported was the cost of £12,300 associated with managing a crisis episode per person with BD (derived from a French study conducted in 1998-99) in contrast to around £19,800 per person with schizophrenia (16, 17)

Previous reviews have demonstrated the paucity of evidence related to the burden of BD on population in the UK (18, 19). For this purpose, cost-of-illness (COI) studies are arguably well placed to systematically quantify the costs associated with BD and highlight the relative magnitude of the impact of illness on society in comparison to the burden of other diseases (20, 21). For instance, the COI of Dementia published in the first Dementia UK report was instrumental in raising the profile of dementia and influenced shaping of national dementia strategies including the Prime Minister's challenge on dementia (22).

Even though cost-of-illness studies can provide useful information to a wider non-technical (non-economists) audience about the burden of disease, evidence on economic evaluations of interventions is of a higher priority to inform effective and efficient public health policies for addressing the disease burden identified through cost-of-illness studies (20, 23).

As briefly mentioned above, BD is associated with relapses, medical co-morbidities, unwanted hospital admissions and reduced social and physical functioning; this can result in unfavourable clinical outcomes and markedly decreased HRQoL (24). The measurement of patient self-reported outcomes including HRQoL is of significant importance in evaluating relative effectiveness, and cost-effectiveness (25). To compare cost-effectiveness estimates across disease areas, current guidelines for economic evaluations recommend the use of quality-adjusted life years (QALYs) based on generic, preference-based measures of HRQoL such as EuroQol-5D with three levels (EQ-5D-3L) (26).

EQ-5D-3L is a widely used preference-based measure of HRQoL that assigns utility weight to an individual's overall preference for a health state; it measures health status on five core domains of general health (i.e. mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) with three level of responses (no problem, some problems, extreme problems) (27). However, the assessment of HRQoL in severe mental health disorders such as BD is not without challenges. HRQoL, as measured by EQ-5D-3L, may not always correlate well with symptom changes and its validity in mental illness has been questioned (28, 29).

Further, there is a paucity of studies reporting EQ-5D-3L outcomes in the field of BD. This could be due to several reasons ranging from reluctance to use EQ-5D-3L to lack of economic perspective. Lack of use of EQ-5D-3L precludes the estimation of QALYs and comparison across disease areas (30). With the growing need of economic evaluations for priority setting in healthcare, statistical techniques such as 'mapping' offer a means of empirically translating different measures of health outcomes

including clinical outcome measures or HRQoL or utilities in BD to one another and thereby facilitating economic evaluation. In the context of economic evaluation, a 'mapping' exercise is a rescue act for trials that collect only clinical outcomes. It is, therefore, considered second best to gathering generic preference-based utility measures along with clinical outcomes in a study (31).

In long-standing mental illness like BD, often associated with a high risk of relapse, international guidelines recommend psychological interventions to enhance patient's awareness, compliance with medicines, prevention of relapse and improve patient functionality (7, 32, 33). The latest NICE guideline on the management of BD in the UK, encourages long-term care models delivered by a cross-disciplinary team comprised of a multicomponent intervention that includes shared decision-making, psycho-education, monitoring of outcomes, relapse prevention and standard protocol-based pharmacotherapy (7).

However, existing research has indicated a dearth of evidence on whether psychological interventions in BD management represents good value for money (34). As one of the research priorities to improve patient care, NICE recommended an evaluation of effectiveness and cost-effectiveness of psychological and collaborative models of care (7). To make a strong economic case for change, it is vital to have robust evidence on the cost-effectiveness of new models of care. To date, there is no comparative economic evaluation of an individual psychoeducation intervention in the UK.

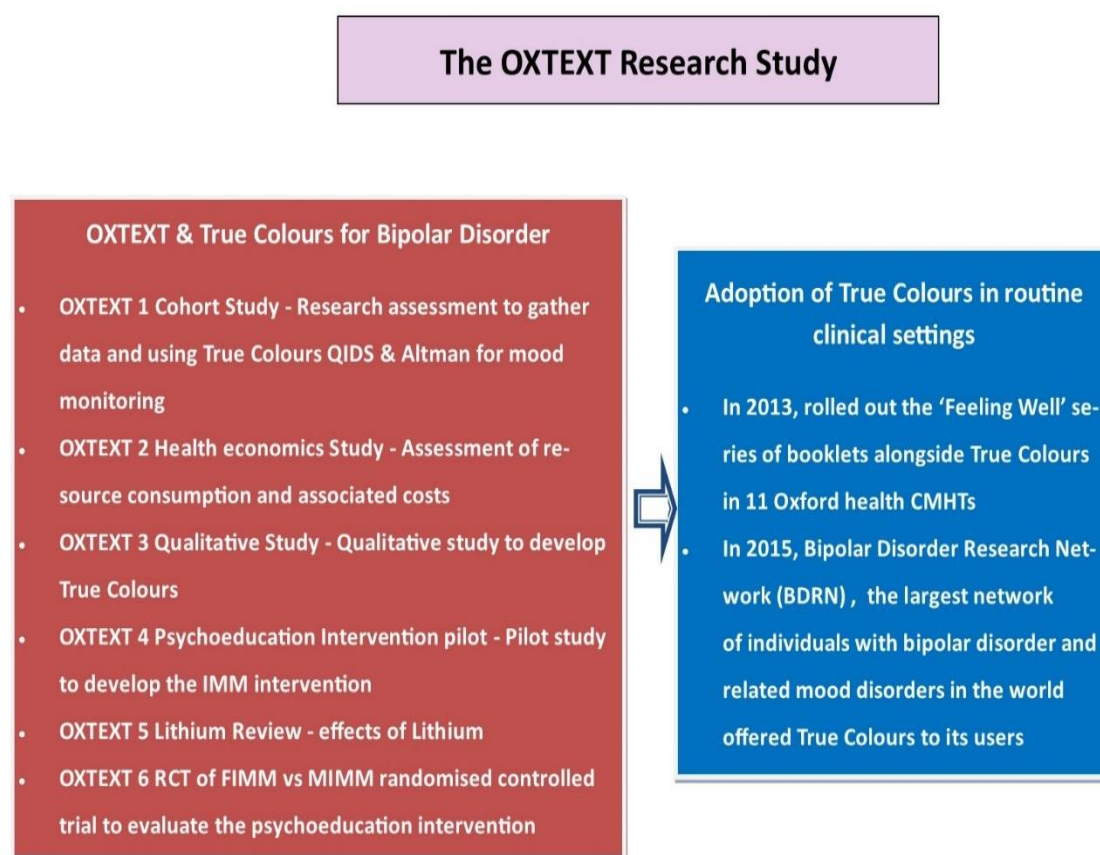
This thesis explores aspects of the health economics of BD that are pivotal for priority setting and decision making: existing economic evidence on BD management, the burden of BD in the UK, appropriateness of the widely-used EQ-5D-3L in quantifying outcomes associated with patients treated for BD and cost-effectiveness of novel modalities of care for BD.

This thesis is primarily based on data collected from an applied cohort study, and the subsequent section will briefly introduce them with **Chapters 5, 6 and 7** describing the relevant sub-studies in further detail. The remaining sections of this chapter will outline the thesis aims and objectives, elaborate the thesis structure and explain how this thesis would contribute to the body of evidence in mental health economics.

1.2 Applied study: the OXTEXT research programme

The OXTEXT research programme (see **Figure 1.1**) is a portfolio of six closely related studies (OXTEXT-1, -2, -3, -4, -5, and -6). The aim of this programme is to develop management strategies for BD along with innovative tools that can detect early signs of relapse in patients with BD. The main innovation is centered around a simple-to-use SMS or web based mood monitoring system, called 'True Colours' (TC), where BD patients self-report their symptoms via short messaging service (SMS) or web-based system on a weekly basis (35). The overarching goal of this program is to assist patients in personalized management of BD

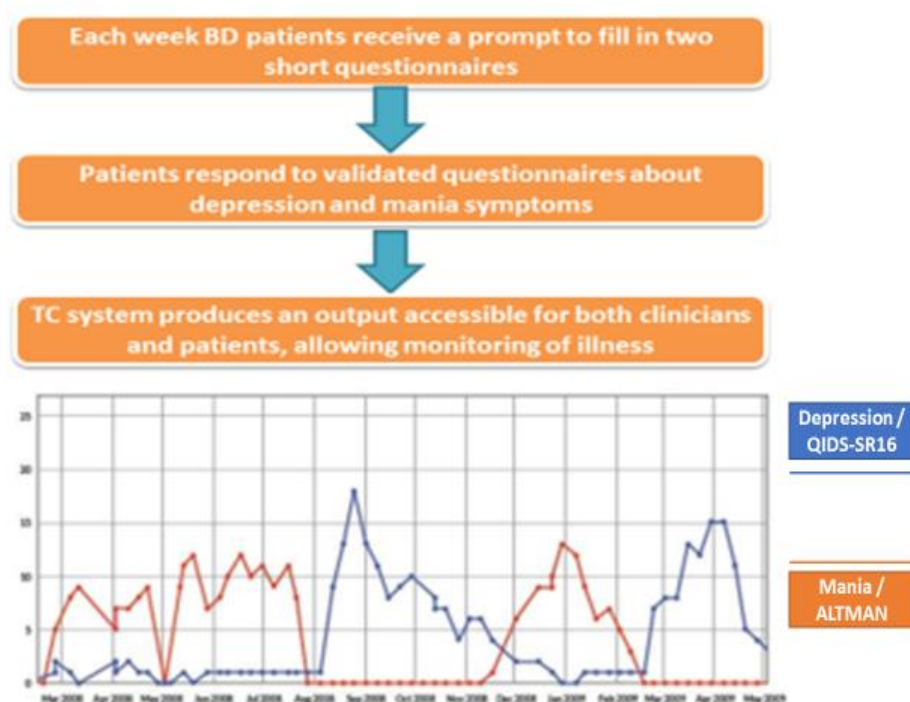
Figure 0.1.1 Summary of the OXTEXT research programme



*IMM: Integrated Mood Management; FIMM: Facilitated Integrated Mood Management; MIMM: Manualised Integrated Mood Management

The TC system allows real-time remote monitoring of symptoms and clinical outcomes. Individuals with BD receive weekly text or email prompts and reply in real time with their mood ratings using validated self-report scales for depressive (the Quick Inventory of Depressive Symptomatology or QIDS-SR16) and manic symptoms (the Altman Self-Rating Scale for Mania or ASRM) (36). The TC system also produces a graphical output accessible to both patients and clinicians allowing them to track the patient's condition and identify early signs of relapse in a collaborative way (see **Figure 1.2**).

Figure 1.0.2 True Colours mood monitoring system



Remote mood monitoring or self-monitoring of mood for adults with BD has been successfully implemented at the Mood Disorders Clinic at the Warneford Hospital, Oxford using the TC mood monitoring (TC-MM) system since 2008 and is now offered to patients as a part of routine clinical care (35, 37).

OXTEXT-1 is a cohort study that aims to understand the factors that influence the prognosis of BD (including the socio-clinical-demographic characteristics of users of the TC system, clinical outcomes, and genetic data) to personalise treatment recommendations. OXTEXT-1 began in 2010, is still ongoing, and had recruited around 300 participants by the conclusion of recruitment for the OXTEXT-6 study in December 2012.

Individuals were recruited across Oxfordshire and Buckinghamshire via systematic and non-systematic recruitment methods. Systematic recruitment involved screening for potential individuals with BD via two routes: community adult mental health teams within the Oxford Health NHS Foundation Trust, and clinicians at the Bipolar Disorder Research Clinic at the NIHR-Clinical Research Facility (NIHR-CRF) at the Warneford Hospital. Individuals were recruited non-systematically via advertisements, posters and flyers in the local community, and via calls for volunteers on the research team websites, and in the local media.

OXTEXT-2 is a health economics study that evaluates resource use data gathered prospectively from a sub-group of the OXTEXT-1 cohort. The evaluation included all health and social care costs, costs to the patients and their families, and wider societal costs due to lost productivity.

OXTEXT-3 is a qualitative study that aims to optimise TC by gathering information and feedback from service users, and OXTEXT-4 develops and then conducts a feasibility assessment of a novel psycho-education intervention called Integrated Mood Management (IMM) that aims to raise knowledge about the illness, factors that are associated with relapse, and strategies to self-manage the illness. The mode of delivery of this intervention is evaluated in OXTEXT-6. OXTEXT-5 is a systematic review of the adverse effects of lithium.

OXTEXT-6 is a randomised clinical trial that evaluates whether Facilitated Integrated Mood Management (FIMM) is better than Manualised Integrated Mood Management (MIMM) for patients with bipolar disorder. FIMM involves regular support from trained facilitators and MIMM, without direct facilitation. Participants in both trial arms had routine access to their clinicians and were engaged in weekly symptom monitoring using TC. The data required for this thesis was obtained from patients recruited in the OXTEXT-2 and OXTEXT-6 sub-studies.

1.3 Aims and objective of the thesis

In current times of constraining healthcare budgets, the economic case of a disease plays a crucial role in shaping public debate and concern and leverage policy decisions. It is vital to understand the relative burden of illness on the society so that appropriate investments are made to identify cost-effective interventions and implement health policy decisions to achieve maximum social benefit. This is particularly important in the case for BD, where there is a sheer dearth of evidence-base on the prevalence of disease, its costs, the relevance of outcomes to capture health benefits, and cost-effectiveness of management strategies in the UK.

The overarching aim of this thesis is to inform decisions about local and national service provision for BD by applying a variety of health economic tools to appraise the economic costs, appropriateness of outcomes, and cost-effectiveness of management strategies for individuals with BD in the UK. The specific objectives are:

1. Identification of the most cost-effective interventions in the field of BD through systematically reviewing and critically appraising economic evaluations in the field of BD management
2. Use prospectively collected individual-level patient data to estimate the societal burden of BD in the UK, and to explore the clinical and socio-demographic factors that drive variations in the costs associated with BD.
3. Estimate the appropriateness (with regards to test-retest reliability, discriminant validity, convergent validity and responsiveness) of applying the EQ-5D-3L outcome measure to assess the HRQoL of BD patients, and explore the feasibility of 'mapping' the EQ-5D-3L from disease-specific instruments.
4. Evaluate the economic evidence for alternative structured psychoeducation interventions in the management of BD

1.4 Outline of the thesis

This thesis has eight chapters. This section provides a synopsis of each of these chapters.

Chapter 2 describes the burden of BD in the UK including epidemiology, life course, and its effect on mortality and disability. The clinical management of BD is discussed along with current management practices and treatment modalities.

Chapter 3 provides a conceptual overview of health economics in mental health, focusing on costs, outcomes and cost-effectiveness. This chapter specifically

elaborates on aspects of cost assessment that are relevant to individuals with BD, such as social care and informal care costs, and productivity losses. Furthermore, attention is drawn to outcome measures with related challenges associated with their use in BD. The key concepts described in this chapter provide the foundation for the evaluations reported in subsequent chapters.

Chapter 4 systematically reviews and critically appraises published economic evidence on BD management strategies. The analysis is split into two parts: part one explores trial-based economic evaluations, while part two focuses on model-based economic evaluations. The review also highlights the strengths and limitations of the different methodologies that have been applied and provides suggestions for good practice when undertaking an economic evaluation in the management of BD.

Chapter 5 presents the results of a cost-of-illness (COI) study conducted to estimate the NHS, societal and social care costs associated with BD in the UK. This analysis is broadly split into three sections. Part one describes the identification, measurement and valuation of all resources associated with the BD patients enrolled in a pragmatic observational study (OXTEXT-2). This analysis takes into account the costs incurred by the NHS (including primary and secondary care), the costs incurred in social care, and the costs associated with lost productivity and informal care.

It also explores the clinical and socio-demographic factors that drive variations in the costs associated with BD, and the uncertainty surrounding these estimates is evaluated in sensitivity analysis. The second section reviews the prevalence of BD in

Western Europe and examines the COI of BD in the UK. In the final section, the results from part one are compared with those reported in other COI studies to assess the practicality of applying standard methods and understand the factors that drive cost variations.

Chapter 6 addresses the issues associated with applying the EQ-5D-3L in BD. The first part of this chapter reports the appropriateness of using the EQ-5D-3L to value HRQoL in patients with BD. This is assessed by estimating the test-retest reliability, construct validity and responsiveness of the EQ-5D measure compared to two disease-specific measures: the Quick Inventory of Depressive Symptomatology-Self-Reported-16 (QIDS-SR16) and Altman Self-Rating Mania Scale (ASRM). The second part of the chapter assesses whether it is feasible to map from the QIDS-SR16 and ASRM instruments onto the EQ-5D-3L in situations where the EQ-5D-3L was not used. Exploratory and confirmatory factor analyses are conducted to assess the conceptual similarity between items included in the disease-specific and generic instruments. This chapter uses data from the OXTEXT-6 sub-study.

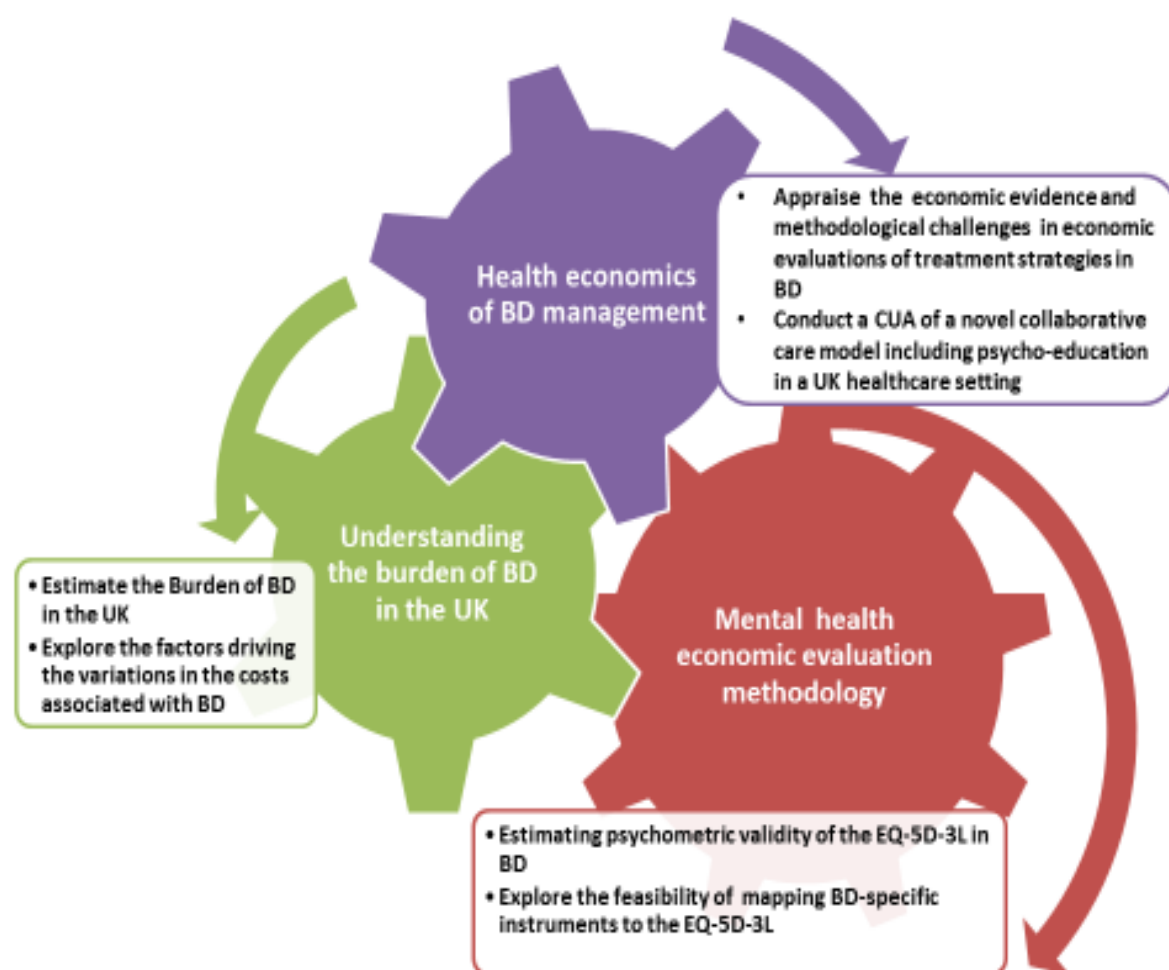
Chapter 7 draws on the previous chapters to bring together evidence on the costs and outcomes associated with alternative treatment strategies in BD management. This evidence is used to conduct a cost-utility analysis (CUA) of a novel structured psycho-education (five psycho-education sessions with a facilitator) intervention also known as FIMM compared with Manualised Integrated Mood Management (MIMM) for patients with bipolar disorder who use remote mood monitoring in a UK healthcare setting. This within-trial economic evaluation uses data from the OXTEXT-6 sub-study.

Finally, **Chapter 8** presents the discussion and conclusion. It provides an overview of the thesis findings, explores the strengths and limitations of the research, and identifies the wider methodological and practical implications of the findings. This chapter also includes some suggested directions for future economic evaluations and related research in the field of Bipolar disorder

1.5 Contribution to the literature

This thesis will contribute to the existing literature in three research domains: the health economics of BD management, understanding the burden of BD in the UK and mental health economic evaluation methodology (see **Figure 1.3**).

Figure 1.3 Contribution to the literature



The first contribution to the health economics literature is made by a systematic review of economic evaluations for BD management strategies. Studies will be critiqued with the aim of identifying the most cost-effective psychological and pharmacological treatment strategies in the management of BD. The results of this review will contribute to the broader task of informing and assisting the decision-maker in allocating scarce health care resources. As per existing guidelines, psychoeducation is the first choice of psychological intervention in the long-term management of BD. The CUA presented in Chapter 7 advances this literature, providing the first evidence of the cost-effectiveness of a novel psychoeducation intervention in a UK healthcare setting.

There is a relative lack of awareness or limited knowledge on the burden of BD in the UK. So, what is the scale of the problem? What are the associated costs and distribution impact across sectors? This thesis will also contribute to an improved understanding of the burden of BD in the UK. The cost-of-illness study in Chapter 5 will provide a comprehensive estimate of the societal burden of BD derived from a pragmatic cohort of individuals with BD in the UK. The chapter will also provide new evidence on the associations between main clinical-demographic factors and healthcare and societal costs for patients with BD.

The third significant contribution is related to mental health economic evaluation methodology. Given the need for consistency across economic evaluations, the National Institute for Health and Care Excellence currently recommends the EQ-5D-3L as a tool for quantifying utility and conducting economic appraisals. However, there is a lack of research examining the use of the EQ-5D-3L instrument in BD. This study will assess the appropriateness of EQ-5D-3L to measure HRQoL in individuals with BD. Further, it will explore the feasibility to map BD-specific clinical scales on EQ-5D-3L in the absence of EQ-5D-3L data.

1.6 Conclusion

This chapter has introduced the thesis, and described its aims and objectives. It also provided an outline of the subsequent chapters and explained how the thesis will contribute to the existing literature. **Chapter 2** will now provide background information on the epidemiology, aetio-pathology and natural history of BD along with the current state of management of the disease. This understanding has a major bearing on recognition of the impact of BD on HRQoL and health care costs and is, therefore, relevant for the purpose of resource allocation decisions related to BD in the UK.

2.1 Introduction

This chapter provides background details on Bipolar disorder (BD) including diagnosis, aetiology, epidemiology, course of illness and the associated health and economic burden that arise as a result of the disease, for the patient, family, healthcare system and the wider impact on society. An overview of the clinical management of the illness is presented with its challenges.

2.2 Disease area

Bipolar disorder (BD) (alternatively known as mood disorder, bipolar affective disorder, manic-depressive illness, bipolar spectrum disorder) is one of the top five leading causes of mental and substance use disorders burden worldwide. Even with a relatively low global prevalence in comparison to other diseases, BD is responsible for more years of life lived with disability than Ischemic heart disease, Cerebrovascular disease, epilepsy, and dementia (38).

BD is a debilitating severe mental illness with frequent recurrences characterised by episodes of severe mood swings ranging from depression to heightened mood. The disease is classified using International Classification of Diseases (ICD-10) or the Diagnostic and Statistical Manual of Mental Disorders (DSM) (39). According to the

fourth edition of the DSM (DSM-IV), BD represents a spectrum of diseases that are heterogeneous from mild depression to mania or more severe forms of mood disorders associated with psychotic symptoms. BD-I and BD-II disorders are the two of the most common forms of the disease, with rapid cycling (cyclothymia) and bipolar disorder not otherwise specified (NOS) being rarer variants (see **Figure 2.1**).

BD-I constitutes one or more manic (mood elation) or mixed episodes, and one major depressive episode. A mixed episode is characterised by at least one week of symptoms that fit the criteria for both major depressive and manic episodes. BD-II features one or more major depressive episodes and at least one hypomanic episode. Hypomania is a milder form of the manic episode that does not result in profound social or physical dysfunction. While, in cyclothymic disorder, patients experience milder episodes of depression and mania alternating for more than two years, patients in rapid cycling form may have ≥ 4 episodes of mania, depression, hypomania or mixed symptoms within a year (40).

Figure 2.1 Diagnostic criteria for bipolar disorder DSM-IV=Diagnostic and Statistical Manual of Mental Disorders, fourth edition (39)

Diagnostic criteria for bipolar disorder (based on DSM-IV)	
Bipolar I disorder*: Presence, or history of, at least one manic (or mixed) episode	
Bipolar II disorder*: Presence, or history of, at least one major depressive episode and at least one hypomanic episode (with no history of a manic or mixed episode)	
The symptoms are not attributable to physical illness or physiological effects of a drug or other substance and are not better accounted for by another psychiatric disorder	
Manic symptoms Elevated, expansive, or irritable mood Increased activity that is goal directed or psychomotor agitation Reduced need for sleep Excessive involvement in pleasurable activities with likely adverse consequences Inflated self esteem or grandiosity Increased or pressured speech Flight of ideas or racing thoughts Distractibility	Depressive symptoms Depressed mood Markedly reduced interest in nearly all activities Increased or decreased appetite or weight Insomnia or hypersomnia Psychomotor retardation or agitation Fatigue or loss of energy Feelings of excessive worthlessness or guilt Impaired concentration or indecisiveness Recurrent thoughts or actions of death or suicide
Manic episode: At least four manic symptoms including altered mood that persists for at least a week and causes marked functional impairment, hospital admission, or there are psychotic symptoms	
Hypomanic episode: As for manic episode but less severe; symptoms persist for at least four days and functioning is noticeably altered but not enough to lead to hospital admission or to greatly impair function. There are no psychotic symptoms	
Major depressive episode: Five or more persistent depressive symptoms (which must include depressed mood or diminished interest), which last for at least two weeks and occur on most days, and that cause serious distress or functional impairment	
Mixed episode: Persistent mood symptoms for at least a week that meet criteria (apart from duration) for both a manic and major depressive episode, which occur at different times or rapidly alternate	
Psychotic symptoms: These may occur during manic episodes in bipolar I disorder (but by definition not during hypomanic episodes) and during depressive episodes in either bipolar I or bipolar II disorder	
*The World Health Organization classification ICD-10 does not distinguish between bipolar I and bipolar II disorder and requires another mood episode in addition to a single manic episode	

2.3 Aetiology

Although BD has been known for centuries, the aetiopathogenesis of this illness is complex and still poorly understood. The multi-factorial risk model of aetiopathogenesis with interactions between the environment/lifestyle factors (i.e. childhood maltreatment, social support, stressful life events, cannabis use etc.) and susceptible genes is the prevailing theory of causation. The recent genome-wide association studies into BD has demonstrated that the heritability of BD is driven by multiple risk alleles, with substantial overlap with other severe mental illnesses including Schizophrenia (41). Beside the multi-factorial risk model, a variety of other aetio-pathological mechanisms have been proposed to cause the development of BD; the most relevant are reported below. Neuro-biochemical imbalance in neurotransmitter pathways (catecholamines, serotonin, gamma-aminobutyric acid (GABA), glutamate etc.), hormones (i.e. brain-derived neurotrophic factor), and dysregulated signalling cascades have traditionally been implicated in the causation of BD (24).

Latest research findings have also indicated that inflammation plays a fundamental role in the pathophysiology of BD, with accumulating evidence supporting the conceptualisation of BD as a multi-system inflammatory disorder. These studies suggest that patients with BD demonstrate high levels of inflammatory cytokines, and successful treatment leading to remission normalises the levels of peripheral inflammatory mediators (42). However, at the same time, studies have reported that the elevated levels of inflammatory mediators is not observed in all patients with BD

(43). Other theories such as around sleep and circadian rhythm are gaining prominence. As phasic dysregulation of mood is a core feature of BD, it has been hypothesised that dysfunction in circadian rhythm plays a major role in the aetiology of BD (44).

This complex aetiopathogenesis makes it difficult to disentangle cause and effect, thereby making it a challenging prospect to develop new drug targets or predict the prognosis of clinical outcomes.

2.4 Epidemiology

A recent multi-national survey of over 60,000 adults in 11 countries across America, Asia and Europe reported a total lifetime prevalence rate of 2.4% for bipolar spectrum disorders that included BD-I (0.6%), BD-II (0.4%) and sub-threshold BD (1.4%). Although the prevalence varied from 0.1% in India to 4.4% in the United States, the authors highlighted that such variations in bipolar spectrum disorders rates are likely to be due to significant under-detection in low income countries as a result of misdiagnosis, mistreatment, and the greater stigma attached to mental illness (45).

There is a lack of epidemiological data on the prevalence of BD in the UK. For estimating the burden of BD in the UK, prevalence data were gathered from a rapid review of population-based epidemiological studies exploring BD prevalence in Europe and analysed using a random effects meta-analysis in **Chapter 5**; this yielded

a prevalence of 0.8% (95% CI: 0.43 % to 1.3%). During the finalising of this thesis, the global burden of disease study group used Bayesian meta-regression modelling to publish prevalence estimates for all countries. In the UK, the age-standardized, past-year prevalence of BD in the year 2013 was reported to be 0.8% (95% CI: 0.7 % to 0.9%) in females and 0.6% (95% CI: 0.5 % to 0.7%) in males (5).

BD can occur at any age but often affects people early in their professional lives with the average age of onset of first symptoms being 20 years of age. Both men and women are equally vulnerable to the severe forms of this disorder (46). BD is often prone to misdiagnosis with 69% of patients diagnosed with unipolar depression. Those who are misdiagnosed, on an average, seek an opinion from four doctors before receiving the proper diagnosis, and more than a third of patients wait up to 10 years for their diagnosis (1). The consequences of misdiagnosis are far-reaching as these patients are rigorously treated with anti-depressants with their associated side-effects, higher rates of suicide and hospitalisation; this results in poor outcomes and unwanted higher costs of iatrogenic healthcare (47).

2.5 Psychiatric co-morbidities

Patients with BD have significantly higher rates of co morbidity with other psychiatric disorders such as anxiety disorder, eating disorders, and substance abuse disorder. The lifetime prevalence of comorbidity of BD along with anxiety disorders is 42%, with substance abuse is 40%, and with eating disorders is 17% (48, 49). According to a recent meta-analysis, pooled lifetime prevalence of comorbid anxiety disorders was

45%, with the prevalence being three times greater than those without BD. Moreover, the review found that individuals with co-morbid anxiety disorders continued to present with functional impairment even in euthymic remission states (50). Individuals with comorbid BD and substance abuse present with further challenges including earlier onset of illness, increased frequency of mixed or rapid cycling episodes, delayed recovery, higher rates of re-hospitalisations and frequent suicide attempts (51). Further, personality disorders can confound the diagnosis and course of illness in BD patients, with patients presenting with severe residual symptoms and poor outcomes; the rate of comorbidity with BD can be as high as 50% (52).

2.6 Medical co-morbidities

In addition to psychiatric co-morbidities, BD patients also report frequent medical co-morbidities and side-effects of treatment. Individuals with thyroid function abnormalities have higher risk of BD (53). Research has also indicated that patients had higher rates of obesity, and metabolic syndrome, thereby increasing risks of cardiovascular illness and diabetes (54). The high rates of physical co-morbidities especially cardiovascular risk factors results in a two to three times higher mortality rate from cardiovascular diseases than observed in the general population (40).

2.7 Course of illness

The life course of BD is poorly understood. Its natural history follows a protracted waxing and waning course which makes the management of this disease a great challenge. Patients are heterogeneous in their symptoms and signs (55). The relapse rates for manic or depressive episodes could be as high as 50 % for the first year to 70% by the fifth year. Additionally, adherence to long term complex medication regimens is inadequate and subsequent outcomes poor. The recurrent episodic nature of illness can cause deterioration of functioning, and an increasing number of episodes is associated with poor disease prognosis (56, 57).

Figure 2.2 elucidates the different phases of the illness along with its multiple relapses and associated impact on psychosocial functioning. There is an overlap of disease symptomatology, and even in the presence of normal mood or euthymic state, individuals can have sub-threshold depressive or manic symptoms. As a result of cognitive impairment, despite symptomatic recovery, functional recovering may lag behind (24). This functional impairment may have impact on the HRQoL of patient with BD as explored in **Chapter 6**.

Figure 2.2 Life course of Bipolar disorder

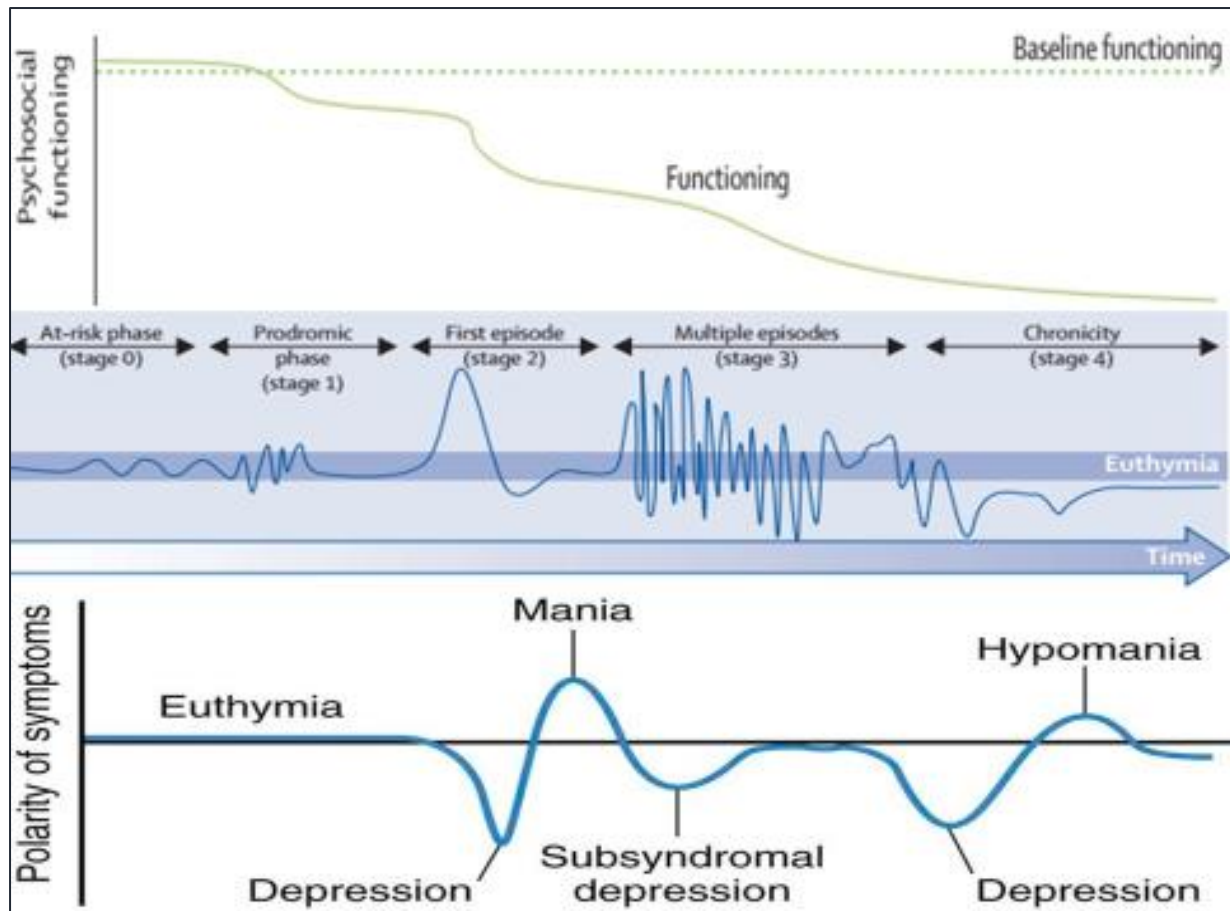


Figure adapted from Grande et al (2016)

Depressive symptoms dominate the course of the disease with depressive symptoms more common than manic (3.5 times) and mixed or rapid cycling (five times) episodes. The depressive episode is also accountable for the largest burden of illness, with 80% of patients at an increased risk of committing suicide (40, 58). Comparison of BD outcomes with that of those suffering from major depressive disorder (MDD) suggested that approximately two-thirds of patients with MDD attain remission in ten years in contrast to less than half of patients with BD (59).

Although traditionally BD has been considered as an episodic condition characterised by recurrent spells of hypomania/mania and depression, studies have highlighted that BD patients represent a continuum of both depressive and manic symptoms and exhibit variation in severity of depressive or manic symptoms (60). It implies that the different phases of BD cannot be considered as discrete bipolar phenomena. Further, a significant proportion of BD patients demonstrate residual sub-syndromal symptoms between major syndromal episodes, and they could be symptomatic for almost half of their long-term follow-up periods (61). The persistence of residual symptoms and co-occurrence of depressive and manic symptoms has implications for estimating and reporting HRQoL as further investigated in **Chapter 6**.

2.8 Mortality

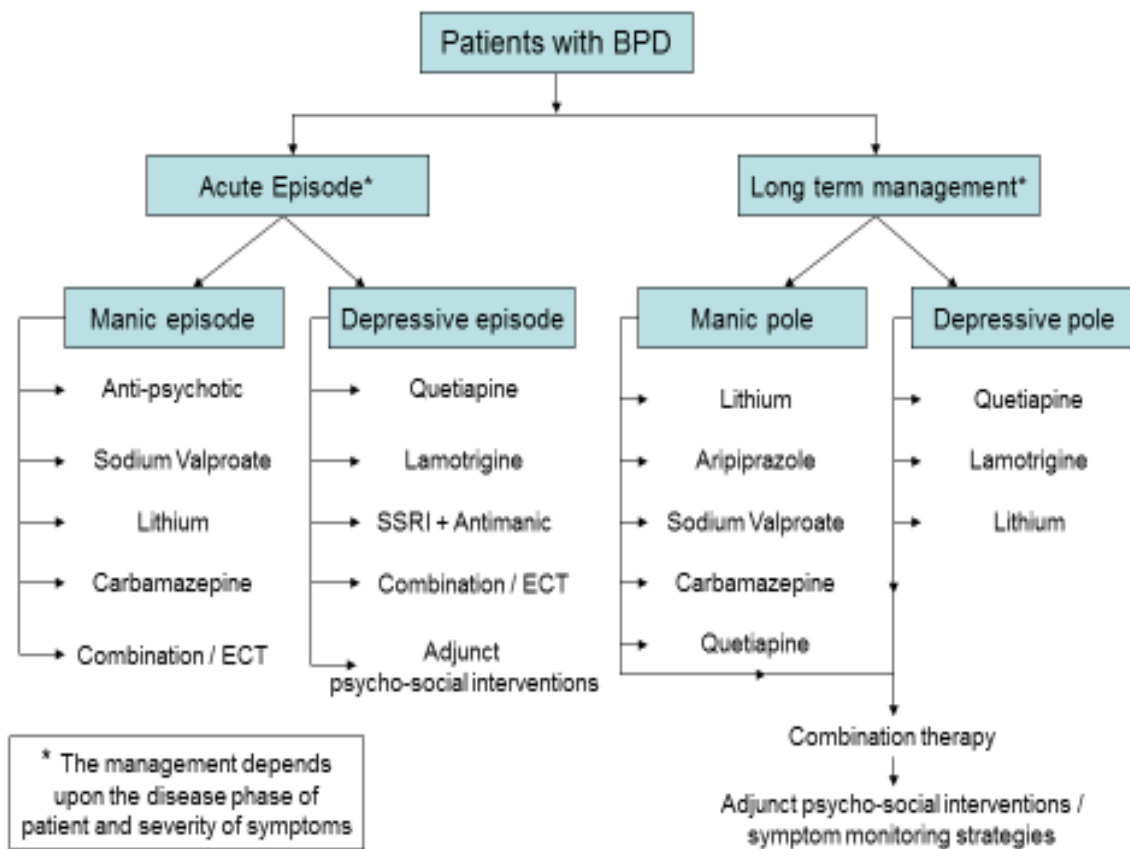
A growing body of evidence has indicated that BD is associated with premature mortality. Traditionally, the increased deaths were believed to be related to unnatural causes including suicide, homicide, or accidents. However, individuals with BD have a shorter lifespan on average of 8.5-9 years lesser in comparison to the general population. Most of these premature deaths were due to cardiovascular diseases and cancer (62-64). A recent Danish nationwide registry study of 5.4 million people including 22,635 patients with bipolar disorder, found that boys and girls with BD at 15 years of age had lower life expectancy by 12.7 and 8.9 years respectively (65).

Individuals with BD, especially those with mixed states are at highest risk of suicide attempts, with one- fourth of BD patients attempting suicide. If BD is left untreated, the suicide rates is around 15% (66). In comparison to the general population, BD is associated with 10-30 times higher suicide deaths, with being a male and first-degree family history strongly associated with the likelihood of suicide deaths (67).

2.9 Existing management strategies

According to the latest National Institute for Health and Care Excellence (NICE) guidelines for BD management, the mainstay of managing BD is appropriate use of pharmaceutical agents as seen in **Figure 2.3** (7).

Figure 2.3: Clinical management of Bipolar disorder



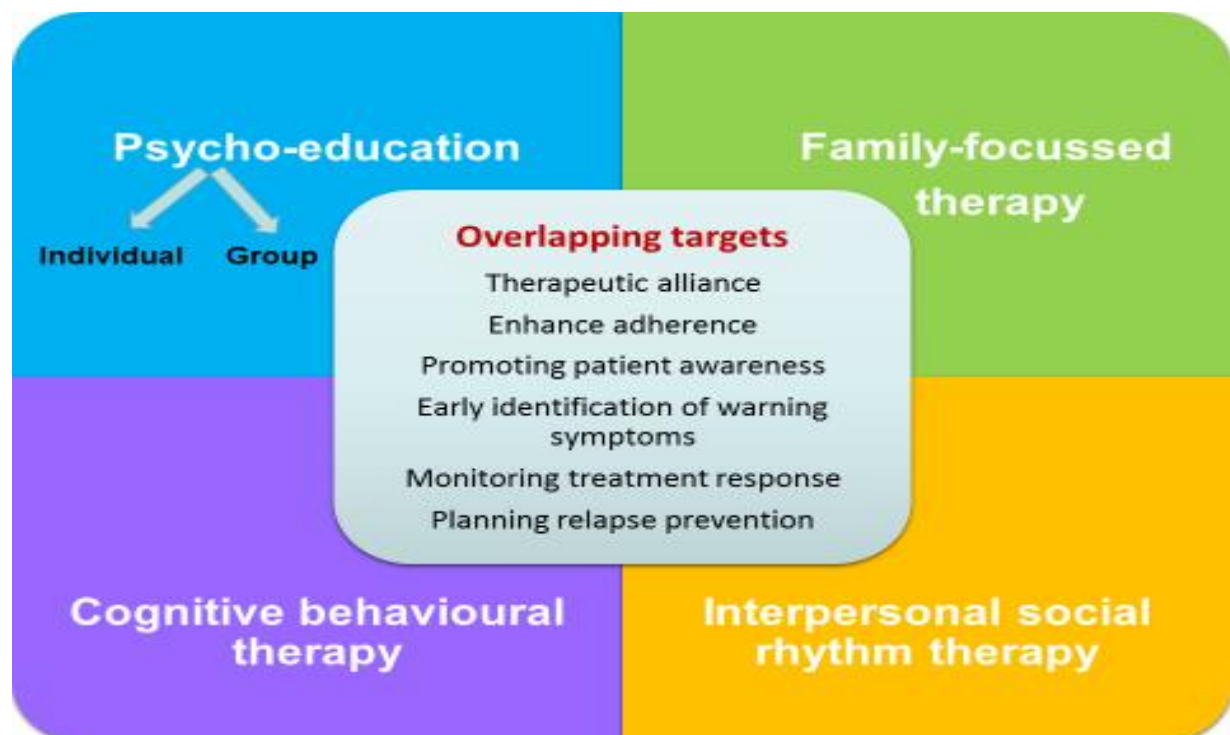
Patients who do not respond to mono-therapies should be tried for combination regimens, and psychological interventions should be used to enhance patient's awareness, compliance with medication regimen and prevention of relapse. The guidelines acknowledge that the long term management of BD is not optimal and requires further research (7). However, given the natural course of BD disease, focus on preventative strategies should be paramount.

Psychological interventions of BD patients have been used as an adjunct to pharmacological therapy to reduce recurrence, improve patient's health outcome and

social functioning. Although, there are four main types of psychological interventions, they have overlapping elements and common goals that are depicted in **Figure 2.4**. Psychoeducation is the recommended first line of psychological intervention in the UK

The results of previously published studies and guidelines equivocally advocate adjunctive psychological interventions which would account for a greater improvement in clinical outcome and quality of life in BD patients than pharmacological treatment alone (7, 68). Nonetheless, there are still questions that remain unanswered especially with regards to the type of the psychological intervention and delivery model to improve patient care and outcomes. In line with the latest NICE guideline's research recommendation, **Chapter 7** will explore the cost-effectiveness of a psychological intervention (7).

Figure 2.4: Broad categories of Psychological interventions for adults with BD



2.10 Monitoring strategies

Due to the complexity of BD management (**Figure 2.2** and **2.3**), decision making for a patient heavily relies on the pattern of past episodes, triggers of illness and patient's response to therapy. Further, a paradigm shift towards redefining BD phenomenology along with a focus on inter-episodic mood instability has led to development of several monitoring strategies that oversees patient's mood fluctuations, drug response and course of illness (10).

There are paper based self-monitoring tools such as the National Institute of Mental Health Life Chart Methodology (NIMH-LCM) and the Clinical Monitoring Form which are used in clinical practise, but they are limited by poor compliance and laborious nature of data entry (69, 70). In the recent years, automated mood charting with the help of software installed on home based computers; hand-held professional digital assistants (PDA), interactive voice response system over the telephone, optical scanning and electronic diaries have been experimented in trial populations. However, these technologies are restricted by their reliance on desktop or hand-held computers and lack of systematic evidence generated from randomised controlled trials and cost-effectiveness analyses. Mobile phone-based self-monitoring is a promising innovation in the field of BD, and have showed promising results in long term disease management of other chronic diseases (35, 71, 72). They are cheaper than their digital counterparts including laptops, desktops, PDAs and have the advantage of being portable, allowing patients to have the freedom to respond anywhere that has mobile coverage.

Lately, there has been a growing interest in electronic self-monitoring of mood using smartphones and internet based technologies. It has the advantage of being a contemporaneous real-time submission of mood data that reduces the possibility of recall bias and errors with data entry. Moreover, it encourages patient engagement and allows a more patient-centric model of care (73). As **Chapter 1** highlighted, this thesis included a cohort of individuals who have been using remote mood monitoring as part of their routine clinical care to follow-up their mood variation and identify significant changes in their mood.

2.11 Disease burden, and functional impairment

As a result of the widespread consequences of BD, the disorder has a major impact on patient well-being along with emotional, health and economic implications not only for the patients, but also their families and the wider society (74). According to the Global burden of disease report 2010, BD was one of the leading causes of disability worldwide and the fourth largest burden of diseases in young individuals aged between 10-25 years (75, 76). The changing demographic pattern including an aging population along with population growth is propelling the burden of BD over time (5).

Using the data from the global burden of disease (GBD) 2013 study, the burden of BD in the UK was briefly explored. The GBD study uses disability-adjusted life years or DALYs to capture the burden of diseases, and risk factors worldwide. DALY is a single metric that quantifies both mortality and morbidity associated with a condition and is measured on a scale that ranges between 0 (representing the best possible health

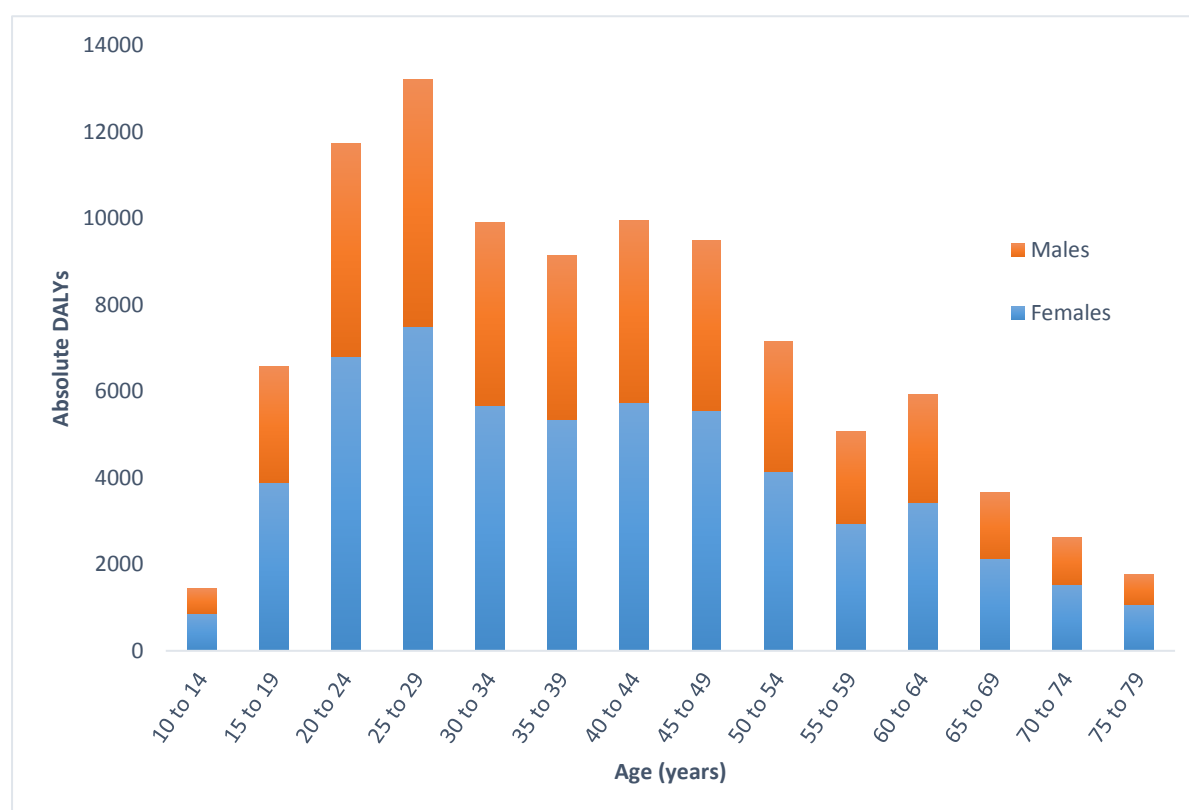
state) and 1 (worst possible health state). It represents the sum of the present value of years of future life lost due to premature death (also called as Years of life lost, YLL), and the average severity of years of life lived in bad health (Years lost to disability, YLD). DALYs represent negative aspects of an illness on health with one DALY equivalent to a loss of one healthy year of life due to a particular illness (77). In the UK, BD was one of the leading causes of years lost due to disability attributing to three percent of YLDs in individuals aged between 15 and 34 years (see **Table 2.1**).

Table 2.1: The Ten leading GBD causes of years lost to disability in individuals aged 15-34 in the UK, both sexes combined, 2013

Causes	YLDs	Attribute proportion
All causes	1534001	100%
Low back pain	162415	11%
Migraine	122847	8%
Major depressive disorder	114648	7%
Asthma	96313	6%
Anxiety disorders	78619	5%
Neck pain	63304	4%
Opioid use disorders	54910	4%
Iron-deficiency anaemia	43157	3%
Bipolar disorder	41462	3%
Falls	30509	2%

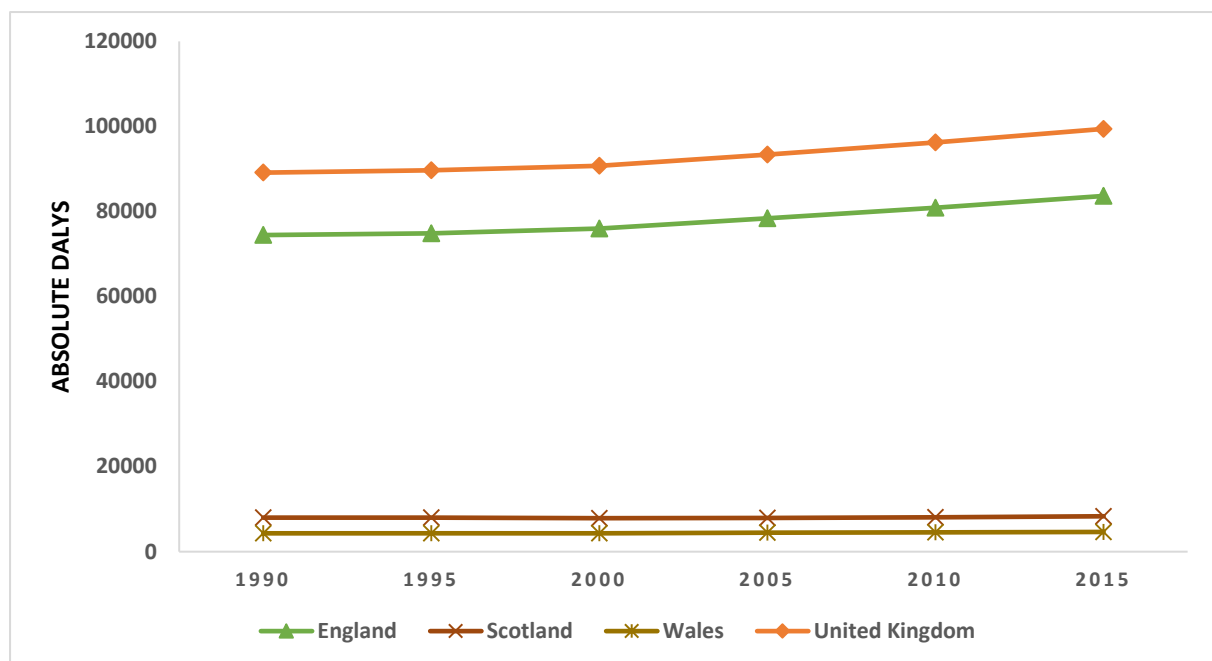
Figure 2.5 displays DALYs for bipolar disorder across the life span for both sexes in the UK. It indicated that the burden was highest in the 25-29 age group (14%).

Figure 2.5: Bipolar disorder disability-adjusted life years (DALYs) in the UK (2013), by age and gender



BD was responsible for 89,133 (54,726-132,468) DALYs in 1990 and 99,421 (61,343-146,945) DALYs in 2013, all accounted for by YLDs. Although the absolute number of DALYs has increased over the time in the UK, no significant time trends for DALY rates of BD was observed between 1990 and 2013 (see **Figure 2.6**); this could be a result of arguably stable prevalence rates of BD used in the GBD study (77).

Figure 2.6: Changes in Bipolar disorder DALYs in the UK from 1990 to 2013



Like other severe mental illness like schizophrenia, individuals with BD experience a high degree of work impairment and less than 50% are in any kind of employment. Even patients who achieve complete clinical remission exhibit difficulties in attaining premorbid levels of functioning (78). The EMBLEM study, for instance, indicated that at baseline, 69% of 3,497 BD patients had high work impairment and who subsequently after two years of follow-up had poorer occupational and overall functional outcomes (79).

Although the reasons for poor vocational outcomes in patients with BD are multifactorial and poorly understood, factors such as residual depressive symptoms and significant cognitive deficits (i.e. working memory and attention span) have been most consistently implicated with functional impairment in BD patients under remission (80-82)

2.12 Economic cost and quality of life

As a major mental health disorder, BD requires multi-disciplinary and perpetual care that places a significant economic burden on constrained healthcare resources (3). The economic burden of an illness is usually estimated by cost-of-illness (COI) studies. COI studies can comprehensively estimate both health and social care costs (i.e. costs for primary care, outpatient visits, inpatient admissions, medicines, social care etc.) along with opportunity costs as a result of lost resources also called as indirect costs (i.e. productivity losses due to illness or premature death as result of disease, and informal care provided by family and friends) (83).

However, cost-of-illness estimates can vary considerably within and across diseases, primarily due to differences in the country of origin, study population, methodology, and cost categories included in each study (19). In the field of BD research, the majority of the COI studies are US-based and there is paucity of UK based COI studies. A recent systematic review of BD studies found that most of the studies used secondary data with lack of transparency and narrow perspective of analysis (23). Even in studies conducted in the same country, significant cost variations could be observed. For example, annual direct cost estimates for BD in the US varied between \$2,651 and \$18,420 per person. In the UK, the annual direct healthcare costs associated with BD per person was reported to be between £670 and £3,916. Such heterogeneity in the published costs of BD could pose challenges for decision making (23). Adopting recommendations of good practise from literature, and using data from

a prospective cohort of individuals with BD, **Chapter 5** will present results of a COI study quantifying the societal costs associated with BD in the UK.

BD also results in significant impairment of health-related quality of life (HRQoL) and individuals exhibit symptoms that disrupt normal day-to-day activities along with their vocation, family and social life. A complete clinical remission or recovery should include a return to an acceptable and realistic functionality and HRQoL (4). Patients with BD have significantly lower HRQoL scores as measured by the generic HRQoL instrument SF-36, compared with the general population (84).

Previous analyses have indicated that individuals spend three times more time in a depressive state than in manic phase and that depressive symptoms seem to be the primary determinant of quality of life in BD (85). Even individuals with hypomania and mania symptoms have lower HRQoL than the general population, and these lower scores are influenced by sub-syndromal depressive symptoms (86, 87). Routine clinical use of HRQoL measures improves clinical management, patient satisfaction and outcome. It is, therefore, important to have valid HRQoL measures that could quantify patient's recovery and functionality (88). Furthermore, preference-based HRQoL measures are fundamental inputs required for economic evaluations. **Chapter 3** provides a description of different instruments used to measure HRQoL in BD, and the validity of the most widely used preference-based HRQoL measure in BD (EQ-5D-3L) is explored in **Chapter 6**.

2.13 Summary and conclusion

This chapter has provided a clinical background for the thesis including an overview of BD, describing the aetiology, epidemiology, life-course of the disease, summarising the current UK modalities of care, describing the burden of illness, economic costs and associated impact on functioning and HRQoL.

BD is a severe mental illness that is characterised by recurrent fluctuations in mood. There are no biomarkers or biological tests to identify the disease. Worldwide BD prevalence estimates are between 0.1 and 4% of the population with the prevalence in the UK reported to be 0.8%, but the prevalence is often underestimated. The disease involves a long delay in diagnosis of up to ten years as it is frequently misdiagnosed with years of depression.

Existing treatment guidelines are largely in agreement over the management of BD. The main aim in the management of BD patients is to prevent relapses, suicides, acute episodes, unwanted hospital admissions and improve social and physical functioning. Over the past decade, there has been an emergence of new management strategy modalities (e.g. adjunct psychotherapies) and service delivery options. Economic evidence related to the management of BD has become an increasingly important aid to decision-making for the allocation of scarce health care resources within the UK and other countries. Details of economic evaluation methods and cost-effective BD management strategies are discussed in further detail in Chapters 3 and 4.

The delay and misdiagnosis of BD, co-morbidity with other psychiatric and medical conditions, and suboptimal treatment of BD significantly contributes to the societal costs and burden of BD. COI studies provide critical information on the economic burden that BD enforces on a society and can highlight the relative magnitude of its impact in comparison to other illnesses. However, the inconsistency of methodologies employed along with a paucity of COI studies conducted in the UK limits a robust estimation of the burden of BD in the UK. To contribute towards highlighting BD burden and facilitate priority setting in the UK, **Chapter 5** will examine the COI of BD in the UK. The depressive phase of BD is the predominant cause of disability among BD patients resulting in severe limitation of HRQoL, social function and premature death. The importance of patient-centred BD management including the assessment of function, social and HRQoL outcomes cannot be over-emphasised. Existing HRQoL estimation tools and validity of a commonly used preference-based instrument EQ-5D-3L is described in detail in **Chapter 3** and **6** respectively.

Guidelines have indicated the emerging need of using multiple modalities of care in long-term management of symptoms. The ideal long-term management for a BD patient should encompass all three core interventions including pharmacological, psychological and mood monitoring. However, there is a lack of evidence for psychological interventions, and there is growing need for appropriate psychological interventions and monitoring strategies to prevent relapse and improve patient outcomes. NICE has recommended further research in psychological interventions of BD. **Chapter 7** will explore in detail the economic evaluation of a standardised psychoeducation intervention using a combination of mood monitoring and psychoeducation.

3.1 Introduction

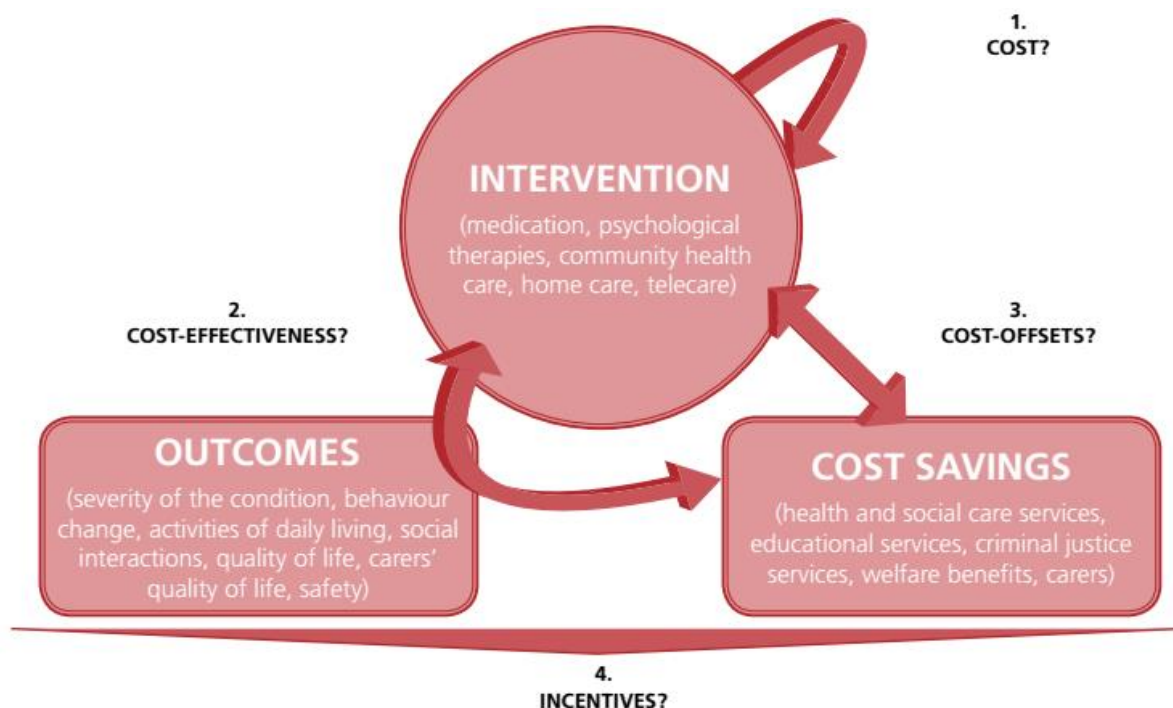
This chapter provides a conceptual framework required to understand the methods proposed in the following chapters that explore aspects of the health economics of BD that are pivotal for priority setting and decision making. This includes a description of cost-of-illness studies, along with the principal elements of an economic evaluation: costs, outcomes and cost-effectiveness. Specifically, the aspects of cost assessment relevant to individuals with BD such as social and informal care costs, and productivity losses are described. Furthermore, attention is drawn to outcome measures with related challenges associated with their use in BD. Finally, frameworks for conducting economic evaluations are described: trial-based economic evaluation and decision-analytical modelling.

3.2 Background to economics and BD

Chapter 1 and **Chapter 2** highlighted that Bipolar disorder (BD) is one of the leading causes of disability with a lifetime prevalence of 1%-2% in the population (89). Individuals with BD have up to ten years of lower life expectancy than the general population, primarily due to significantly higher levels of preventable physical health conditions, and medical co-morbidities (90).

BD poses a significant emotional, functional, economical and health impact not only for the patients and their families, but also has system-level financial consequences that can affect provision of services (91). Mental health services in the UK are under severe pressure. The waiting times are too long, and in some regions, there is a lack of specialist mental health services. In the current climate of constrained healthcare resources, policy makers may prefer implementing preventive or curative strategies that provides the maximum value for public resources while improving the health and wellbeing in individuals with mental illness. To shape policy development, influence value-based commissioning and provide efficient service delivery, it is important to build the economic case for BD.

Figure 3.1 Key Mental health – economic questions



Source: The economic case for better mental health (92)

Figure 3.1 highlights the key elements to assess the economic evidence for mental health action. This simple framework outlines the four questions that economists address to build up an economic case. The magnitude of illness expressed in terms of burden of disease or cost-of-illness (COI) is often the first step in articulating the problem statement with the public and policy makers on the relative burden of an illness (92-94). COI studies can also be helpful for identifying a range of cost items for further economic evaluation studies. However, COI studies alone are insufficient for allocating resources as they provide no indication of how to achieve efficiency gains or health benefits (20)

The American economist Paul Samuelson defines economics as:

“The study of how men and society end up choosing, with or without the use of money, to employ scarce productive resources that could have alternative uses, to produce various commodities and distribute them for consumption, now or in the future, among various people and groups in society. It analyses the costs and benefits of improving patterns of resource allocation.” (95).

This definition encloses some of the core concepts at the basis of health economics and highlights the importance of scarce resources, alternative use and decision making. The available resources will never be sufficient to implement all the needed programs or treatments and the resources once deployed to implement a strategy, could not be used for the next best alternative. In this sense, economists define the forgone value from this decision as the opportunity cost. In other terms, implementing a programme will limit our resources to fund another innovative treatment. Therefore,

resources should be carefully deployed to maximise the benefit for the population health. In this framework, the assessment of the available interventions should not exclusively focus on the costs or effectiveness alone but should evaluate both these aspects together.

In the last decades, COI studies, and economic evaluations (EEs) in mental health have jointly influenced policy makers to recognise a particular problem and take decisions comparing alternatives assessing both their costs and consequences (8, 22). Further, a variety of approaches have been used to assess the health outcomes, leading to different economic evaluations. The following sections aim at describing cost-of-illness studies, the main elements of economic evaluation, different types of EEs and the framework employed to conduct EEs.

3.3 Cost-of-illness (COI) studies

"The aim of COI studies is descriptive: to itemize, value, and sum the costs of a particular problem with the aim of giving an idea of its economic burden."

As Jefferson et al. (2000), note that the primary aim of a COI is to elaborate the relative burden of a disease (96). COI are not economic evaluations as they do not compare interventions. Instead, they are used to aggregate information and highlight the total cost (economic burden) associated with an illness.

3.3.1 Nomenclature of COI studies

There are different types of COI studies and they are primarily categorised according to:

a) The type of epidemiological data used: prevalence-based or incidence-based

The prevalence-based approach focuses on costs incurred by all existing patients (prevalent cases) due to a specific disease over a defined period, usually one year, regardless of date of onset. An incidence-based approach estimates potential life-time costs avoided if new cases (incident) of a disease are averted. In other words, this approach calculates life-time costs of incident cases occurring typically within a defined year. An incidence-based approach indicates how costs vary as disease progresses thus allowing assessment of the costs associated with the management of illness from onset until recovery or death. It is based on more assumptions than a prevalence-based approach and is suitable for chronic conditions with stable course over time (97). The prevalence-based approach is regarded as the more suitable methodology for highlighting the total economic burden of a disease, more specifically, for drawing attention to major cost areas that could be amenable for potential interventions (98).

b) The type of estimation method chosen to quantify resource use and costs:

top-down or bottom-up approach

Top-down or gross-costing estimates are derived from assigning a proportion of total health expenditure from a given perspective, for instance, total mental health expenditure for the entire population, to the disease under investigation according to its prevalence. It is relatively easy to undertake, but typically multiple sources of information are used to calculate proportion of resource use/costs that should be assigned to the specific disease, thus the precision of results largely depends upon the rigour of apportioning costs to the condition of interest. Further, the gross-costing approach suffers from the inherent inability of national/regional healthcare databases to capture all the desired cost categories associated with a complex disease, for example informal care and productivity loss costs. Additionally, this approach cannot account for condition-specific co-morbidities (i.e. metabolic syndrome in BD patients).

Bottom-up costing estimates the average patient-level healthcare resource utilization/costs and indirect costs for a given time-period and extrapolates it to the whole population using appropriate prevalence data. As costs tend to be derived directly from patients, statistical analyses of costs can be performed to adjust for likely confounding factors and detect differences between patients for important socio-demographic characteristics. However, the bottom-up approach is time consuming and there are limitations whilst extrapolating when the sample is not representative of the population of interest (99, 100).

3.3.2 Limitations of cost-of-illness studies

COI studies suffers from the lack of consistency and transparency in input parameters. The static snapshot of costs presented in COI may be of limited value as they do not directly guide on how resources should be allocated to improve efficiency (20, 101). At the same time, the process of policy making is complex, and it has been argued that, COI studies provide a more suitable “point of entrée” into discussions around important resource allocation decisions than cost-effectiveness analysis (21, 102). **Chapter 5** will provide evidence on the societal costs associated with BD in the UK and explore the main factors influencing those costs in patients with BD.

3.4 Main elements of an Economic Evaluation

Drummond et al (2005) defined an economic evaluation as: “the comparative analysis of alternative courses of action in terms of both their costs and consequences” (103). EE can be performed to compare both costs and outcomes of healthcare interventions to identify the alternative that represents the best value for money. This section aims at describing the main elements of an economic evaluation such as: comparators, perspective of the analysis, costs and outcomes.

3.4.1 Comparators

Due to a scarcity of resources, the deployment of one programme implies that another one misses out. Therefore, the selection of comparators could be crucial for the output of an EE. The recommendation from guidelines for EEs is to choose the current active treatment or most common clinical practise as a comparator (104). In this sense, an inappropriate selection of the control may bias the analysis and lead to a sub-optimal decision. If there is a disagreement on which is the best intervention between the current treatments, more than one comparator could be used. However, there may be a lack of evidence on the effectiveness and cost-effectiveness of the included comparators or no treatments available for the specific disease. In these specific cases, using the “do nothing” or placebo are acceptable comparators. In the case of lack of evidence, this choice is justifiable as we could not assume the current treatment to be better than the “do nothing” option.

3.4.2 Perspective

One of the first questions that health economists should be considering when developing an EE is to identify for whom the costs and benefits of the intervention are relevant. The perspective adopted in health economics are the third-party payer/healthcare provider, patient or the societal perspective. The National Institute for Health and Care Excellence (NICE) in the UK prefers the NHS and personal social services perspective (i.e. the healthcare provider perspective); this includes primary (e.g. GP and nurse visits), secondary care costs (e.g. hospitalisation and outpatient

visits) and social care costs. The societal perspective could include but not limited to costs incurred to the healthcare provider, social care, patients' personal costs, carers, productivity losses and wider costs in education and criminal justice settings (105).

Costs could be broadly categorised as follow:

a) Direct healthcare costs: these are costs that are directly related to delivery of healthcare and where actual payments are made. It usually includes costs for outpatient visits, inpatient admissions, medicine costs, costs related to diagnosis, tests, devices and treatment, and costs incurred in primary care. There are some other costs such as transportation costs to healthcare facilities, costs to the criminal justice system and cost of social care that are often categorised as direct non-healthcare costs (106).

b) Indirect costs: these are costs where no payments are made but still resources were consumed. This will typically include productivity losses to employer (sick leave) and society due to illness or premature death as result of disease, and informal (unpaid) care (time taken by carer to look after patients) provided by family and friends (106). Although the productivity cost may represent a considerable burden on both the society and on patients/carers; its use in the full economic evaluation is limited due to methodological uncertainty in estimating the productivity loss. The inclusion of indirect costs (i.e. loss of productivity) has raised concerns related to equity and double counting in Health Related Quality of Life (HRQoL) benefits. (107).

In addition to the above two categories, an often-quoted category is that of intangible costs which are defined as the costs attributed to pain and suffering because of illness or intervention. Although these costs are difficult to quantify, they can be captured using health related quality of life instruments to derive utility or willingness-to-pay methods (103).

However, limiting an evaluation to NHS perspective alone may exclude some of the costs typically falling on individuals with mental illness such as the costs relevant to informal care, productivity losses, costs in criminal justice settings, etc. Excluding these costs may lead to a biased conclusion. For example, excluding the carers' cost could result in adopting a community-based rather than residential care intervention. Even though the NICE acknowledges this problem, the inclusion of these costs is also limited as a result of the complexity of the data capturing costs beyond the health and social care system (108).

3.4.3 Costs

Once a perspective is adopted, the cost should be identified, measured and valued consistently with the point of view of the EE analysis.

Identification

The first step in EEs is usually the identification of resource use. The perspective adopted will strongly influence the inclusion of a resource items. The most

straightforward approach is to pose the question: “for whom are the resources important?” Relevant methods to identify resources could be clinical expert opinion, literature, patient focus groups, standard costing guidelines and templates (104, 109).

Measurement

The resource items should be appropriately measured with appropriate discrete units according to their nature. For example, physiotherapy sessions could be measured considering the number of sessions, a surgical intervention can be measured in terms of minutes spent in the theatre. The data can be provided from a variety of sources such as: ad-hoc health economics questionnaires (an example is provided in **Appendix 8**), Client Service Receipt Inventory, patient diaries and clinical records (110, 111).

Valuation

The valuation step relies on the results from the identification and measurement of the resource items. The cost of the intervention is calculated by multiplying the quantity of each resources by an appropriate unit cost. The unit cost should include the provision of the services (including supervision, administration, multi-agency meetings, training costs, clinical salaries and equipment) and directly estimated for better accuracy or obtained from authoritative sources (e.g. NHS reference cost, Personal Social Services Research Unit - PSSRU) (110).

3.4.4 Measurement of Health Outcomes

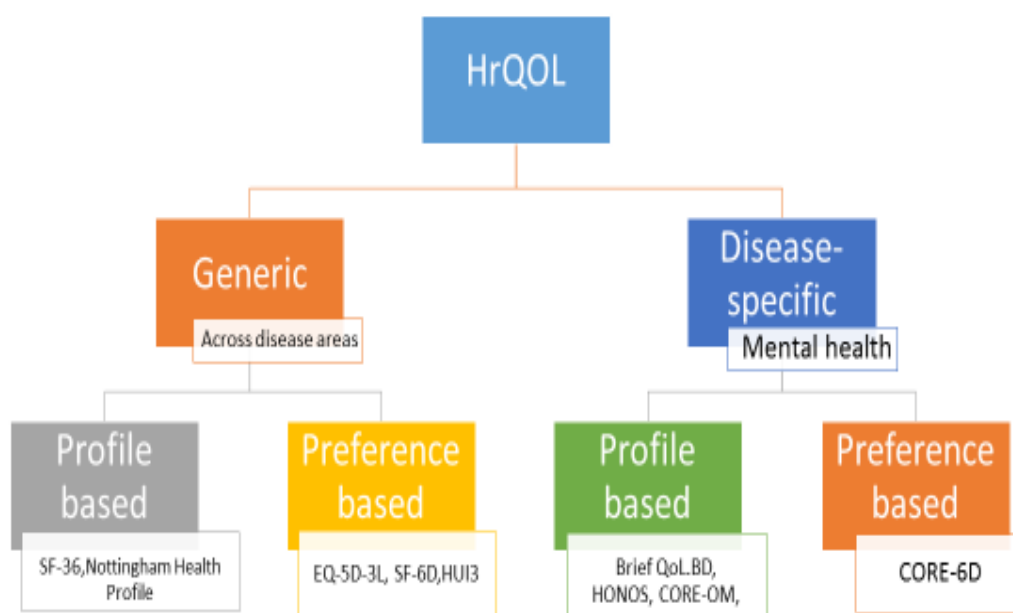
NICE recommends that the outcomes of an intervention in health economics evaluations should account for all health effects. In this context, outcomes are broadly classified as clinical and HRQoL outcomes. Although a disease-specific clinical outcome measure (e.g. depression score as measured by Quick Inventory of Depressive Symptoms–Self Report, QIDS-SR16) could be appealing for its sensitivity and rigorousness, these outcomes are single-scale measures and may not allow for comparison with different disease areas. Moreover, clinical measures are focused on disease specific aspect relevant for disease management and may neglect broader aspects of patients' quality of life (104).

HRQoL is a multidimensional construct. The instruments that measure HRQoL are broadly categorised as generic (used across patient populations with different diseases or healthy general population) or disease-specific (focussing on a particular disease (109). Both generic and disease-specific instruments can be further classified as a) profile-based, or b) preference-based measures (see **Figure 3.2**).

Profile-based measures are used to compare the relative burden of diseases either across disease areas or within a broad disease area such as mental health. They generally include aspects of physical functioning, self-care, psychological status, levels of pain, and amount of social integration. They are designed to yield profile scores on the multiple dimensions of HRQoL. These questionnaires allow comparisons of interventions in different population groups scores disease areas and

can contribute to assessment of comparative effectiveness (112). On the other hand, preference based measures combine different dimensions of health to generate multiple health state values which are then valued based on population preferences to yield a single summary score or utility. Utilities are anchored by a value of one indicating full health and zero for being dead (health states considered worse than dead can be represented by negative values) (113).

Figure 3.2 Type of Health-related quality of life (HrQoL) instruments



Short form-36 (SF-36); EuroQol-5D (EQ-5D-3L), short-form 6D (SF-6D), Health Utilities Index 3 (HUI3), Brief BD-specific QoL instrument (Brief QoL.BD), Health of the Nation Outcome Scales (HoNOS) Clinical Outcomes in Routine Evaluation - Outcome Measure (CORE-OM), Clinical Outcomes in Routine Evaluation-6D (CORE-6D)

Generic Questionnaires

The generic questionnaires aim to capture the health status of patients independently from their specific condition. Many generic questionnaires have been developed and used in BD over the last decades (e.g. EuroQoL instruments, Short Form questionnaires, WHO-QOL-BREF) (114-116). Between them the EuroQoL instruments (e.g. EQ-5D-3L) and Short Form questionnaires (e.g. SF-36) can be used to derive utility score in different population groups and allow for comparison within and across disease areas (more about in utility in subsequent section). Although the generic HRQoL questionnaires may capture relevant changes in people's quality of life regardless of age, treatment and measure the patient's well-being, they could be insensitive to small changes in HRQoL. This lack of precision in estimating the changes in HRQoL could be overcome by administering jointly a generic and a specific HRQoL instruments (25).

In the UK, the NICE recommends the use of the generic preference-based multi-attribute utility instrument EQ-5D-3L for conducting economic evaluation across disease areas (26). The EQ-5D-3L is based on a descriptive system composed by 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 3 levels: 1 (no problem) to 3 (severe problems). To generate a health profile, the levels of each dimension dimensions are combined (e.g. 11232). A single utility value, representing population-derived preferences for different health states, can then be attached to each profile. These utility values have been estimated using respondents in a number of different

countries; the UK index was developed by Dolan et al, which has a range from -0.594 to 1.00 (117) .

3.4.5 HRQoL Measures Specific to BD

As **Chapter 2** highlighted, there is ample evidence to indicate that HRQoL in BD patients is severely impaired (4). Whereas there is no single standard instrument used in BD, most of the times, HRQoL is captured by using a variety of generic questionnaires. Some of the commonly used instruments in BD are Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q), Short form 36 questionnaire (SF-36), WHO Quality of Life (WHO-QOL) and EQ-5D-3L(118). Of late, a relatively new condition specific measure: Quality of Life in Bipolar Disorder scale (QoL.BD) has been tested in clinical practice and research (88).

The disease-specific HRQoL instruments could be specifically designed for the disease and ensure a high sensitivity to small changes in patients' quality of life. However, these tools have a narrow scope, and it is not possible to carry out comparisons across different types of interventions or disease areas, which therefore limits their usefulness in the wider NICE decision-making context.

However, the assessment of HRQoL is particularly problematic in BD with no consensus on the right tool for measurement. BD presents with a heterogeneous symptomatology which makes it challenging to measure the resulting multi-factorial

impairment in HRQoL. HRQoL, as measured by generic measures such as EQ-5D-3L, may not always correlate well with symptom changes and its validity in mental illness has been challenged (29, 119). In the of field of BD HRQoL research, there is lack of research to support generic measures of HRQoL and it has been recommended that efforts should be made to develop disease-specific instrument to assess HRQoL (119). Nevertheless, given the ambiguity around the appropriateness of EQ-5D-3L **Chapter 6** will explore whether the EQ-5D-3L is appropriate to measure HRQoL in individuals with BD.

3.4.6 Utility measurement

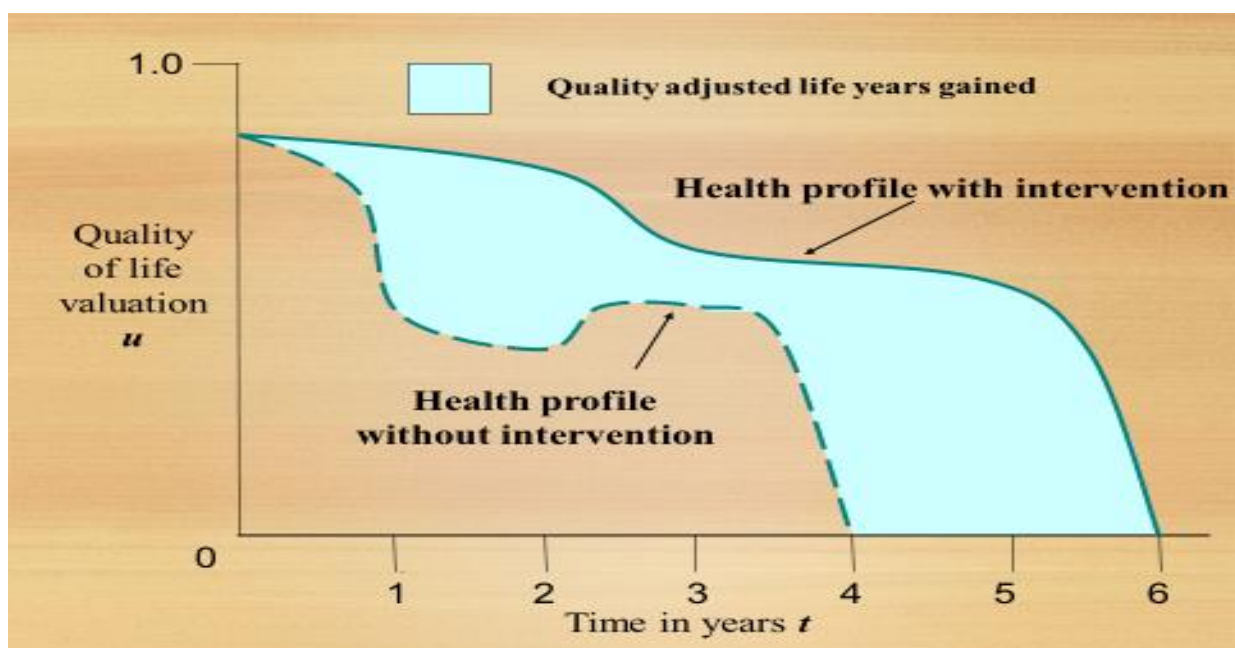
As previously reported, one of the advantages of using utilities is their ability to capture patients' satisfaction and preference for the possible health states regardless of the disease they are experiencing. In health economics, these preferences could be measured and combined with the effectiveness unit (e.g. life years) to reflect the individuals' preferences for a treatment. Utilities could be measured employing two methods: standard gamble (SG) and time trade off (TTO) (120).

3.4.7 Quality Adjusted Life Years (QALYs)

The quality adjusted life year (QALY) is the most comprehensive measure of benefit to compare interventions as it can combine both quality (quantified by utilities) and duration of survival into a single outcome measure. Once the utility values are obtained, these can be used to adjust the survival of the patients. **Figure 3.3** illustrates

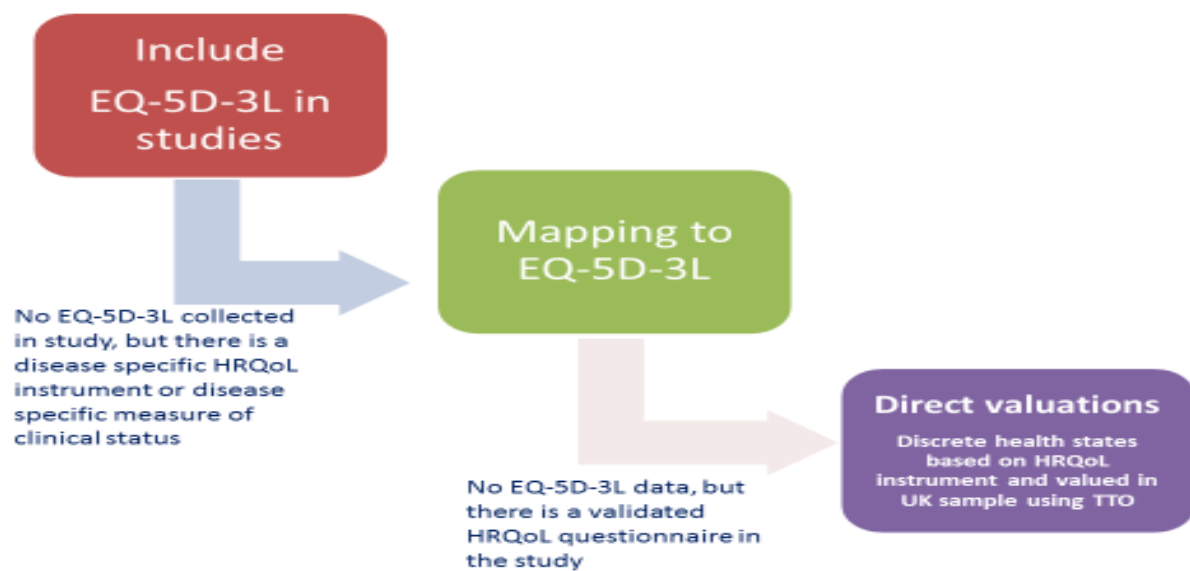
quality adjusted life profiles. The number of QALY's is calculated by multiplying a person's life expectancy by the value of the utilities obtained at each time-point. Utility index score as measured by the EQ-5D-3L indicates 0 as equivalent to death and 1 being in full health. Scores can be less than zero for health states regarded as worse than death. Being in depression, for instance, may be assigned a quality adjustment value of 0.5. A six-year period of living with depression on a particular treatment strategy is three QALY's and this is assumed to be equivalent to someone living for three years in full health. For more complex health profiles, involving transitions between states of health, the area under the survival adjusted curve (shaded area in **Figure 3.3** is estimated to determine the Quality Adjust Life Years (QALYs) accrued by each subject).

Figure 3.3 Comparison of health profile between intervention and no-intervention



QALYs are a fundamental part of the reference case for economic evaluations of treatments submitted to the NICE. Although, EQ-5D-3L is the preferred method of obtaining utility in the UK, in the absence of its inclusion, utilities can be mapped or derived using the TTO valuation method in population based samples (see **Figure 3.4**) (26).

Figure 3.4 Hierarchy of utility methods as recommended by NICE(26)

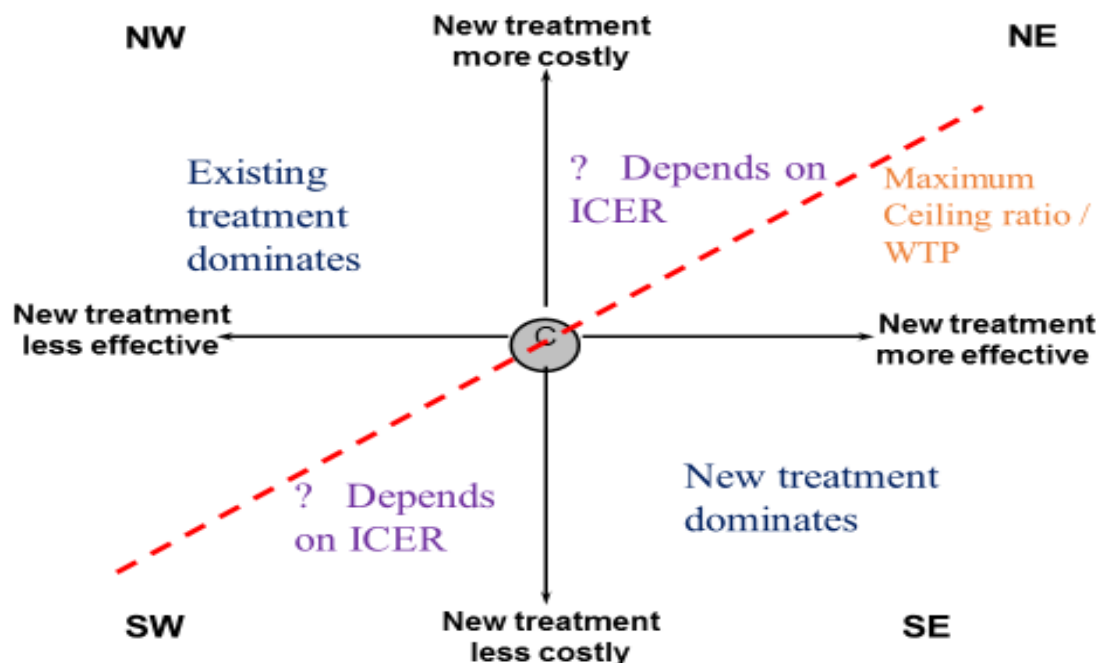


HRQoL: health-related quality of life; TTO: time trade-off

3.4.8 Incremental Cost Effectiveness Ratio (ICER)

In addition, to presenting the results of costs and outcomes separately, for decision making purposes, it is advisable to present results in a framework that encapsulates costs and effectiveness in a single index, also known as incremental cost-effectiveness ratio (ICER). ICER is the additional cost of achieving an additional unit of outcome (incremental costs divided incremental outcomes). **Figure 3.5** presents a graphical representation of the decision problem in cost-effectiveness analysis.

Figure 3.5 The cost-effectiveness plane



Costs and effects of two alternative strategies are compared: existing and new. In the northwest (NW) and southeast (SE) quadrants a case of dominance exists either in favour of or against the treatment under evaluation, and the decision rule is apparently clear. In the SE quadrant, the new strategy is less costly than the alternative and more effective. In the NW the new treatment is more costly and less effective than the old treatment. In the northeast and southwest quadrants, the decision is less clear, a trade-off has to take place. In the north east (NE) quadrant, the new treatment is more effective but also more costly. The ICER ratio allows decision makers to assess whether the joint distribution of cost and effectiveness for an intervention is cost-effective at a given willingness-to-pay (WTP) threshold or ceiling ratio. The WTP threshold adopted by NICE in UK is between £20,000 and £30,000 per QALY. Although this value may vary case by case, the probability of the intervention being

accepted is higher if the ICER is smaller than 20,000 per QALY. If the ICER is larger than £30,000 per QALY the probability of being accepted will be lower (120).

3.4.9 Timeframe and discounting

Economics evaluation could be performed alongside clinical trial, which limits the time-horizon to the trial timeframe. However, the results could be extrapolated over the trial length up to lifetime horizon. Adopting an appropriate time horizon is pivotal to make an unbiased decision. In this sense, the benefits of an intervention may not be relevant within the trial timeframe, but they could become significant over a lifetime horizon. Therefore, the discount rate plays an important role to reflect the preferences toward immediate health or financial gains, rather than those may occur in the future. The NICE set the current discount rate at 3.5% both for health outcomes and cost (121).

3.4.10 Uncertainty analysis

Although there are several classifications of uncertainty, in the context of decision-making, there are broadly two types of uncertainty; the uncertainty associated with an ICER and with the trade-off of between related costs and benefits that are acceptable to policy-makers (105).

ICER Uncertainty

The accuracy of the incremental cost-effectiveness ratio (ICER) relies on the costs and outcomes accuracy. However, the ICER is a difficult measure to deal with. First, in the case of equal effects between the treatment and the control group, the denominator would be equal to zero leading to an intractable result. Secondly, the ICER may adopt values which could have different meaning. Negative values could be due to a lower effectiveness of the innovative treatment or to lower cost of the innovative treatment. Likewise, positive values could be due to a higher incremental effectiveness and a higher cost of the innovative treatment; or due to a lower incremental effectiveness and lower cost of the innovative treatment. This problem makes it difficult to calculate confidence interval for the ICERs. Therefore, it is important to quantify uncertainty with adequate methods (e.g. non-parametric bootstrap, Monte Carlo simulation method) (122).

Non-parametric Bootstrap

ICER estimation is reliant on the cost and outcome variables; these are deterministic variables and any observed change in these variables can influence the ICER, and the resulting decision. It is therefore, important to subject the most uncertain variables to vary within a reasonable and plausible range of values to ascertain robustness of results. To get insight into the uncertainty surrounding the ICER, statistical techniques such as non-parametric bootstrapping are recommended in a trial setting. Briefly, bootstrapping is a technique involving a large number of repetitive draws from the study sample to empirically estimate the shape of the distribution (123).

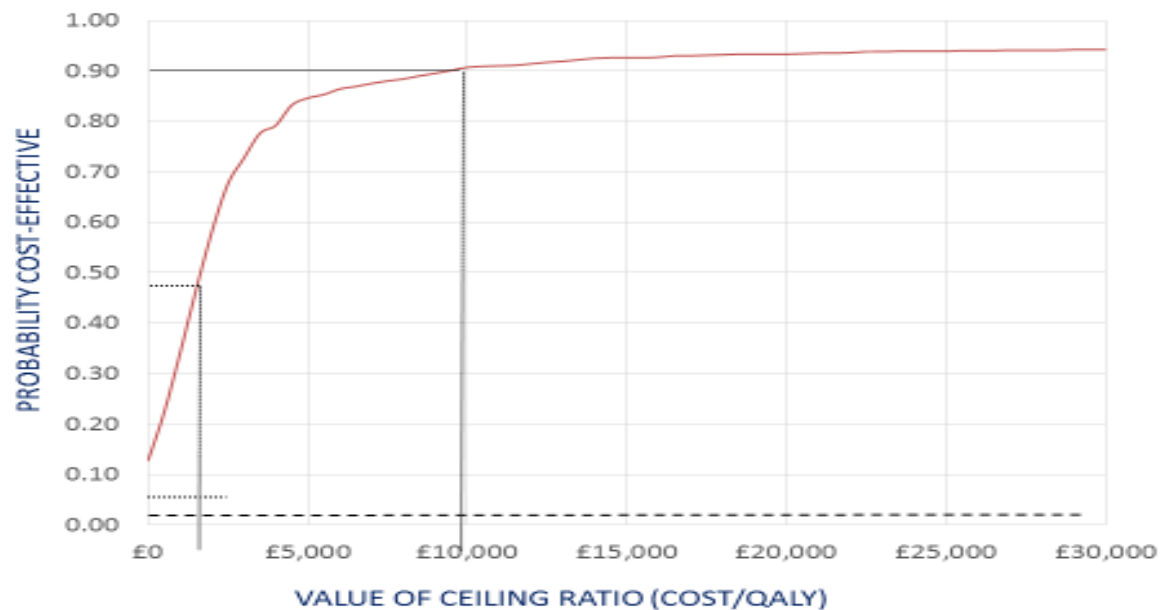
Probabilistic sensitivity analysis

A different approach is desirable for decision models. Probabilistic Sensitivity Analysis employs the Monte Carlo sampling methods to reproduce the results of the decision model while simultaneously taking into account the uncertainty around each parameter (124). Regardless of the employed method, a number of simulations for both cost and effectiveness are obtained. Plotting them on the cost-effectiveness plane will provide a cloud of ICER iterations that graphically represent the uncertainty around the result.

WTP uncertainty

From a decision maker's perspective, the probability that a new treatment is cost-effective varies depending on what society is willing to pay per unit of health gain (the so-called ceiling ratio). This information is provided by plotting cost-effectiveness acceptability curves (CEAC) (120). The CEAC is obtained by plotting the probability of the treatment being cost effective compared to an appropriate alternative at a given WTP threshold. The probabilities are calculated considering the number of ICER replicates which are inferior to the given WTP threshold. An example of a hypothetical CEAC curve is presented in **Figure 3.6** below; the x-axis gives the value of the ceiling ratio, and the y-axis provides the probability that the intervention is cost-effective. In this example, at a WTP of £10000 per QALY gained, there is a 90% that the hypothetical treatment is cost-effective.

Figure 3.6 Cost-effectiveness acceptability curves



3.5 Types of economic evaluation

This section aims to provide a brief overview on the most common economic evaluations in health care.

3.5.1 Cost-consequences analysis

Cost-consequences analysis (CCA) reports the cost and the consequences of the interventions assessed in a disaggregated way. The decision maker should combine and interpret the results. In mental health, there are several outcomes other than a change in health status that may be of interest to decision-makers, CCA can present costs and a range of outcomes without offering any judgement on which outcome is

better than another. The major limitation of CCA is the lack of interpretation and synthesis of the results. Besides, CCA assumes the decision-maker is capable of deciding the weight to put on the different outcomes. Because of these limitations, CCA in isolation is relatively uncommon (120). However, a combination of methods including CUA and CCA can provide clarity and an easier way of communicating with policy-makers (105).

3.5.2 Cost-minimisation analysis

Like CCA, cost-minimisation analysis also considers both cost and consequences of the treatments. However, in this case, it is assumed or there is evidence to prove that the effect between the two interventions is identical. Therefore, the option with the lower cost is preferred. Given the interest of health economists in the joint distribution of cost and effects, rather than the single cost or effectiveness distribution; this method could provide biased estimates. As a result, cost-minimisation analysis is not particularly used in health economics (125).

3.5.3 Cost-effectiveness analysis

Cost effectiveness analysis (CEA) consider costs and effects of two or more interventions. The incremental cost and effects are calculated and presented as ICER. The effectiveness of the treatments in CEA are calculated using disease-specific clinical outcomes. This limits the generalisability of the CEAs, and does not allow for comparison across different disease areas.

3.5.4 Cost–utility analysis

Cost–utility analysis (CUA) relies on the CEA fundamentals, but employs the QALY as measure of effectiveness. As explained in this chapter (please, see **Sections 3.4.5-3.4.7**), the QALY is based on the patients' survival adjusted by the HRQoL utility values. These values are obtained by generic questionnaires (e.g. EQ-5D-3L) which ensure the comparability of CUAs performed in different disease area. An empirical analysis using CUA is presented in **Chapter 7**.

3.5.5 Cost–benefit analysis

Both CUA and CEA do not place a monetary value on the outcomes, but they report which of the assessed intervention will bring the biggest benefit in term of health outcomes. Cost-benefit analysis (CBA) attempts to assess the monetary value related to an intervention. One of the advantages of CBA is to allow the comparison of treatment benefits beyond the health budget boundaries. For example, the monetary benefits of a CBA performed in the healthcare sector could potentially be compared with those obtained in the transport or education sectors. Even though there are apparent comparability advantages of CBA, the most common used EEs are CEA and CUA. This preference for CEA and CBA is mainly due to conceptual and practical problems in obtaining a monetary evaluation of life or quality of life (126). Further, the aversion of placing an explicit monetary value on health or on life has contributed to limit the usage of CBA among the EEs.

3.6 Frameworks for conducting economic evaluation

An economic evaluation can be based on a clinical trial or a decision analytic model synthesizing available evidence from several sources or a hybrid of both the methods.

3.6.1 Trial-based economic evaluations

The primary aim and scope of clinical trials are to investigate the safety and efficacy of interventions in a highly regulated setting or effectiveness in case of a pragmatic trial. They typically use a protocol driven experimental design to randomly allocate participants to different groups. This design is considered the “gold standard” for testing efficacy of interventions, and offers an opportunity to collect health economic data at a relatively low additional cost. Even though randomized controlled trials (RCTs), especially explanatory trials, appear a suitable vehicle for economic evaluations by reducing bias and confounding and offering high internal validity, their selective recruitment and protocolised follow-up compromises their external validity. However, economic evaluations conducted along with pragmatic trials offer an opportunity to measure cost-effectiveness in typical clinical practice, thereby allowing relatively easier implementation in ‘real-life’ settings (122).

Nevertheless, sample sizes obtained for RCTs are primarily based on clinical outcomes, and therefore, provide limited power to detect any significant differences in costs or cost-effectiveness (127). Data requirements for economic evaluations are also different from those of a clinical RCT. Outcomes explored in clinical trials are often

based on disease specific scales and are not generic enough for comparisons across different types of interventions or disease areas. NICE recommends the QALY as an appropriate measure of benefit to compare interventions. Trial-based economic evaluations are also limited by the duration of primary clinical trials, which in mental illness is usually for a year or a maximum of a few years. Such short time-frame may not be appropriate for an economic analysis of interventions that may have impact on diseases over a longer time horizon (128).

Although trial-based economic evaluations have limitations, they play an important role in producing reliable estimates of cost-effectiveness from individual patient-level data. This allows the use of different statistical and econometric techniques to explore the relationship between costs, outcomes and associated factors. Furthermore, trial-based evidence allows construction and validation of decision analytical models; these models can then be used as complimentary vehicles to extrapolate shorter-term trial-based cost-effectiveness information for chronic diseases (129).

3.6.2 Decision analytical models

Decision analytical models (DAMs) represent a simplified representation of real-world problems, and offer an opportunity to combine a range of evidence from various sources to inform the decision-making process in healthcare resource allocation. They facilitate comparison of all existing comparators, allow extrapolation of costs and effects over the life course of illness, and provide an iterative framework that could be harnessed to reduce uncertainty in real life decisions. The evidence required for DAMs may include, but are not limited to data on clinical effectiveness, baseline clinical

characteristics, adverse events, mortality, resource use, unit costs and utility estimates. The data sources can be as diverse as the “gold standard” meta-analysis of RCTs for clinical effectiveness to case studies or expert opinions for resource use (120).

There are different modelling methods including decision trees, Markov models, and discrete event simulations (DES). Their usefulness depends on the disease area under research, nuances associated with available data and resource constraints (129).

There are limitations to the usefulness of model-based economic studies including a lack of transparency in modelling assumptions, inappropriate structure (poor reflection of disease progression), need for advanced expertise (DES models), poor quality data and lack of rigor in obtaining evidence to populate models (130, 131). In order to be confident about the outcomes of models, it is paramount to address all of the four principal types of uncertainty, namely a) methodological, b) structural, c) parameter, and d) heterogeneity (132). However, if studies adhere to robust standards and good practice guidelines, DAMs could be useful complements and extensions to trial-based economic evaluations and pivotal for rational healthcare planning

3.7 Discussion and conclusion

This chapter introduced the subject of health economics together with specific details about cost-of-illness and economic evaluation. In the broader context of priority-setting

in mental health, the economic evidence consists of several important concepts, and within this chapter a focus has been given to: COI, cost, outcomes and utility, the ICER and the issue of uncertainty, the types of economic evaluation and frameworks for conducting an economic evaluation.

As outlined earlier in the chapter, COI and economic evaluation both are valuable tools in the health economist toolbox, especially in mental health as there is still a lack of 'parity of esteem with physical health (15). Estimating the burden of disease, and identifying cost-effectiveness interventions are complementary methods to build the economic case for BD.

COI and CUA will be seen in practice in **Chapters 5** and **7** respectively. The empirical work in these chapters will use frameworks described in this chapter. The COI is a prevalence-based bottom-up study using a prospective cohort of individuals with BD, and the CUA is conducted alongside an RCT. To address one of the methodological challenges highlighted earlier in the chapter, **Chapter 6** will provide evidence on the appropriateness of using EQ-5D-3L in individuals with BD.

Identification of cost-effective interventions in the field of BD represents one of the key considerations into the broader task of informing and assisting the decision-maker in allocating scarce health care resource. In the following chapter, fundamental concepts of economic evaluation detailed in this chapter are considered, to review and critique published economic evaluations of both pharmacological and psychological therapies in the management of BD.

Chapter 4: Economic evaluations in bipolar disorder: A systematic review and critical appraisal

4.1 Introduction

As **Chapter 2** highlighted, Bipolar disorder (BD) is a chronic illness and one of the leading causes of disability worldwide (89). The life expectancy of patients living with BD is less by ten years compared with the general population due to significantly higher levels of medical co-morbidities, substance abuse disorders, and increased suicide rates (90). BD is often a lifelong condition and incurs a substantial economic burden on the patients, health and social care system, and the wider society. However, few studies have evaluated the economic impact of competing strategies in the management of BD.

Chapter 3 elaborated the importance of economic evaluation. Recognising needs of a population in relation to a disease area is often the first step in understanding the problem statement of an illness. However, it is important to identify cost-effective management strategies that help reduce the burden for patients while providing rational choices for policy-makers (25). Resources are scarce, demonstrating effectiveness alone of treatments is no longer a viable option, it is critical to have evidence on those interventions that offer good value for money, and also have the potential to generate cost savings in health and social care (133). As **Chapter 2** and **3** outlined, providing cost-effectiveness interventions to individuals with BD will also

mean that greater costs elsewhere (i.e. productivity losses) beyond the healthcare and social care system are avoided.

Although, there are an increasing number of trials assessing efficacy of various therapies in BD, the literature is limited on economic evidence (134-138). Due to the dearth of economic studies, it has been proposed that attempts should be made to systematically collate existing evidence to facilitate decision making under uncertainty (11). A health technology assessment report published in 2007 was the only study to review the available economic evidence for the management of BD. This review was limited to literature prior to 2005 and managed to identify only three studies (139). In recent years, many innovations such as second-generation antipsychotics, combination therapies, collaborative care, psychological therapies and mood monitoring have been introduced in the management of BD. A recent comparative review on their cost-effectiveness, however, is lacking and an update of existing evidence is therefore required.

The aim of this chapter is to systematic review and critically appraise published literature on trial and model-based economic evaluations in the management of BD. The focus is on treatments where alternative management strategies exist for BD patients and include pharmacological, psychological and collaborative models of care. By reviewing and critically appraising the available published economic evidence on BD management, this chapter provides a strong evidence base for a range of management strategies in BD which could produce better outcomes at lower cost.

Drawing on the fundamental concepts of economic evaluation discussed in **Chapter 3**, the review comprises two sections with the first section focusing on economic evaluations conducted within or alongside a clinical trial. The second section presents modelling-based studies of alternative treatment strategies. The review starts with a description of the methodology employed.

4.2. Methodology

4.2.1. Search strategy

A comprehensive, systematic literature search was undertaken to identify all English language articles published between January 1980 and March 2012 that provided data on the economic evaluation of any management strategy for BD. The decision not to include earlier publications was taken to obtain data likely to be relevant to current healthcare settings and costs.

The recommendations of the “PRISMA statement” were followed in carrying out and reporting the process of the literature search and selection strategies (140). In total, seven electronic databases were searched: National Health Services - Health Economic Evaluation Database (NHS-HEED), Health Technology Assessment Database, Excerpta Medica database (EMBASE), MEDLINE, Pre-MEDLINE, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and PsychINFO. Additional searches were carried out in Google Scholar and Google search, and reference lists of key articles and recent reviews were checked for further relevant publications. Citations were also tracked using Google scholar. The search strategies

included both key words and corresponding medical subject headings (MeSH words). The details of individualised search strategies for each database are provided in **Appendix 1.**

4.2.2 Study selection criteria

The following selection criteria were applied to all identified studies:

- a) Only economic evaluations that compared two or more BD management strategies and considered both costs and consequences - i.e. cost-minimization analyses (CMA), cost-consequences analyses (CCA), cost-effectiveness analyses (CEA), cost-utility analyses (CUA), or cost-benefit analyses (CBA) - were included in the review. Studies that only considered costs, clinical effects (i.e. time to relapse or improvement in clinical status as measured by disease-specific instruments) or quality of life alone were not eligible.
- b) Other studies that limited the ability to assess the methodological quality of studies, e.g. letters, comments, conference abstracts or abstract publications were not considered.
- c) Studies that focused only on methodology or were literature reviews were excluded. However, primary studies from previous literature reviews that met the inclusion criteria were tracked for completeness.

4.2.3 Data extraction and quality assessment

Once a study was considered appropriate for appraisal (based on inclusion/exclusion criteria), data were extracted using a bespoke data extraction form detailed in **Appendix 2**. I extracted the data initially with a random subsample of studies undergoing double data extraction by a second researcher (Judit Simon). Disagreements were resolved by a third researcher (Jane Wolstenholme). To ensure that required data were uniformly and thoroughly collected, abstraction forms were pilot tested for four studies. The data abstraction forms included the following items: descriptions of study objective, study design, study population, study perspective, time horizon, interventions and comparators, data sources (baseline characteristics, clinical effect size, costs, measure of benefits), study results, sensitivity analyses, study limitations, generalizability and study quality scores as assessed by standard quality assessment checklists (see below).

Given the inconsistency in methodology, conduct and reporting of economic evaluations, methodological quality of the studies was also assessed. The quality of included studies was appraised as per recommendations of the Cochrane Collaboration (141). Trial-based evaluations were assessed using the Centre for Reviews and Dissemination's checklist (142) which is based on the commonly used Drummond's checklist (104). This checklist comprises 36 items covering three domains: a) study design, b) data collection, and c) analysis and interpretation of results. These dimensions cover aspects on study perspective, patient population, clear description of alternative treatment strategies, primary outcomes, measures taken to address uncertainty, and generalisability (checklist attached in **Appendix 3**).

The internal validity of trial-based economic evaluations depends on the extent to which estimates for economic, clinical and epidemiological inputs are rigorously measured and analysed utilising appropriate statistical techniques. On the other hand, external validity refers to generalizability of results to either a) a different group of individuals other than the study participants in similar study settings, or b) a different group of individuals other than the study participants in diverse settings (143).

For model-based economic evaluations, best practice guidelines proposed by Philips et al.(2004) was followed (132). This 61-item checklist broadly consists of three domains covering the a) structure, b) input data and c) consistency of a model, and includes 15 different dimensions for quality assessment (checklist attached in **Appendix 4**). Overall, the dimensions cover reflection of the disease process, type of model, the source of information used to populate the model, related assumptions, and uncertainty analyses.

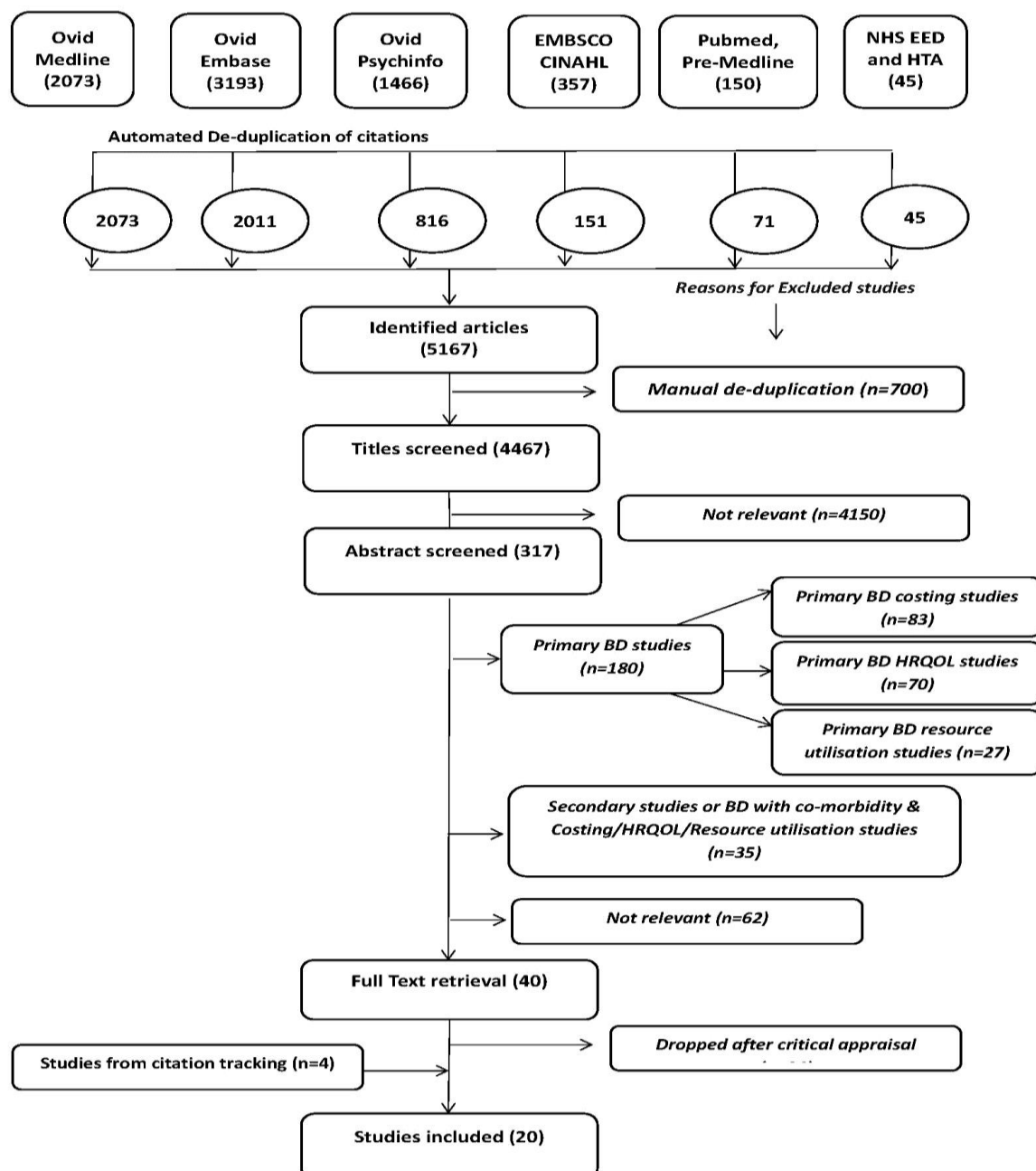
Items in both checklists were marked as 'Yes' (scored 1), 'No' (scored 0) or 'Not Applicable' (NA) against their respective quality criteria. Equal weights were used for all questions. Further, studies were classified as with a) high risk of bias: fulfilled < 50% of applicable quality criteria, b) medium risk of bias: fulfilled between 50 and 80% of applicable quality criteria, and c) low risk of bias: fulfilled >80% of quality criteria.

4.3 Results

A flowchart of the search and selection strategies is shown in **Figure 4.1**. Seven electronic databases were systematically searched yielding a total of 7284 citations.

Many of the citations were duplicate records found in different databases. Endnote bibliographic management software was utilized to match records that refer to the same article in multiple databases. This process of automated de-duplication of citations yielded a total of 5167 citations.

Figure 4.1: Flowchart of search strategy and included studies

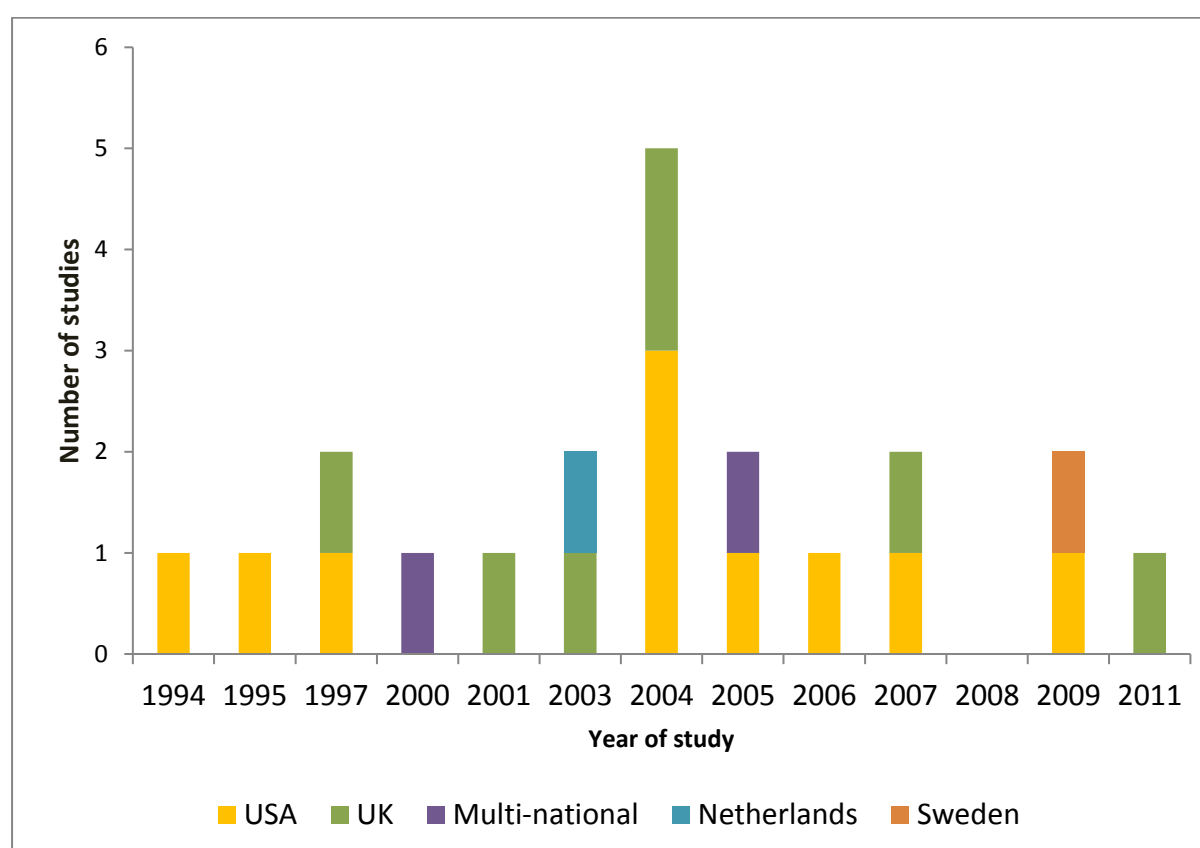


Out of 5167 citations, a further 700 duplicate publications were removed manually. After a review of the titles of the remaining 4467 papers, 317 articles were selected as potentially relevant and their abstracts screened. 277 abstracts were discarded as they investigated either costs or outcomes in isolation, or studied mental disorders as a whole making it impossible to identify BD-related costs and outcomes, or the studies were reviews, commentaries and expert opinions.

Full texts were retrieved for 40 articles. Citation tracking identified four more potential publications. Out of the final 44 studies, 20 studies met all the inclusion criteria. Data extraction and quality appraisal was performed for these papers. **Appendix 5** presents the list of excluded studies along with the reasons for their exclusion. Tables containing list of studies evaluated against explicit quality criteria are provided in **Appendix 6 and 7**.

Of the 20 eligible studies, 9 (45%) were from United States and 7 (35%) from the UK. Two studies (10%) were multinational, comparing different WHO defined regions including developing and developed countries. The remaining two studies (10%) were from the Netherlands and Sweden (**Figure 4.2**). Pharmaceutical management strategies were the most commonly investigated interventions (n=14, 70%), followed by collaborative care or system-level interventions (n=4, 20%). One further study looked at adjuvant psychosocial intervention and one study evaluated a screening strategy. In terms of disease states, 15 studies (75%) focused on prophylactic treatments.

Figure 4.2: Study setting and year of study



Two studies (10%) considered the treatment of acute mania, one study evaluated interventions for both acute and maintenance phases and one study evaluated a screening strategy for BD. Considering methodology, five studies (25%) were trial-based economic evaluations and 15 studies (75%) were model-based economic evaluations. Overall, four studies undertook CCA, six undertook CEA, five undertook CUA, and five studies performed both CEA and CUA. Details of the individual studies including participant and methodological characteristics and results are provided in **Tables 4.1 and 4.2**, and are summarised below by trial or model-based economic evaluations.

Table 4.1 Summary of methods and conclusions of included trial-based economic evaluations

Study/ Year	Setting/ funding	Sample	Clinical Study design	Comparators		Economic study design	Costs	Outcome	Results	Conclusions
				Intervention	Alternative					
Revicki et al, 2003	US, Multicenter	18-65 with BD-I who were hospitalized for acute	12 week double blinded RCT	OLZ 10-20 mg/day	DVP 20 mg/kg/day	Design: CCA	\$US	a) SADS-MRS scores	No statistically significant differences noted between the two groups for primary out comes. Economic outcomes: No significant difference between the two groups for overall medical costs (p=0.88). Mean outpatient costs was significantly lower for DVP than OLZ (\$541 Vs. \$638, P=0.004).	DVP is similar in clinical outcomes and HRQOL to OLZ at a lower outpatient cost. Risk of bias: High Fulfilment of quality criteria: 13/18 /5 (42%)
	Secondary care	Mania, and SADS-MRS score >25		N=25 patients	N=27 patients	Perspective: Third party payer	Cost year NR	b) Hamilton rating scale for depression score		
	Funding: Abbott Laboratories					Intervention: Medicines	Direct	c) HRQOL measured by Q-LES- Q		
Revicki et al, 2005	US, Multicenter	≥18 with BD-I and who were hospitalized for acute	Open label Pragmatic RCT; Follow up: 12 months	LI 1800 mg/day (acute phase) or LI=900-1200mg/day during remission	DVP 15 to 20 mg/kg/ day	Design: CCA	\$US	a) Number of BD free months	No statistically significant differences noted between the two groups for primary out comes and costs.	DVP is comparable in costs and clinical outcomes to LI Risk of bias: Medium Fulfilment of quality criteria: 16/14 /6 (53%)
	Secondary care	Mania or mixed episode		N=97 patients	N=104 patients	Perspective: NR	Cost Year 1997	b) MCS scores of SF-36		
	Funding: Abbott Laboratories					Intervention: Medicines	Direct			
						Sensitivity analysis: NR				

Study/ Year	Setting/ funding	Sample	Clinical Study design	Comparators		Economic study design	Costs	Outcome	Results	Conclusions
				Intervention	Alternative					
Lam et al, 2005	London, UK Secondary care Funding: South London and Maudsley Hospital's research and development fund	18-70 with BD-I with at least two episodes in the previous 2 years or three episodes in the previous 5 years.	RCT, Follow up: 12 months and 30 months	CBT+ Care as usual N=51 patients	Care as usual N=53 patients	Design: CCA Perspective: NHS Intervention: Psycho-logical+ Care as usual Sensitivity analysis: Univariate and PSA	£GBP Cost Year 1999/2000 Direct	Number of BD free days	Post 12 months: CBT group spent 62.3 fewer days with BD than the care as usual group. Post 30 months: CBT group spent 110 fewer days with BD than the comparator. Economic outcomes: Post 12 months: probability of CBT more cost-effective than standard treatment was more than 85%. Post 30 months: 80 %	Adding CBT to standard care is more cost-effective than standard care alone Risk of bias: Medium Fulfilment of quality criteria: 16/10 /4 (69%)
Bauer et al, 2006	US, Multicenter Secondary care: VAMC Funding: Veterans Affairs	≥18 and who were hospitalised for BD	Single blinded RCT, Follow up: Three years	Collaborative care delivered by 0.5 FTE nurse specialist & a 0.25 FTE psychiatrist Components a) Support providers b) Group Psycho-education c) feedback N=166 patients	Care as usual N=164 patients	Design: CCA Perspective: Healthcare provider Intervention: Service level Sensitivity analysis: NR	\$US Cost Year 2004 Direct	a) Number of weeks in manic or depressive episode b) Mean ratings of mania and depression using the 6-point PSR scale c) Generic HRQoL as measured by SF36 every 24 weeks	1) 6.2 fewer weeks in an affective episode over three years compared with usual care, including 4.5 fewer weeks of manic episodes. No differences observed in the duration of depressive episodes. 2) No statistically significant differences in mean depression or manic PSR scores. 4) MCS scores from SF-36 were significantly higher in intervention arm than the care as usual group (37.6 Vs. 34.1, P value=0.01).	Collaborative care provides 4.5 additional weeks free of Manic episodes at no extra cost Risk of bias: Medium Fulfilment of quality criteria: 18/15 /3 (55%)

Study/ Year	Setting/ funding	Sample	Clinical Study design	Comparators		Economic study design	Costs	Outcome	Results	Conclusions
				Intervention	Alternative					
									PCS scores were similar in both groups.	
									Economic outcomes: Total average treatment costs were not statistically different between the groups.	
Simon et al, 2006	US, Multicenter	Patients aged ≥18 with any diagnosis of bipolar spectrum disorder during the prior 12 months	Single blinded RCT, Follow up: Three years	Collaborative care delivered by 3 nurse specialists	Care as usual included individual psychotherapy, or group psychotherapy, but no coordinated care	Design: CEA Perspective: Healthcare provider Intervention: Service level Sensitivity analysis: NR	\$US Cost Year NR Direct	a) Mean mania and depression scores on the 6-point Psychiatric Status Rating (PSR) over 104-weeks b) Weeks spent in Mania or depression as defined by PSR scales	Whereas the mean mania PSR scores were significantly lower in the intervention group (z=2.09, P=.04), there were no statistically significant differences noted between the two groups for mean depression PSR scores. At an incremental cost of \$1251 (95% confidence interval, \$55-\$2446), people in collaborative care had 5.5 additional weeks free of severe manic symptoms than the control group	At an incremental cost of \$1251, collaborative care program can significantly reduce the frequency and severity of mania in BD Risk of bias: Medium Fulfilment of quality criteria: 17/16 /3 (52%)

*Centre for reviews & dissemination checklist for assessing economic evaluation; Abbreviations: BD=Bipolar disorder, CBT= Cognitive behavioral therapy, CCA=cost consequences analysis, CEA= cost-effectiveness analysis, DVP=Sodium Valproate, FTE=Full time equivalent, HMO: health maintenance organization, HRQOL=health related quality of life, LI=Lithium, MCS=mental component summary, NHS=National health service, NR= not reported, OLZ=Olanzapine, PSR =Psychiatric status rating, Q-LES-Q= Quality of Life Enjoyment and Satisfaction Questionnaire, RCT= randomized controlled trial, SADS-MRS=Schedule for Affective Disorders and Schizophrenia- Mania Rating Scale, SF36=Short form-36, VAMC= Veteran affairs medical center, YMRS: Young Mania Rating scale

Table 4.2: Summary of methods and conclusions of included model-based economic evaluations

Study/ Year	Setting/ funding	Sample	Primary source of data	Comparators		Economic study design	Costs	Outcome	Results	Conclusion
				Intervention	Alternative					
Keck et al, 1996	US Second ary care Funding : Abbott Labs	Simulated cohort of 10000 patients with BD-I hospitalis ed for acute mania	Clinical parameters: RCTs , and expert opinion Costs: physicians Fee, 'Red Book', admin records Utility: NA	LI	Placebo	Study type: CEA Design: DAM Intervention: Pharmaceuticals Perspective: TPP Time horizon: one year Sensitivity analysis: Univariate	\$US Cost Year 1994 Direct	The initial response rate to therapy & rate of side effects	Initial hospital length of stay is 18.4 days (LI) and 14.3 days (DVP). Initial response rate is 0.49 (LI), and 0.59 (DVP). Relapse rate is 0.56 (LI), 0.56 (DVP). Mean total costs were \$43,400 and \$39,643, respectively, for LI and DVP	<i>DVP is cost- effective in comparison with LI Risk of bias: Medium Fulfilment of quality criteria: 25/20/17 (51%)</i>
Bridle et al, 2004	UK Second ary care Funding : NHS R&D	Patients with an episode of acute mania	Clinical parameters: MTC of RCTs Costs: BNF, HES, published cost of illness study Utility: NA	a) LI b) HLP c) OLZ d) DVP e) QTP	No main- tenance treatment	Study type: CEA Design: DAM Intervention: Pharmaceuticals Perspective: UK NHS Time horizon: three weeks Sensitivity analysis: Univariate & PSA	£GBP Cost Year 2001- 2002 Direct	≥ 50 % improve- ment in a patient's baseline YMRS score	In the deterministic analysis, LI, DVP, and QTP were dominated by HLP. In comparison with HLP, OLZ had an ICER of £7179 per additional responder. At a WTP of £10,000, OLZ has a probability of being 43 % cost-effective.	<i>HLP is the most cost effective therapy if WTP is < £7179/ additional responder. If the WTP is >£7179 /additional responder then OLZ is the most cost-effective option</i>

Study/ Year	Setting/ funding	Sample	Clinical Study design	Comparators		Economic study design	Costs	Outcome	Results	Conclusions
				Intervention	Alternative					
										<i>Risk of bias:</i> <i>Medium</i> <i>Fulfilment of</i> <i>quality criteria:</i> 24/19/29 (57%)
<i>Chisholm et al, 2005</i>	Global 13 WHO sub- regions Primary/ seconda ry care Funding : WHO	Mathemat ical model of different stages in the course of BD	Clinical parameters: GBD 2000 , literature and Delphi consensus Costs: various resources, WHO- CHOICE project Utility: GBD 2000, Dutch utility weights	a) LI 1200 mg b) LI + psycho- social care c) DVP and d) DVP + psycho- social care	a) LI 900 mg b) OLZ 12.5 mg c) PL	Study type: CUA Design: Mathematical model Intervention: Pharmaceuticals/ psychosocial care/system Perspective: Healthcare provider Time horizon: lifetime Sensitivity analysis: Univariate	\$US (adjust ed for PPP) Cost Year 2000 Direct	Cost per DALY averted. Community based treatment was more cost effective than hospital based care. The most cost effective treatment was LI with adjuvant psychosocial therapy with an average of \$2165– 3830/ DALY averted in high-mortality developing sub regions, \$3953–6475 in low-mortality developing sub-regions, and \$5487–21123 in developed sub-regions	<i>Community based adjuvant LI with psychosocial therapy was the most cost- effective option</i> <i>Risk of bias:</i> <i>Medium</i> <i>Fulfilment of</i> <i>quality criteria:</i> 30/20/12 (62%)	
<i>Calvert et al, 2006</i>	US Second ary care Funding : GSK	Simulated cohort of 1000 BD-I patients aged ≥ 18 years who start in an euthymic state	Clinical parameters: metaanalysis of 3 RCTs using direct and indirect comparison/ literature Costs: various	LAM 200 mg/d	No treatment	Study type: CEA/CUA Design: Markov Model Intervention: Pharmaceuticals Perspective: TPP Time horizon:18 months	\$US Cost Year 2004 Direct	a) Number of episodes, b) Number of euthymic days c) QALYs	LAM dominated PL &OLZ for all the three measures of benefit. The ICER of LAM Vs. LI was \$2,400 per episode avoided, \$30 per euthymic day gained, and \$26,000 per QALY. The ICER of OLZ Vs. LI was \$200 per euthymic day gained, \$7,000 per	<i>Using conventional WTP thresholds, LAM is the most cost-effective strategy</i> <i>Risk of bias:</i> <i>Medium</i>

Study/ Year	Setting/ funding	Sample	Clinical Study design	Comparators		Economic study design	Costs	Outcome	Results	Conclusions
				Intervention	Alternative					
			sources, 'red book' and physician survey Utility: literature			Transition period: quarterly cycles Sensitivity analysis: Univariate & Bivariate			acute episode avoided, and \$374,500 per QALY.	<i>Fulfilment of quality criteria: 39/11/12 (76%)</i>
<i>NICE economic Model, 2006</i>	UK Second ary care Funding : NICE	Simulated cohort of 1000 BD-I patients aged ≥18 years and who start in an euthymic state	Clinical parameters: metaanalysis of RCTs, longitudinal study, GDG opinion, literature Costs: various standard sources, BNF Utility: literature and assumptions	a) LI 1000 mg b) DVP 1250 mg c) OLZ 10 mg	PL	Study type: CEA/CUA Design: Markov Model Intervention: Pharmaceuticals Perspective: UK NHS Time horizon: 5 years Transition period: monthly cycles Sensitivity analysis: Univariate & PSA	£GBP Cost Year 2004-2005 Direct	a) Number of acute episodes averted b) Number of days free from acute episodes. c) QALYs	LI is likely to be a cost-effective long-term treatment option for patients with BD-I when WTP for an additional episode averted or an additional day free from episode is below £17,500 and £150 respectively. There is 90 % probability that OLZ is highly likely to be the most cost-effective option among all treatments assessed for all patients at a WTP threshold of £20,000/QALY.	<i>Using conventional WTP thresholds, OLZ is the most cost-effective strategy Risk of bias: Medium Fulfilment of quality criteria: 41/15/6 (73%)</i>
<i>Klok et al, 2007</i>	Netherla nds Second ary care	Simulated cohort of 10000 patients with BD-I who were hospital-ised for	Clinical parameters: pooling of data from RCTs / literature/ author's assumptions	a) QTP 600 mg b) LI 800 mg c) DVP 1000 mg	LI 300 to 1800 mg/d	Study type: CEA Design: DES Intervention: Pharmaceuticals Perspective: Healthcare provider Time horizon: 100 days	€ Euros Cost Year 2003 Direct	a) Time to remission b) Time to response c) Length of stay d) Proportion	Combination therapy options were cheaper and more efficacious than PL & monotherapies. QTP+LI was costlier than OLZ+LI & RIS+LI, but was associated with	<i>QTP+LI is more cost-effective than combination of LI with OLZ or RIS Risk of bias: High</i>

Study/ Year	Setting/ funding	Sample	Clinical Study design	Comparators		Economic study design	Costs	Outcome	Results	Conclusions
				Intervention	Alternative					
	Funding : Astra-Zeneca	acute mania	Costs: Dutch costing guidelines and drug tariffs Utility: NA	d) QTP + LI 600mg +800 mg e) OLZ + LI 15 mg+800 mg F) RIS + LI 6 mg+ 800 mg		Sensitivity analysis: NR		≥ 50 % improvement in a patient's baseline YMRS score e) Number of serious side effects	highest proportion of responders (54.7%, P <0.05). QTP+LI generated an ICER of €1203/ serious adverse event prevented against RIS+LI, and an ICER of €3481/serious adverse event prevented in comparison with OLZ+LI combination.	<i>Fulfilment of quality criteria:</i> 19/20/23 (49%)
<i>McKendrick et al, 2007</i>	UK Secondary care Funding : Eli Lilly	Simulated cohort of 1000 BD-I patients aged ≥ 18 years who start in an euthymic state	Clinical parameters: Synthesis of results from RCT and literature Costs: various sources, literature and clinical opinion Utility: NA	OLZ = 5 to 20 mg/d	No treatment	Study type: CEA Design: Markov Model Intervention: Pharmaceuticals Perspective: UK NHS Time horizon: 12 months Transition period: monthly cycles Sensitivity analysis: Univariate & PSA	£GBP Cost Year 2003 Direct	a) Number of acute mood episodes, and b) Number of days in hospital c) Hospitalizations due to an acute mood event	OLZ significantly reduced the annual number of acute mood episodes per patient from 0.81 to 0.58 (95% CI: -0.34-0.12) at a lower incremental annual cost of £799 (95% CI: -£1824- £59). OLZ is more effective at a lower average cost.	<i>OLZ is more cost-effective than LI in preventing the number of acute mood episodes in patients.</i> <i>Risk of bias: Medium</i> <i>Fulfilment of quality criteria:</i> 33/16/13 (67%)
<i>Soares-Weiser et al, 2007</i>	UK Secondary care	Simulated cohort of 1000 patients	Clinical parameters: MTC of RCTs	a) LI: 800 mg b) DVP: 1250 mg	PL + LI or DVP combination	Study type: CUA Design: Markov Model	£GBP Cost Year	a) Number of acute episodes averted	1) <u>Recent depressive episode</u> ; DVP was the cheapest non-dominated option. Against DVP, LI	<i>Using conventional WTP thresholds, LI+IMI is the</i>

Study/ Year	Setting/ funding	Sample	Clinical Study design	Comparators		Economic study design	Costs	Outcome	Results	Conclusions
				Intervention	Alternative					
	Funding : NIHR HTA program	with stabilised BD-I following an acute episode	Costs: various standard sources NICE guidelines, HES, expert opinion, and BNF Utility: literature and author's assumptions	c) OLZ: 10 mg d) CBZ: 600 mg e) LAM: 200 mg f) IMI: 150 mg g) LI+IMI		Intervention: Pharmaceuticals Perspective: UK NHS Time horizon: 60 years Transition period: annual Sensitivity analysis: Univariate & PSA	2004-2005 Direct	b) Number of days free from acute episodes. c) QALYs	generated an ICER of £10,409 /QALY. Against LI, LI+IMI generated an ICER of £21,370 /QALY. PSA indicated 41% probability for LI and 53% for LI+IMI being cost-effective at a £30,000 threshold. 2) <u>Recent manic episode</u> : OLZ and LI dominated rest of the strategies. In comparison with OLZ, LI generated an ICER of £ 11,359 / QALY. At £20,000, there is 77% likelihood of LI being the most cost-effective option.	<i>most cost-effective strategy for patients with a recent depressive episode. For patients with a history of manic episode, LI is the most cost-effective strategy</i> <i>Risk of bias: low</i> <i>Fulfilment of quality criteria: 50/11/1 (82%)</i>
<i>Fajutrao, et al, 2009</i>	UK Second ary care Funding : Astra-Zeneca	Simulated cohort of 1000 BD-I patients ≥ 18 years in an euthymic state	Clinical parameters: metaanalysis of RCTs, expert opinion and literature Costs: national sources, and BNF	QTP, 400 to 800 mg/d + LI or DVP	No screening	Study type: CEA/CUA Design: Markov Model Intervention: Pharmaceuticals Perspective: UK NHS Time horizon: 24 months Transition period: quarterly cycles	£GBP Cost Year 2007 Direct	a) Number of acute mood episodes, b) Hospitalisations due to an acute mood event c) Costs per QALY	QTP + LI or DVP combination dominates PL+ LI or DVP combination. When compared with PL+LI or DVP, the QTP + LI or DVP combination generates an ICER of £506 per additional acute mood episode prevented, £4261 for an addition hospitalization prevented,	<i>QTP+LI/DVP combination is more cost-effective than LI or DVP alone</i> <i>Risk of bias: Medium</i> <i>Fulfilment of quality criteria: 41/17/4 (71%)</i>

Study/ Year	Setting/ funding	Sample	Clinical Study design	Comparators		Economic study design	Costs	Outcome	Results	Conclusions
				Intervention	Alternative					
			Utility: literature			Sensitivity analysis: Univariate & PSA			and -£7453 per additional QALY gained.	
Menzin et al, 2009	US Primary care: MCO Funding : Astra-Zeneca	Simulated cohort of 1000 patients >18 with first episode of MDD	Clinical parameters: literature, authors calculations and expert opinion Costs: various standard sources and literature Utility: NA	Screening using self-administered questionnaire : MDQ	a) PL + LI/DVP combination	Study type: CEA Design: Decision Tree Intervention: Screening Perspective: TPP Time horizon: 5 years Sensitivity analysis: Univariate & PSA	\$US Cost Year 2006 Direct	Total number of correctly diagnosed patients with BD	One time screening with MDQ detects 38 additional patients with depression or bipolar depression (440 with screening Vs. 402 without screening) among 1000 patients presenting with first time symptoms of a MDD in primary care. When compared with not screening arm, the screening arm generated a five year/patient discounted cost saving of \$1937 (\$34,107 with screening Vs. \$36044 without screening).	One time screening program for BD in patients presenting with symptoms of MDD is more cost-effective than no screening. Risk of bias: Medium Fulfilment of quality criteria: 28/16/18 (66%)
Woodward et al, 2009	US Secondary care Funding : Astra-Zeneca	Simulated cohort of 1000 stabilized BD-I patients who start in an euthymic state	Clinical parameters: Synthesis of results from RCT, literature and expert opinion Costs: various sources, 'red	QTP (400-800 mg/d) plus LI (0.5-1.2 mEq/L) or DVP (50-125 µg/mL)	a) PL + LI/DVP combination b) No maintenance treatment c) LI d) LAM 200 mg/d e) OLZ 12.5 mg/d	Study type: CEA/CUA Design: Markov Model Intervention: Pharmaceuticals Perspective: TPP & societal Time horizon: 24 months Transition period: quarterly cycles	\$US Cost Year 2007 Direct	Costs per QALY	QTP + LI or DVP combination dominates (i.e. cost saving and produces more QALYs) Placebo + LI/DVP. While the QTP + LI or DVP combination generates 1.49 QALYs at a cost of \$US 12,930, Placebo + LI/DVP generated 1.44 QALYs at a cost of \$US	QTP+LI/DVP combination is more cost-effective than LI or DVP alone Risk of bias: Medium Fulfilment of quality criteria: 35/23/4 (62%)

Study/ Year	Setting/ funding	Sample	Clinical Study design	Comparators		Economic study design	Costs	Outcome	Results	Conclusions
				Intervention	Alternative					
			book', and literature Utility: literature		f) ARP 30 mg/d	Sensitivity analysis: Univariate & PSA			12937 over a two years period.	
Woodward et al, 2010	US Second ary care Funding : Astra- Zeneca	Simulated cohort of 1000 BD-I patients average aged 40 years and who start in an euthymic state	Clinical parameters: pooled results from RCTs, literature and expert opinion Costs: various sources, 'red book', and literature Utility: literature	a) QTP XR - 600mg/d + LI (1100 mg/d) or DVP (1220 mg/d)	ARP: 15 mg	Study type: CEA/CUA Design: Markov Model Intervention: Pharmaceuticals Perspective: TPP & societal Time horizon:24 months Transition period: quarterly cycles Sensitivity analysis: Univariate & PSA	\$US Cost Year 2009 Direct & produc tivity losses	a) Number of acute mood episodes, b) Hospitali- zations due to an acute mood event c) Costs per QALY	QTP-XR + LI or DVP dominates LAM, OLZ, ARP monotherapies and PL. When compared with PL+LI or DVP, QTR-XR + LI or DVP generates an ICER of \$US22 959/ QALY & dominates from a direct & societal perspective respectively. When compared with LI, QTP-XR + LI or DVP generates an ICER of \$US100 235/QALY, and ICER of \$US 81712 /QALY from a direct & societal perspective respectively	Using conventional WTP thresholds, LI monotherapy is the most cost- effective strategy Risk of bias: Medium Fulfilment of quality criteria: 41/16/5 (74%)

Study/ Year	Setting/ funding	Sample	Clinical Study design	Comparators		Economic study design	Costs	Outcome	Results	Conclusions
				Intervention	Alternative					
Kasteng <i>et al</i> , 2011	Sweden Second ary care Funding : Bristol- Myers Squibb	Simulated cohort of patients with BD with or without metabolic syndrome	Clinical parameters: Synthesis of results from RCTs & literature Costs: various sources, databases, expert opinion & literature Utility: literature	OLZ :10 mg	No main- tenance treatment	Study type: CUA Design: Markov Model Intervention: Pharmaceuticals Perspective: Societal Time horizon: lifetime Transition period: annual Sensitivity analysis: Univariate & PSA	SEK Cost Year 2009 Direct & produc tivity losses	a) Life years gained b) QALYs	From a societal perspective, ARP combination dominates OLZ. When compared with OLZ, ARP generates an incremental cost saving of 28447 SEK while producing 0.04 additional Life years 0.09 additional QALYs gained.	<i>With a particular focus on metabolic side effects of treatments including CHD & DM, ARP is more cost effective than OLZ Risk of bias: Medium Fulfilment of quality criteria: 41/14/7 (75%)</i>
Chisholm <i>et al</i> , 2012	Sub- Saharan Africa & South East Asia Primary/ Second ary care Funding : WHO	standard populatio n of one million people	Clinical parameters: GBD 2004 , literature and Delphi panel Costs: various resources, WHO- CHOICE project Utility: GBD 2004	a) LI (1200 mg) b) LI + psycho- social care c) DVP and d) DVP + psycho- social care	Placebo	Study type: CUA Design: Mathematical model Intervention: Pharmaceutical/psyc ho-social care/ Service level Perspective: Healthcare provider Time horizon: lifetime Sensitivity analysis: Univariate & PSA	US \$s (adjust ed for PPP) Cost Year 2000 Direct	Cost per DALY averted.	LI and LI + psychological care dominated rest of the strategies.	<i>At a national GDP/capita threshold of \$2000, use of LI in community settings was found be "very cost effective" Risk of bias: Medium Fulfilment of quality criteria: 32/18/12 (64%)</i>

Study/ Year	Setting/ funding	Sample	Clinical Study design	Comparators		Economic study design	Costs	Outcome	Results	Conclusions
				Intervention	Alternative					
Ekman et al, 2012	UK Second ary care Funding : Astra- Zeneca	Simulated cohort of 10000 patients with BD	Clinical parameters: metaanalysis of RCTs, literature and a cohort study Costs: various sources, guidelines, HES, expert opinion, & BNF Utility: literature & assumptions	a) QTP 300mg b) QTP 600mg c) QTP+MS d) OLZ e) OLZ + FLU f) OLZ+MS g) ARP h) RIS: i) LI j) LAM k) DVP l) PAR + LI m) VEN + LI		Study type: CUA Design: DES Intervention: Pharmaceuticals Perspective: UK NHS Time horizon: 5 years Sensitivity analysis: Univariate & PSA	£GBP Cost Year 2011 Direct	a) Number of mood events including depressive & manic episodes b) QALYs	<u>Starting in depressive episode:</u> QTP was more cost effective than OLZ in all phases of BD generating an ICER of £8600/QALY. Comparing with other strategies, QTP remained dominant treatment in majority of situations depending on the phase of BD. <u>Starting in Remission/stable state:</u> QTP was more cost effective than OLZ with an ICER of £27400/QALY. QTP dominated other strategies most of the times.	<i>Using conventional WTP thresholds, QTP is more cost-effective than OLZ in the maintenance treatment of BD & Bipolar depression. Risk of bias: low Fulfilment of quality Criteria: 50/3/9 (94%)</i>

* ARP= Arpiprazole, BNF=British national formulary, CBZ= Carbamazepine, CHD= coronary heart disease, CHOICE= CHOosing Interventions that are Cost Effective, CUA= cost utility analysis, DALY= disability adjusted life years, DAM=decision analytical model, DES= discrete event simulation, DM= diabetes mellitus, DVP=Sodium Valproate, FLU=Fluoxetine, GBD= global burden of diseases, GDG= guideline development group, HES=hospital episode statistics, IMI=Imipramine, HLP=haloperidol, ICER= incremental cost-effectiveness ratio, LAM=Lamotrigine, LI=Lithium, MDD= major depressive disorder, MDQ= mood disorder questionnaire, MTC=mixed treatment comparison, NA=not applicable, NHS R& D HTA program = national health service research and development health technology assessment program, OLZ=Olanzapine, PAR=Paroxetine, PL=placebo, PPP=purchasing power parity, PSA=probability sensitivity analysis, QALY= quality adjusted life years, QTP=Quetiapine, QTP-XR= Quetiapine extended release, RCT=randomized controlled trial, RIS= Risperidone, TPP=third part perspective, VEN=Venlafaxine, WTP=willingness-to-pay, YMRS=Young mania rating scale WHO= World health organization

4.4 Trial-based economic evaluations

4.4.1 Study design

Of the five trial-based economic evaluations, four of the studies were CCA (144-147) where costs and effects are presented in a disaggregated manner and usually a range of outcomes are reported alongside costs. One study was a CEA, where costs were presented in an aggregated form alongside a single disease-specific outcome measure, BD free days. There were no CUAs where the outcome is presented as QALYs or disability adjusted life years (DALYs) (148). All but one study (145) failed to state explicitly their type of economic evaluation, and none of the studies justified the type of economic analysis they undertook. Two studies focused on pharmacotherapies (144, 146), two on system-level collaborative interventions (147, 148), and one investigated adjuvant psychotherapy (145).

The two collaborative intervention studies used “usual care” as the comparator but did not provide standardised protocols for treatment or follow-up reviews for patients in this arm of intervention (147, 148). While some trials investigating pharmacotherapies included patients who had acute episodes of BD (144, 146), other studies included participants in a stabilised phase of BD (145, 147, 148). The number of participants included in the economic evaluations varied from 52 to 331 (144, 148). Other than one study (145), all studies provided explanations for study sample size (144, 146-148). Sample sizes were based on clinical outcomes and therefore likely to be insufficiently powered to detect any meaningful differences in economic outcomes. This may be why

only the study by Simon et al. (148) found any statistically significant differences in costs between the comparators.

Although all trial based studies reported attrition in study participants (7%–53%) for the economic analysis, none of the studies used appropriate methods as outlined in **Chapter 3** to address missingness in data. This loss of information further reduces the power to detect any significant differences in economic outcomes. The follow up period of these studies varied from 12 weeks for drugs used in acute phase to 12-36 months following randomisation for therapies used in the maintenance phase (144-148). For all trial-based studies, issues with randomisation processes such as lack of blinding or single blinding may have influenced study participation, resource use and outcomes, and therefore the external validity of their results should be treated with caution. However, at the same time, it is challenging to implement blinding for psychological therapies.

4.4.2 Data: Costs

Three studies were from the healthcare payer's perspective (145, 147, 148), one from the third party payer's perspective (i.e. institutions other than patients or healthcare providers such as insurance providers who provide re-imbursements) (144) and one study did not explicitly state the choice of perspective (146). Only Lam et al. (145), reported resource use and unit costs separately. Costs related to adverse effects were not taken into consideration in four out of five studies (144, 146-148), and as mentioned in **section 4.4.1**, none of the studies reported information on missing data. Appropriate statistical analyses including bootstrapping, regression modelling and nonparametric

tests were undertaken by all studies to address skewed cost data. Discounting of costs was undertaken in one study out of the three studies with longer than one-year time horizons (147).

4.4.3 Data: Outcomes

The majority of trial-based evaluation studies reported a range of diseases specific outcome measures such as duration of episode, number of BD free days/months, or ratings on a mania/depression scale. Three studies also presented health-related quality of life (HRQoL) outcomes obtained from the Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q) (144) and the Short Form-36 (SF-36) (146, 147). While SF-36 is a generic measures of health status as described in **Chapter 3**, Q-LES-Q is a widely used self-reported instrument in mental health research to measure degree of satisfaction and enjoyment (149). However, none of the studies derived preference-based generic measure of benefits such as QALYs or DALYs, and no reasons were provided for choosing the outcomes measures used. Discounting for outcomes was not undertaken by any of the included studies.

4.5 Analysis and interpretation of results

Neither selection nor randomisation process was clear for studies comparing pharmaceuticals, thereby making them susceptible to selection bias (144, 146). For the studies investigating collaborative care and psychological therapy, blinding was limited to care providers or interviewers, and patients were aware of their intervention status. This could have influenced the overall study participation, resource use and outcomes,

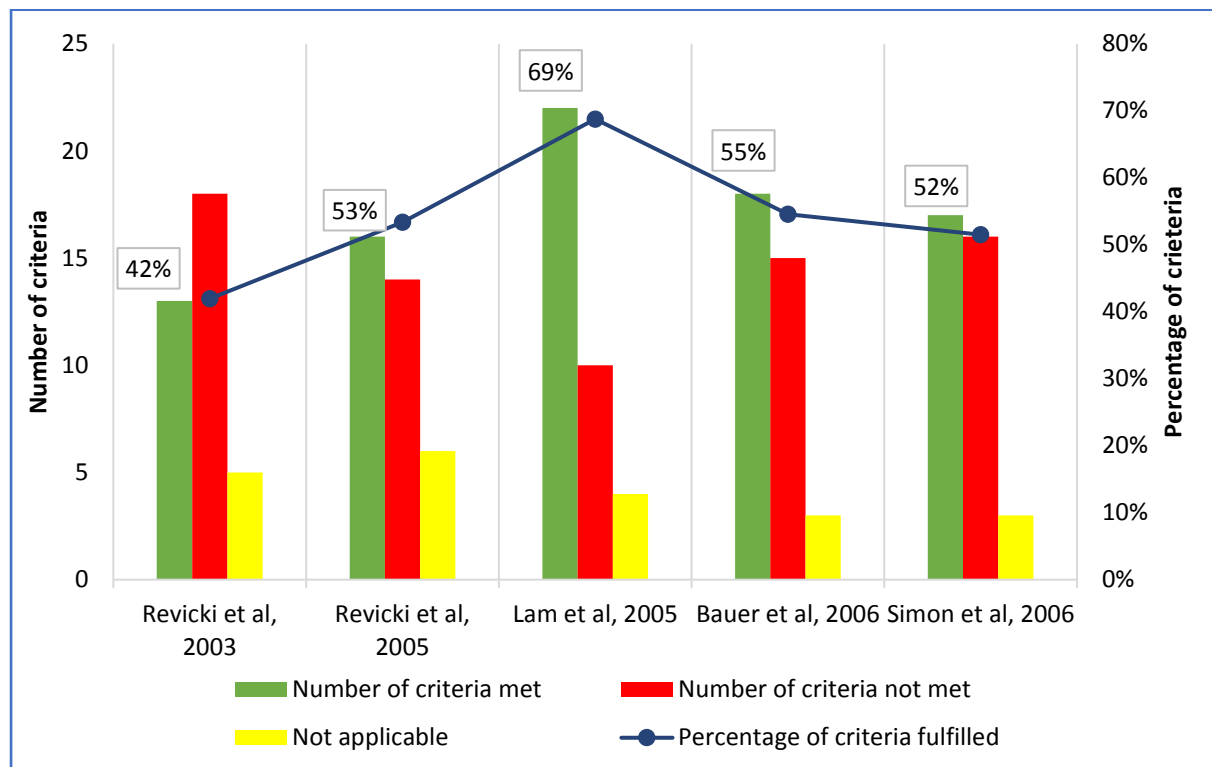
but on the other hand, it is difficult to blind participants in these types of collaborative studies.

Except the study by Simon et al. (148), none of the studies presented aggregated results (in the form of incremental cost-effectiveness ratios) with respect to the primary outcome (144-148). While difference in outcomes among alternative pharmaceutical options could not be found two studies (144, 146), studies evaluating collaborative and psychological therapies all presented favourable results in comparison to care as usual (145, 147, 148). Sensitivity analysis was undertaken in one study. Authors of this study used a net benefit approach to obtain cost effectiveness acceptability curves. This study was also the only one that discussed drawbacks with respect to transferability of results to other settings (145).

4.5.1 Overall quality scores

The overall quality score of each study was plotted along the percentage of criteria it fulfilled in **Figure 3.3**. In general, the two trial-based pharmacoeconomic studies investigating the cost-effectiveness of sodium valproate (DVP) against lithium (LI) and olanzapine (OLZ) scored poorly against the recommended checklist indicating a lack of methodological rigor(142, 144, 146). Although the studies investigating collaborative care satisfied higher percentage of quality criteria, none of the trial-based economic evaluations could be classified as having a low risk of bias (**Figure 3.3**).

Figure 3.3: Chart showing the fulfilment of quality criteria by trial based studies



The economic study by Lam et al. (145) scored highest on the quality criteria. This was a study of adjuvant CBT based on a well-conducted randomised controlled trial and showed adjuvant CBT as a cost-effective treatment option for preventing relapse in patients with BD. This result was also consistent with the recently published clinical evidence that indicated that individual psychological interventions including CBT are effective in relapse prevention (68). Nevertheless, the lack of transparency with regards to “usual care” and the use of “bipolar-free-day” as the measure of outcome may restrict its external validity and the comparability of the cost-effectiveness results across disease areas.

4.6 Model-based economic evaluations

This review identified 15 model based economic evaluations, 12 (80%) of which evaluated pharmaceutical interventions (13, 139, 150-159), two (13%) investigated adjuvant psychotherapy in community and hospital settings (160, 161), and one study explored the benefits of screening in the management of BD (162).

Pharmacological evaluations included treatments primarily for prophylaxis in euthymic BD patients (9 studies, 75%) and for managing acute episodes (three studies, 25 %). For the acute phase, three different drugs – divalproex or sodium valproate (DVP), haloperidol (HLP) and quetiapine with lithium combination (QTP+LI) - were found cost-effective in three different settings: US, UK and Netherlands -, and in three different years: 1997, 2004, and 2007 -, respectively (13, 156, 159). For prophylactic treatments, QTP or QTP combined with either with LI or DVP was determined to be cost-effective in four models(150, 152-154), while lithium, olanzapine (OLZ) or lamotrigine (LAM) were deemed cost-effective in the rest of the five prophylactic studies which used different settings and cost-effectiveness thresholds(139, 151, 155, 157, 158). Furthermore, two studies undertaken at global level identified adjuvant psychotherapy with LI to be cost-effective (160, 161).

However, these diverse results cannot be considered without exploring and understanding the credibility of the modelling methodologies used to inform cost-effectiveness of the aforementioned strategies. The modelling studies utilised different structures, data resources and methods to address uncertainty, as well as internal and external validity. Using the framework discussed in **Chapter 3** and the good practice

guidelines for reporting decision analytical models (DAMs) (132), the salient features of the included models are provided below.

4.6.1 Model structure

All modeling studies in this review had a clear definition of the decision problem, model objectives and scope. Different modeling approaches were used to address the study question including one decision tree model (162), eight cohort based Markov models (139, 151-155, 157, 158), two discrete event simulation (DES) studies (150, 156), two publications using the same mathematical model (160, 161), and two studies which failed to explicitly state their approach or structure (13, 159). All studies except Klok et al.(156), specified the primary decision maker and perspective of study. Nine studies adopted a healthcare provider perspective (13, 139, 150, 154, 155, 157, 158, 160, 161), three studies utilized a third party payer perspective (158, 159, 162) one adopted a societal perspective (151), and two reported results using both third party payer and societal perspectives (152, 153). Six of the modeling studies (40%) compared two alternative treatments; two compared four alternatives, while seven studies (47%) compared more than five alternatives. Only six (40%) studies evaluated all existing relevant treatment options (13, 150, 152, 156, 159, 162).

Decision tree models:

Menzin et al. (162) used a decision tree to determine the benefits of a screening tool (the MDQ questionnaire) to identify BD among patients suffering from a depressive episode. They used a time horizon of five years to compare outcomes and costs

associated with screening versus no screening. The structure of the tree was simple, transparent and was coherent with the disease process.

Markov models:

All the Markov models except one (151) agreed on at least three states of the disease including euthymic, manic, and depressive phases. Out of the eight Markov model studies, five studies (63%) used model structures with transition states that sufficiently reflected the disease pathway in BD including 'death' as one of the model states (139, 151-153). Two of these studies used the same model structure (152, 153), one study focused solely on side-effects of drug treatments (151), while the remaining model was developed for a health technology assessment (HTA) commissioned by the National Institute for Health and Clinical Excellence (NICE) (139). Only three out of eight Markov models incorporated the adverse events of medications in the form of either a decrement of utility (disutility) or transition to an adverse event state (139, 151, 157). The model used by Kasteng et al. (151) was unique in that it was constructed exclusively to compare pharmacotherapies with a sole focus on their side effects including metabolic syndrome, coronary heart disease and diabetes mellitus.

The time horizon adopted by the models varied between three weeks to lifetime, but only four studies (139, 151, 160, 161) used a timeframe of 10 years or longer allowing them to also capture the longer-term effects and cost associated with BD. Two of the modeling studies used three weeks to 100 days to investigate the cost-effectiveness of treatments in acute manic phase of BD (13, 156). The time horizon of five of the Markov models was based on the time duration of respective clinical trials, which ranged

between 12 to 24 months (152-155, 158). These studies did not extrapolate their findings beyond the trial period. The choice of cycle length in models varied between monthly to annual transition periods.

Only Soares-Wesier et al. (139) accounted for patients' history by splitting acute episodes into separate states and allowed switching over to medications with second-line treatments in case of non-response. All Markov models used constant transition probabilities implying that patients with BD would have a constant annual likelihood of dying irrespective of their length of stay in a given health state.

Discrete event simulation (DES) models:

The two DES models identified were of contrasting quality. Both models evaluated all the existing BD treatments and incorporated adverse events and mortality into their model, but were limited by their chosen time horizons. The DES model adopted by Klok et al. (156) simulated a large number of individuals with BD-I, where individual health state transition probabilities were determined by patient's attributes such as age, gender, number of previous episodes, existing comorbidity, etc. However, structural assumptions behind the modeling process lacked transparency and it was not clear how the authors calculated and incorporated "time-to-event" probabilities. The time horizon of this model was only 100 days which may be sufficient to evaluate the acute phase treatment of BD, but not long enough to cover the costs and effects related to adverse events, relapse, etc.

The model by Ekman et al. (150) was the only model that fulfilled all the structural quality criteria as outlined by the Philips checklist (132). The structure of the model was transparent and appropriate distributions were considered to estimate individualized time-to-event probabilities. 10,000 patients were individually simulated. Transitions between four health states (including stable state (remission), depressive episode, manic episode, and mortality) were determined by a time-to-event risk function. Patients could also discontinue their treatment due to reasons other than relapse. The study used robust survival analytical techniques for the extrapolation of short-term trial data. A five-year time horizon was considered which may be insufficient to comprehensively assess the benefits of long-term psychiatric medications and truly estimate relapse rates and associated adverse events. While the DES model built by Klok et al. (156) satisfied the lowest number of quality criteria (49%), the DES model developed by Ekman et al. (150) fulfilled the highest proportion (94%).

Mathematical models:

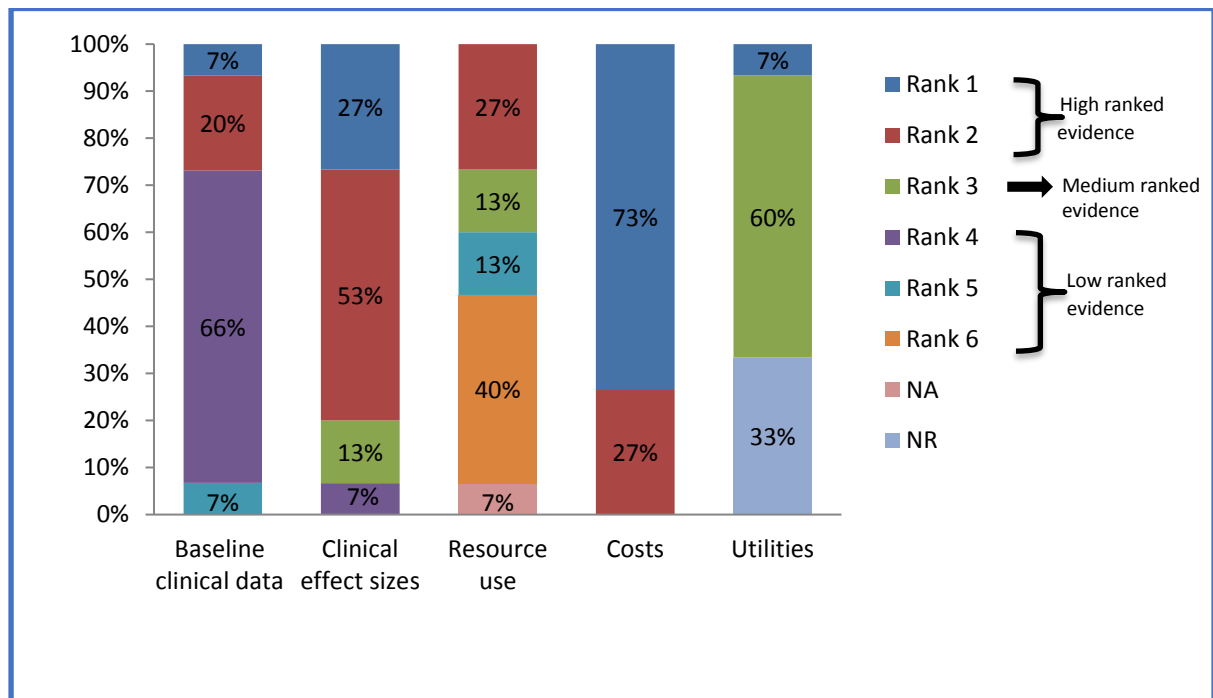
Two of the modeling studies used a recognized static mathematical model (163) to investigate the cost-effectiveness of alternative therapies across different WHO defined worldwide regions (160, 161). The modeling structure did not include all the relevant health states such as death or adverse events. The authors used a lifetime perspective, but did not report how they extrapolated the treatment effect beyond the stated 10-year period for the alternative strategies. Further, they did not consider all relevant pharmaceutical comparators in the management of BD. The mathematical model employed by both these studies investigated the same five alternative interventions, adopted similar data inputs, (160, 161) and had similar weaknesses except in terms of

presenting the incremental cost-effectiveness ratios and using a probabilistic sensitivity analysis. As a result, the quality criteria improved only marginally from 60% to 64% between 2005 and 2012. Despite their shortcomings, these studies were unique in not only eliciting cost-effectiveness of different interventions in BD management, but also comparing the results against cost-effective treatment options for competing neuropsychiatric disorders in resource limited settings.

4.6.2 Data

DAMs require a range of evidence from different sources to inform the parameterisation of models and the subsequent cost-effectiveness analysis of alternative strategies. The outcomes and reliability of a model are therefore subject to input bias dependent on the process of appropriate data identification and incorporation. It is well recognised that there is a hierarchy of effectiveness evidence; following Cochrane collaboration's recommendations to report the quality of sources of evidence used in the included models **Figure 4.4** displays the hierarchy of evidence observed in the decision analytical models (DAMs) included in this review (141, 164).

Figure 4.4 Stacked bar showing hierarchy of evidence used in DAMs as adapted from Copper et al. (164)



Most of the studies either used meta-analyses (139, 150, 155, 157) or individual RCTs (13, 151-154, 156, 158, 159) to identify relative treatment effect sizes. Although individual effect sizes from RCTs may not be generalizable to other populations or settings, given the resource restraints, they are still considered as highly ranked evidence (164).

Only four studies utilised data from “high ranked” reliable administrative datasets or prevalence surveys to incorporate baseline clinical data (156, 159-161). There was also a marked variation in how resource use parameters were identified, with only four studies incorporating information from recently published prospective data collections(152, 153, 158, 162). All included DAMs used “high ranked evidence” such as standard national sources (NHS Reference Costs, British National Formulary, Personal Social Services Research Unit) to identify unit costs.

Five (13, 155, 156, 159, 162) out of the 15 modelling studies did not use generic measures of health outcome such as QALYs or DALYs. Instead, they used natural units such as time to response, improvement of scores on mood scales, and recurrence of mood episodes. These are easier to estimate, but difficult to compare with other strategies within BD management or strategies across other disease areas. One model (158) utilised “high ranked” evidence to identify utility estimates.

Five modelling studies (13, 150, 151, 154, 158) sufficiently described and referenced their model parameters, and the process of data incorporation was transparent. Four models (154, 156, 160, 161) failed to sufficiently describe the data used for modelling, and five studies (13, 156, 160-162) did not discuss or justify the assumptions and choice of data inputs used.

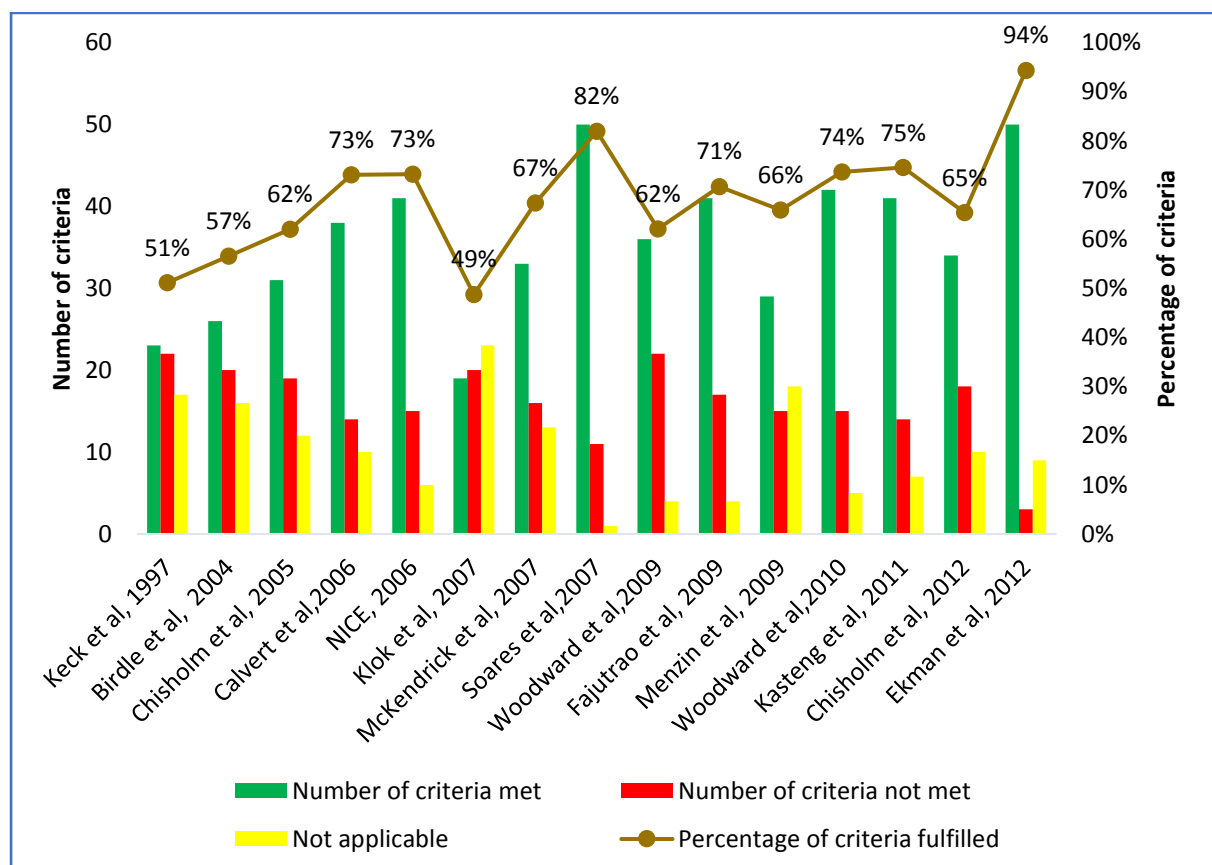
4.6.3 Model uncertainty

Six studies (139, 150, 151, 157, 161, 162) addressed methodological uncertainty by running alternative scenarios for different time horizons, discount rates or analytical perspectives. Parameter uncertainty was not addressed in four studies (13, 156, 159, 162). Seven studies undertook the current gold standard probabilistic sensitivity analysis to address parameter uncertainty (139, 150-152, 154, 160, 162). None of the studies explored structural uncertainty (e.g. by using different states) and type of model used. Heterogeneity was dealt with in seven models by running analyses for different patient sub-groups (139, 150, 151, 157, 158, 160, 161)

4.7 Overall quality of models

The overall quality of a model is described in terms of its internal and external consistency. It covers issues ranging from structural frameworks and assumptions to data inputs required to populate the model. Six studies assessed the mathematical logic of their models related to their internal consistency by running analyses using different modelling assumptions(139, 150, 156, 157, 160, 161). External consistency relates to the comparison of results with previous models, or externally validating a model by using other published epidemiological data. None of the studies attempted an external validation of their respective results. Overall, the quality of the modelling studies varied significantly, ranging from 49% to 94% (**Figure 4.5**).

Figure 4.5: Chart showing the fulfilment of quality criteria by model based studies



Only two studies scored above 80%, which included one Markov model and one DES model(139, 150). The common characteristics of both of these models were valid structural assumptions reflecting the natural disease process, robust ways to synthesise data and appropriate methods to handle uncertainty.

4.8 Discussion

The aim of this chapter was to provide an overview of existing economic evaluations, and critically appraise the evidence to identify interventions that have potential to improve health and wellbeing in individuals with BD. Despite methodological issues and associated caveats (discussed below in **Section 4.8.1 and 4.8.2**), the consistent finding in included studies indicated that Lithium (LI) or LI in combination with atypical antipsychotics like Quetiapine (QTP) are cost-effective pharmacotherapies in BD management. In addition, psychological interventions such as a collaborative care model (including a group psychoeducation) and adjuvant CBT can be cost-effective options in the long-term management of BD in the described study settings.

This systematic literature review was based on comprehensive searches in seven electronic databases and was conducted in line with the current gold standards of systematic reviews of economic evaluations and included recommended checklists for the quality appraisal of eligible studies (132, 141, 142). Use of standardised appraisal tools also allowed identification of methodological issues arising in health economic evaluations of management strategies in BD.

This review was further complemented by an assessment of the hierarchy of clinical evidence included in model-based analyses (164). These checklists, however, are inherently subjective and do not offer any cut-off scores for quality assessment. Furthermore, the Philips checklist is not ideal for evaluating DES or mathematical models, and does not explicitly consider the quality of evidence incorporated in models. Publication bias may also play an important role in the dearth of economic evidence for BD, but this could not be assessed based on published literature.

Twenty eligible studies were identified, five of which were trial-based and 15 of which were model-based economic evaluations. The studies evaluated a range of pharmacological, psychological and system-level interventions and varied largely in their methodology (different study designs, settings, costing) and outcomes. Twelve studies were sponsored by pharmaceutical companies.

In line with the aim of the thesis, **Table 4.3** brings together the overall results in terms of cost-effectiveness of all the included studies. Five studies (all model-based) reported favourable results for quetiapine (QTP) or QTP and lithium (LI) combination (150, 152-154, 156). One study favoured collaborative care model for managing BD patients, and another study in the UK supported adjuvant CBT over care as usual.

Table 4.3: Permutation matrix of the cost-effectiveness results based on Nixon et al. (164)

Health Outcomes			
Costs	Poorer	Same	Better
Higher	No studies	No studies	1) OLZ/HLP* (Bridle et al. 2004) (13) 2) Community based adjuvant LI with psychosocial care (Chisholm et al. 2005) (160) 3) LAM Vs. LI/OLZ/placebo (Calvert et al. 2006) (157) 4) OLZ Vs. various (NICE. 2006) (156) 5) Collaborative care Vs. usual care (Simon et al, 2006) (147) 6) QTP+LI (Klok et al. 2007) (155) 7) LI+IMI and LI [§] (Soares-Weiser et al, 2007) (138) 8) QTP-XR+LI or DVP (Woodward et al. 2010) (151) 9) LI in community setting (Chisholm et al. 2012) (159) 10) QTP Vs. various (Ekman et al. 2012) (149)
Same	No studies	DVP Vs. OLZ (Revicki et al. 2003) (143) DVP Vs. LI (Revicki et al.2005) (145)	11) Adjuvant CBT Vs. usual care (Lam et al.2005) (144) 12) Collaborative care Vs. usual care (Bauer et al. 2006) (146) 13) OLZ Vs. LI (McKendrick et al. 2007) (154)
Lower	No studies	No studies	14) DVP Vs. LI (Keck et al. 1996) (158) 15) QTP+LI or DVP (Fajutrao, et al. 2009) (153) 16) Screening Vs. no screening (Menzin et al. 2009) (161) 17) QTP+LI or DVP (Woodward et al. 2009) (152) 18) ARP Vs. OLZ (Kasteng et al. 2011) (150)

Studies in bold font are trial-based and studies in normal font are model based economic evaluations
 * OLZ or HLP depending on WTP threshold; § LI+IMI for depressive episodes and LI for manic episodes
 ARP= Arpiprazole, DVP=Sodium Valproate, IMI=Imipramine, HLP=haloperidol, LAM=Lamotrigine, LI=Lithium, OLZ=Olanzapine, QTP=Quetiapine, QTP-XR= Quetiapine extended release

Reject intervention	Incremental analysis required	Accept intervention

Depending on the nature of episodes and individual response, existing best available clinical evidence (see **Figure 2.3** from **Chapter 2**) recommends that along with LI, and anticonvulsants including DVP and LAM, atypical antipsychotics like QTP, OLZ and Arpiprazole (ARP) may be used as first-line treatment in the management of BD(165). There is evidence that patients who do not respond to monotherapies may benefit from trying combination regimens. Psychological interventions may be used to enhance patient's awareness, compliance with medication regimen and prevention of relapse. **Table 4.3 indicates** that the results presented in this chapter are reflected in current clinical guidelines.

However, unlike clinical measures, several caveats need to be taken into account before drawing overall conclusions from the results of these cost-effectiveness studies. This structured critical review confirms that the existing economic evaluations in BD suffer from several methodological challenges that are highlighted below; the findings from this chapter add to the recommendations that have already been outlined in recent literature to address these methodological issues (105, 133)

4.8.1 Trial-based economic evaluations

Although there have been several guidelines published in the last decade to streamline the conduct, analysis and reporting of trial-based economic evaluations, the five trial-based studies included in this review failed to display consistency (166). These studies were not without limitations and suffered from issues related to both internal and external validity. Severe mental illness like BD have multiple impacts beyond healthcare (i.e. productivity losses, care burden). Although a societal perspective is theoretically

ideal for evaluating therapeutic interventions in mental disorders, a minimum requirement is to undertake the analysis from a health and social care perspective including personal social services costs (108, 128). Further, the lack of transparency in the choice of comparators, resources use data and costing parameters would limit validity of intended economic appraisal.

The time horizon of included clinical trials was limited and driven by the occurrence of either intermediate clinical endpoints or events. This short time-frame is not adequate in patients with severe mental illness like BD and schizophrenia who not only experience long-term consequences such as drug induced diabetes, obesity, dyslipidemia but also higher mortality than the general population (167). Appropriate time-horizon should be adopted and extrapolation of findings using DAMs should be considered using relevant treatment strategies.

A common limitation observed in included trials was a failure to include generic preference-based measures like EQ-5D-3L that are recommended for evaluating outcomes (128). As **Chapter 3** outlined, use of clinical outcomes alone restricts the scope of comparisons. For example, cost-effectiveness analyses using outcomes such as 'bipolar free day' used in Lam et al. (2005) only allows comparisons with other strategies that can be expressed using the same metric and are not informative beyond BD-specific context (145).

Although use of EQ-5D-3L in severe mental illness has been challenged, until the appropriateness of EQ-5D-3L or similar measure is proven in BD, outcome

measurement should include contextually validated clinical instruments along with a suitable multi-attribute utility scale (29, 128).

Further, Lack of joint representation of costs and effects (i.e. ICERs) and failure to account for uncertainty (sensitivity analyses) in resulting estimates would limit decision maker's ability to arrive at decisions on cost-effectiveness of treatments based on a willingness-to-pay threshold.

4.8.2 Model-based economic evaluations

This review identified 15 modeling studies. Patients with BD have two-three times higher risk of death than the general population, and major goals of BD therapy are to prevent not only relapses but also premature deaths (67). Learning from the current review, any future model structure should reflect BD progression and **include all relevant comparators**. Specifically, the models should incorporate the probabilities of death and adverse events as ignoring these aspects of disease process could lead to considerable variation in modeling results. It is important to explicitly define all assumptions regarding relationships between model components, study settings, analytical perspective, primary outcomes, time horizon, and cycle length.

Due to the very nature of decision analytical models (reliance on data from various sources, different model structures, perspective, time-horizon), outcomes of DAMs were susceptible to variation and uncertainty. However, none of the models addressed all potential types of model uncertainties (as highlighted in **Chapter 3**) and most failed to justify their omission. Several assumptions were made quoting consistent lack of data.

Most DAMs also lacked rigorous validation processes of both their internal and external consistencies.

Another shortcoming in the majority of the Markov models reviewed in this study was the 'lack of memory', also known as the Markovian assumption, which meant that the model did not account for patient's history of illness or severity of episode to determine the probability of moving between health states (129). In BD, the type of recurrence is influenced by severity and duration of past episodes. Ignoring this critical information is likely to favour less effective treatments. Furthermore, individuals can only transit to different states at a fixed cycle length which represents a major assumption given that the course of BD is highly unpredictable, and no two individuals will share the same risk profile.

Recent research suggests that BD is often not episodic or two dimensional (i.e. Manic or depressive states.), instead, it represents a spectrum of mood with a prolonged sub-syndromic phase and varying peaks and troughs of low, elated and mixed mood (24, 168). In the area of BD, DES modelling framework may therefore have distinct advantages over the cohort Markov model by including customisation of outcomes with individual attributes, flexibility with unrestricted time period and the inherent ability to overcome the Markovian assumption (169).

The results of this chapter are consistent with two recently published reviews, one specifically reviewing modelling studies of pharmacotherapies in BD and another appraising both trial and model-based economic evaluations of pharmacotherapies in

the management of BD (170, 171). The critical review of model-based economic evaluations warned that the inappropriate and unjustified use of modelling techniques limits the comparability of existing BD management strategies (171). The other study by Mavranetzouli et al. (2016) reiterated similar findings of their inability to draw robust conclusions on relative cost-effectiveness between treatment strategies. Nevertheless, they also reported that pharmacotherapy is more cost-effective than no treatment. In particular, Lithium appeared to be more cost effective as a maintenance treatment, whereas Lithium in combination with atypical antipsychotics were alternative cost-effective options in the long-term management of BD (170).

4.9 Conclusions

This chapter presented a critical overview of existing economic valuations in the management of BD. The importance of economic evaluation to inform and assist the decision maker in allocating scarce health care resources was highlighted in the introduction.

The last two decades have seen a growing number of economic evaluations in BD, but there is a great deal of methodological heterogeneity in existing studies and ambiguity about the correct evaluation methods. Studies were conducted in different settings and their transferability across settings remains questionable. Given the variability in the methods and quality of the identified studies, it is difficult to provide a conclusive recommendation for the most cost-effective therapy for BD based on existing evidence. Nevertheless, with regards to pharmacotherapies, only two model based studies were

of high quality and concluded that LI or LI and imipramine (IMI) combination and QTP are cost-effective in the management of BD.

Both system-level collaborative care studies (147, 148) generated favourable outcomes, but were of insufficient time duration and scored poorly on the quality checklist. Similarly, adjuvant psychological therapies appear cost-effective in the given study settings (145, 160, 161), but additional research with longer time horizon is required to validate these findings. In addition, there is a definitive need for further empirical research in primary effectiveness studies that can feed into economic evaluations. The apparent lack of high quality studies, especially for psychological interventions and reluctance in adopting the QALY as an outcome measure implies that there is inadequate evidence on the comprehensiveness of available management strategies for individuals with BD.

This chapter provide the methodological background to this thesis, and these findings will be taken forward in the following chapters. To address one of the methodological issues highlighted by this chapter, **Chapter 6** will advance the knowledge base on appropriateness of the EQ-5D-3L in BD. **Chapter 7** will add to the limited body of evidence on psychological interventions by conducting an economic evaluation of a RCT that evaluated alternative psychoeducational management strategies for individuals with BD who use electronic monitoring of mood. For now, to highlight the economic case of BD, the attention of this thesis will turn towards **Chapter 5**, which will provide estimates of a cost-of-illness of BD derived from an observational cohort, its associated costs distribution along with exploration of factors that have the greatest impact on costs.

5.1 Introduction

Chapter 2 highlighted that BD is a severe chronic mental illness characterised by a remitting/relapsing course with episodes of mania and depression. There is no definitive cure, and all existing management strategies primarily aim to prevent relapse. The first onset of BD typically tends to appear in late adolescence to early adulthood. The public health burden of BD is widespread, relating to the significant impairment of HRQoL and well-being along with social and economic harms over a lifetime(24, 172). However, there is limited research quantifying the burden of BD, and the present estimates are subjected to substantial uncertainties with heterogeneous results reported even in the same country.

Evidence regarding the economic burden of a disease on the society can be captured from a cost-of-illness (COI) study and it could arguably offer an important decision-making aid for policy makers (173) As with any methodology, COI studies has its limitations, but as highlighted in **Chapter 3**, using a standardised approach to estimate COI across disease areas can help overcome some of the challenges.

Estimating the costs of BD can provide valuable insights into its relative burden, costs distribution along with the opportunity to explore factors that have the greatest impact on resource utilisation. This evidence may be critical in raising awareness of the

disease, thereby contributing to discussions and decisions around efficient resource allocation.

Using a prospective cohort of BD patients, the objectives of this chapter are to conduct a COI study that quantifies the societal costs associated with BD and gain a better understanding of the main factors influencing costs in patients with BD. The chapter begins by providing a summary of existing evidence in the UK.

5.2 Current evidence

A recent literature review of the cost of brain disorders in Europe estimated that mood disorders cost an overwhelming 113 billion Euros amounting to 14% of the total cost of brain disorders in 2010. The majority of the costs (63%) were indirect costs including productivity losses (174). It is also evident from studies in the United States (US) that the cost of BD, including direct medical costs along with social care costs and indirect costs (productivity losses and informal care) are imposing a major challenge to the society (175).

Although there is a growing body of evidence on the estimation of cost-of-illness for BD, most of these studies have either been conducted in the US or have explored only the direct healthcare costs of hospitalised patients (19, 175, 176). According to recent systematic reviews of cost-of-illness of BD, few studies have comprehensively assessed the societal costs of BD in Europe, and only three studies (see **Table 5.1**) have attempted to estimate the burden of BD in the UK (18, 23)

Table 5.1: Existing studies describing the annual Costs of BD, UK

Study	Year of original study and study perspective	Type and perspective of study	Estimated prevalence, n (%)	Data source	Average per person healthcare costs	Average per person societal costs	Total healthcare costs (NHS perspective)	Cost-of-illness	Cost of hospitalizations proportional to healthcare costs	Proportion of COI attributed to indirect costs
Gupta and Guest, 2002	1997-1998 societal	Top-down ; prevalence-based	297,000 (0.5%)	DIN-link database, HES	£670	£6919	£199 million	£2 Billion	50 %	86%
McCrone et al, 2008*	1998-1999 societal	Mixture of top-down & bottom-up; prevalence-based	1.14 million (2.2 %)	Report based on published literature	£1404	£4561	£1.6 billion	£5.2 billion	15 %	69%
Young et al, 2011	2007-2008 healthcare	Top-down; prevalence-based	87,338 (0.14%)	IMS Disease analyser, HES, Mental Health Minimum Dataset statistics	£3916	NA	£342 million	NA	60 %	NA

* Costs calculated for England and estimates obtained from Lam et al. (2005); HES: Hospital Episode Statistics

The original data used in the two previous societal assessments for BD in the UK are more than 15 years old (8, 177). Gupta and Guest (2002) used a definition of BD that included schizoaffective disorder and recurrent unipolar depression and did not include informal care costs (a critical parameter for this illness) (177). McCrone et al. (2008) used prevalence estimate from the US, derived their COI estimates from a previously published trial that only included BDI patients and assumed estimates from the PRiSM Psychosis Study to measure informal care costs (8). A more recent study by Young et al. for year 2007/08 only considered direct healthcare related costs, which the previous studies showed to represent only a minor fraction of the total societal costs (178). All three UK-based studies gave very different healthcare cost estimates most likely due to the differing prevalence estimates for BD, inconsistent case definition of BD and variance in the applied study methodologies (**Table 5.1**).

As described in **Chapter 2** and **4**, Over the past decade, there has also been an emergence of new therapies (e.g. second generation antipsychotics), modalities of care (e.g. adjunct psychological therapies) and service delivery options (e.g. collaborative care, reduced admission beds) (6). Given the methodological heterogeneity and paucity of UK based COI studies, it is likely that all three existing UK-based COI studies may be misrepresenting the current economic burden of BD in the UK; therefore an updated COI study that follows good practise recommendations is now required (18).

This chapter seeks to comprehensively estimate the societal costs of BD in the UK for the year 2010/11 through prospectively following-up a representative cohort of BD

patients over a 12-month period. The main clinical and socio-demographic factors driving the variations in the individual annual cost of BD are explored using regression techniques. Average costs are projected to the UK population using relevant prevalence estimates.

5.3 Methodology

This research was a prevalence-based COI study with an estimation of economic burden over a defined period. A societal perspective was adopted to prospectively collected patient-level data (bottom-up approach) for the estimation of direct and indirect costs (**see section 3.3** for more on COI study designs/methods).

5.3.1 Patient Cohort

This chapter uses data obtained from the OXTEXT-2 sub-study which is a part of the OXTEXT research program that follows a cohort of patients. Patients were invited to take part between May 2010 and June 2011 from the larger OXTEXT-1 cohort study; a sample of ninety-one patients participated. As described in **Chapter 1**, the OXTEXT-1 study is a pragmatic longitudinal cohort of patients with a confirmed diagnosis of BD who attend psychiatric clinics in Oxford Health NHS Foundation Trust (OHFT) and are monitored with a user friendly SMS or web-based system called 'True Colours' (TC) (35). Implementation of TC has been supported by OHFT since 2008 and is now offered by the Trust across all its community mental health teams (CMHTs) to patients as a part of routine clinical care (37).

Eligible criteria were: males and females aged over 16 years old, with a primary diagnosis of BD according to the DSM-IV-R criteria, who were willing and able to provide informed consent to participation in the study. The study received ethical approval by the local Research Ethics Committee (Oxfordshire REC A, Reference: 10/H0604/13).

Clinical status of patients was measured using the five-item Altman Self-Rating Mania (ASRM) scale and 16-item Quick Inventory of Depressive Symptoms–Self Report (QIDS-SR) scale. The ASRM scale (scores range from 0-20) consists of five questions marked between 0-4 with increasing severity. In the literature, cut-off point of less than or equal to 5 is indicative of ‘being not in a manic state’, whereas a score of greater than 5 indicates a ‘state of mania’ (179). The QIDS-SR consists of 16 items each scored between 0-3. It has nine criterion symptom domains and is scored by summing the highest response in each of a set of questions related to sleep, weight and psychomotor symptoms and then adding the remaining items. Scores range from 0-27 and can be categorised as a) No depression: 0-5, b) Mild depression: 6-10, c) Moderate depression: 11-15, d) Severe depression: 16-20, e) Very severe depression: 21-27 (180).

5.3.2 Data collection

The data on the use of health and non-health care resources were collected using self-reported patient questionnaires at five different time points: baseline assessment then at 3, 6, 9 and 12 months (sample questionnaire is provided in **Appendix 11**). In addition to information on socio-demographics, disease status and duration, data were

collected on contacts made with primary, secondary, tertiary and social care; psychiatric medication; patients' out-of-pocket (e.g. private dietician, acupuncture) and travel costs; and productivity losses such as sick leave and informal care. Premature mortality related productivity losses could not be included due to the size and timeframe of the study; no episodes of self-harm or completed suicides were recorded during the period of study. Top-down mortality rate estimates were not included as robust estimates of relevant mortality rates are not available from the literature. Further, inclusion of mortality based production losses incurred in future is controversial for prevalence based COI studies as it does not relate to other time-defined costs that are usually measured over a year (23).

5.3.3 Cost estimation

Both the National Health Service (NHS) and a societal perspective was adopted in this analysis. Average annual per patient costs and societal costs were estimated using British pounds sterling (£) for the year 2010-2011. For this, resource use categories were multiplied by their relevant national-level unit costs. Unit costs (see **Table 5.2**) were obtained from various sources including the National Schedule of Reference Costs (2010–2011), Unit Costs of Health and Social Care (2011), and the British National Formulary 65 (181-183). Data on distance travelled to psychiatric health care providers and the time informal carers (e.g. family members and friends) spent providing unpaid care and support was also collected for each study participant.

Table 5.2: Unit costs of resource use, 2010-11

Resource use		Unit costs		Source of estimate
Non mental health in-patient services				
Non-mental health inpatient admissions		£58-£1610	per day	Curtis & Netten (2010) , NHS reference costs (2010-2011)
Non-mental health out-patient & emergency services				
Outpatient		£54-£171	per visit	NHS reference costs (2010-2011)
Laboratory investigations		£3-£177	per test	NHS reference costs (2010-2011)
Diagnostic imaging		£29-£173	per test	NHS reference costs (2010-2011)
A & E		£107	per visit	NHS reference costs (2010-2011)
Ambulance		£220	per call-out	NHS reference costs (2010-2011)
Mental health out-patient				
Psychiatrist		£3.00	per minute	NHS reference costs (2010-2011)
Clinical psychologist		£0.60	per minute	Curtis & Netten (2010)
Mental health in-patient				
Psychiatry		£327	per night	NHS reference costs (2010-2011)
Psychiatric intensive care unit (PICU)		£655	per night	NHS reference costs (2010-2011)
Mental health day-care		£21-127	per session/ per day	Curtis & Netten (2010) & NHS reference costs (2010-2011)
Community-based services				
Community psychiatrist		£1.85	per minute	NHS reference costs (2010-2011)
Community psychologist		£0.60	per minute	Curtis & Netten (2010)

Community psychiatry nurse	£0.50	per minute	Curtis & Netten (2010)
Community support worker	£0.25	per minute	Curtis & Netten (2010)
Drug and alcohol services	£50.00	per contact	NHS reference costs (2010-2011)
GP	£32.00	Per surgery consultation lasting 17.2 minutes;	Curtis & Netten (2010)
GP practise nurse	£12.00	Per clinic consultation lasting 15.5 minutes;	Curtis & Netten (2010)
Home care worker	£0.40	per minute	Curtis & Netten (2010)
Housing worker	£0.26	per minute	assuming national average salary, ONS (2010-2011)
Social worker	£0.88	per minute	Curtis & Netten (2010)
Voluntary worker	£0.26	per minute	assuming national average salary, ONS (2010-2011)
Other services	£8.8-£192		Curtis & Netten (2010) & NHS reference costs (2010-2011)
Other			
Sick leave	£83.00	per day	NICE & Sainsbury MH centre 2010-2011
Informal care	£15.6	per hour	assuming national average salary, ONS (2010-2011)

In this study, lost productivity costs were estimated using the human capital approach. All participants were asked the question: “During the specified period, how many days have you been unable to work because you have been unwell?” Sick days or days off work were multiplied by the average daily cost of sickness absence in the UK (99, 184, 185). These costs were estimated only for those who were employed or self-employed (n=53, 58.2%). Informal care costs were estimated by multiplying the average hourly salary derived from the average annual gross UK wage rate for 2010/11 with the number of hours of unpaid work; use of this opportunity cost approach to estimate

informal care costs was previously reported in the Personal Social Services Research Unit's report "Unit Costs Health Social Care 2012" (186, 187). Unit costs were attached to patient reported resource use to generate total annual average costs incurred by BD patients.

5.3.4 Statistical analysis

A p value of ≤ 0.05 was used in all hypothesis tests of statistical significance. All statistical analyses were undertaken using STATA 12.0 (StataCorp, College Station, Tex., USA).

Addressing missingness:

The data utilised in this chapter is from a longitudinal cohort, where same individuals provided repeated measurements over time. Unfortunately, as commonly observed with longitudinal designs in mental health research (188), these measurements were not available for all participants at all time-points. **Chapter 3** provided a brief background into different types of missingness and ways to address it. In this chapter, missingness in observed data was assumed to follow a missingness at random (MAR) pattern, and multiple imputation (MI) was undertaken in Stata 12 using the `mi` command. After declaring the dataset to be ready for MI, those variables that were missing at any time-point and subjected to imputations were registered as "mi impute", and other variables which were not imputed recorded as "mi regular".

Variables used for imputation included the clinical measures (mania and depression related scores) and the various cost categories (i.e. GP, AE, outpatient, inpatient costs etc.). The regression covariates used were age at baseline, type of BD, gender, duration of illness at baseline, baseline and concurrent clinical measures, and the value of various cost categories at baseline. The MI by chained equations procedure (“MI impute”) with predictive mean matching to account for skewed distribution in cost categories was applied to perform 100 imputations (189). The “MI estimate” command then consolidated estimates from each imputed dataset producing an overall estimate, variance and confidence intervals using Rubin’s combination rules (190).

Specifying a model to identify determinants of overall cost:

To determine the influence of major socio-demographic and clinical characteristics such as age, gender, type of BD diagnosis, duration of illness, and clinical status on overall costs and the other main cost categories, Generalised Estimating Equations (GEE) were used on the fully imputed data. GEE is a recommended method to study the evolution of primary outcomes (i.e. total costs, healthcare costs and indirect costs) over time, while adjusting for potential predictors, to take into account the correlation between repeated measurements on the same individual. In particular, as cost outcomes are skewed, log transformation or other linear regression models may be inappropriate. GEE with a gamma distribution and log link can provide a suitable modelling method for skewed costs, particularly when there is a dependency between observations for the same subject (191, 192).

The GEE approach provides population level parameter estimates in which the correlation between observations within an individual is treated as a nuisance that needs adjustment. GEE is conducted in two stages: the first stage involves estimating the appropriate correlation matrix; the second stage uses the specified working correlation structure to derive the parameter estimates and standard errors of the regression model while adjusting for autocorrelation (193).

A multivariate regression with gamma distribution and log-link function was used to address the skewness of the cost data and an exchangeable covariance matrix was found appropriate to account for the correlation between repeated covariates and their influence on the primary outcomes which were measured on multiple occasions. The final model was based on selecting a model with the lowest “Quasi-likelihood under the Independence model Criterion” (QIC score). In contrast to the Generalised linear models (based on maximum likelihood theory), the GEE method adopted in this study is premised on the quasi-likelihood theory and therefore the likelihood-ratio test is usually inappropriate (194). The QIC score is analogous to “Akaike information criteria” that is often used for choosing the best set of variables and assessing goodness-of-fit in the models (195). Following GEE multivariable regression analyses, post-estimation adjusted (adjusted for explanatory variables) costs were estimated using the 'margins' command (196).

5.3.5 Prevalence estimates

Globally, the prevalence estimates of BD vary widely between studies, with lack of any population-based prevalence studies in the UK (197, 198). The world mental health

survey undertaken in 11 countries provided an annual prevalence rate of 0.4% for BPI, 0.3% for BPPI, 0.8% for sub-threshold BD, and an overall annual prevalence rate of 1.5% for BSD, with minimal evidence of a gender difference (199). Recent reviews have suggested estimates ranging from 0.5% to 1.1%, but few other epidemiological studies have claimed contrasting BD prevalence ranging as high as 5.1% to 6.0% (19, 200-203).

A rapid literature review was undertaken to identify any 12-month prevalence estimates of BD relevant for the UK. The electronic databases Embase, and Medline were searched for studies published between 1990 to April 2014. An established filter for BD used in the Bipolar disorder guideline developed by the National Institute for Health and Care Excellence along with key indexing terms for prevalence, and epidemiology were used (157).

The search was limited to the Northern , Western and Southern European countries as defined by the United Nations Development Statistical Division (204). Other inclusion criteria were: a) population or community-based studies using random sampling methods, b) cohort of people aged ≥ 16 years and c) studies in any of the languages spoken in the defined geographical area were included. Exclusion criteria were: a) studies using clinical samples, b) studies using a narrow age band (< 10 years). A pooled prevalence estimate was derived using weighted inverse variance methods: the DerSimonian-Laird method (205). The I^2 statistic was used to assess the approximate proportion of total variability in observed prevalence estimates that could be attributed to heterogeneity other than that due to chance; random effect models

was used if I^2 statistic >75%, indicating high level of heterogeneity (206). The prevalence estimates were subjected to both univariate and multivariate probability sensitivity analyses (see section on sensitivity analyses).

5.3.6 Cost of illness estimates:

The total average annual cost per patient was multiplied with the derived prevalence measure to estimate the societal cost of BD in the UK.

5.3.7 Sensitivity analyses

One-way sensitivity analyses were performed to explore how uncertainties in key input parameters individually impact the base case cost and COI results. Prevalence was varied using the lower and upper bound of the pooled meta-analytic prevalence estimate. Cost estimates that attributed considerably to the overall cost burden (e.g. indirect costs, hospitalisation costs, drug costs, etc.) were also varied according to their lower, and upper 95 % confidence interval limits. Secondly, a complete case analysis was also carried out to compare the results of MI with those of the complete case scenario. To check the impact of joint uncertainty in the input data and assumptions, a multivariate probabilistic sensitivity analysis (PSA) using Monte Carlo simulation was also carried out using Microsoft Excel 2010 (124).

Gamma distribution for costs and triangular distribution for prevalence estimates were utilised. Shape and scale parameters for gamma distributions were based on the mean, and SE derived from the base case analysis following imputation, while maximum, minimum and most likely values for triangular distribution were derived from

the mean and confidence interval of the pooled meta-analytic prevalence estimate. All parameters were simultaneously varied, and 5000 iterations were repeated to generate 95% CI for the total cost, and COI estimates.

5.4 Results

5.4.1 Study population

A total of 91 patients completed the baseline questionnaire. Resource use data was available for 85% (77/91), 78% (71/91), 76% (69/91) and 75% (68/91) of the sample at 3, 6, 9 and 12-months follow-up, respectively. 53% (48/91) of participants had complete data on all 4-prospective time-points, whereas 75% (68/91) of patients had cost data on at least three data-points. Overall, the longitudinal design provided 364 complete measures of response variables and covariates collected over the four time-points.

Baseline sample characteristics of the study sample (n=91) alongside the characteristics of the total OXTEXT1 cohort (n=271) are provided in **Table 5.3** The baseline characteristics of the included participants were similar to the original cohort. Most of the participants were females, 58 (63.7%); the average age at baseline was 45.5 years (SE: 1.46, 95% CI: 42.5-48.3), and 41.8% were either unemployed or retired.

Table 5.3: Demographic and clinical characteristics of participants

Characteristics of participants	OXTEXT-2 (n = 91)	*Complete cases (n=48)	OXTEXT1-cohort (n=271)
Mean age (SEM, range), years	45.5 (1.5, 20-75)	51 (1.7, 20-74)	40.4 (0.8, 17-76)
Female gender	58 (63.7)	33 (68.8)	184 (68%)
Ethnicity, n (%)			
Caucasian	88 (97%)	46 (96%)	256 (94%)
Non-Caucasian	3 (3%)	2 (4%)	15 (6%)
Employment status, n (%)			
Employed/self-employed	45 (49.5)	25 (52.1)	136 (50.2)
In education	8.0 (8.8)	3 (6.3)	33 (12.3)
Unemployed/retired	38 (41.8)	20 (41.7)	102 (37.5)
Mean years of education (SEM, range)	15.8 (0.31, 11-19)	15.9 (0.47, 11-19)	15.9 (0.16, 7-19)
Diagnosis, n (%)			
BDI	55 (60.4)	31 (66)	171 (63.1)
BDII	36 (39.6)	17 (36)	100 (36.9)
Mean QIDS score, W0 (SEM, range)	8.2 (0.51, 0.0-22)	6.9 (0.79, 2.0-22)	9.9 (0.40, 0.0-24)
Mean ASRM score, W0 (SEM, range)	3.0 (0.34, 0.0-16)	3.3 (0.59, 0.0-16)	3.5 (0.31, 0.0-20)

*patients had complete data at all four time-points

The average duration of illness was 14 years (SE= 1.13). Out of the 91 patients, 55 patients (60.4%) met BDI criteria, 36 individuals (39.6%) were classified as BDII patients. The average baseline QIDS score for depression was 8.2 (SE=0.51). The average ARSM score was 3 (SE=0.34). Higher scores based on these two scales indicated greater severity of depression and manic symptoms respectively. There were no reports of deaths during the study period.

5.4.2 Costs of BD

The average annual costs of BD per patient are presented in **Table 5.4** by resource use category.

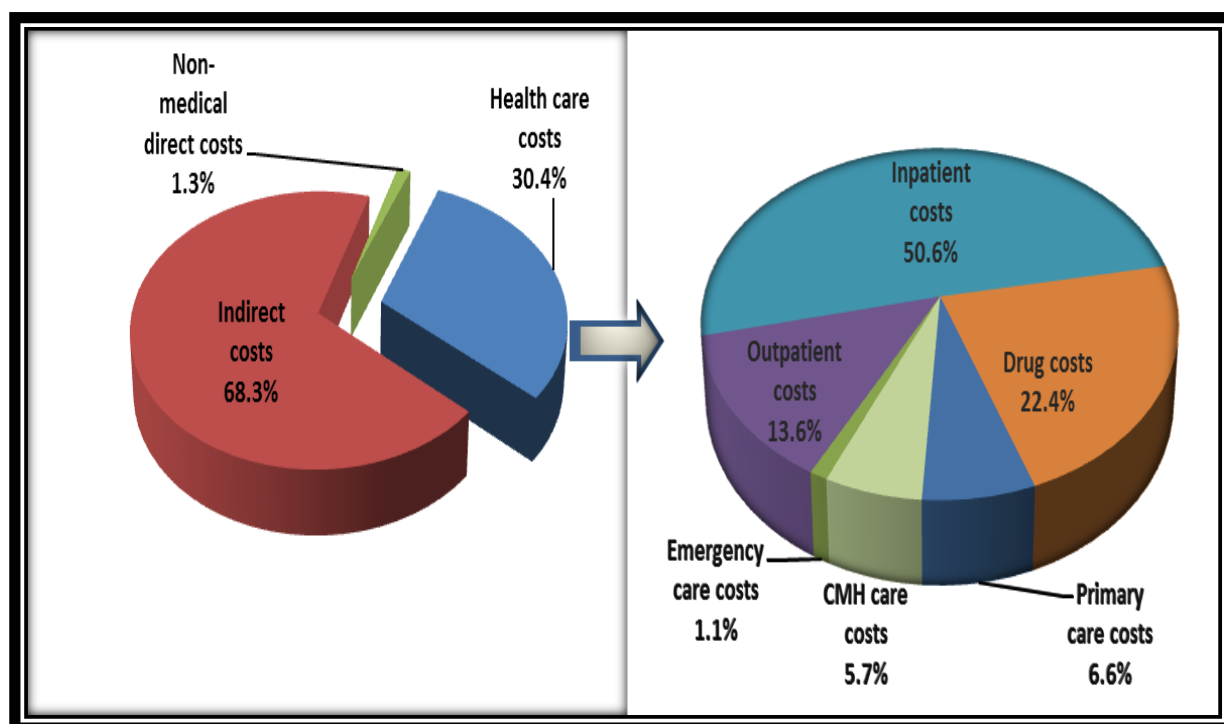
Table 5.4: Average annual costs of BD per patient based on imputed data

Resource use categories		Imputed data (n=91)			Complete case analysis (n=48)		
Primary care costs		Mean	95% CI		Mean	95% CI	
General practitioner		223.1	198.1	248.2	217.7	152.6	282.7
General practise nurse		30.2	25.3	35.0	34.0	21.1	47.0
Community mental healthcare costs							
Community psychiatric nurse		63.7	48.2	79.1	59.9	24.4	95.5
Community psychologist		37.6	21.9	53.2	21.8	-3.3	46.8
Community psychiatrist		54.2	36.2	72.2	39.8	14.3	65.3
Community drop-in centre		10.6	6.0	15.1	9.2	0.0	18.3
Support group		54.8	31.3	78.2	21.3	2.2	40.3
Emergency care costs							
Ambulance costs		23.7	12.5	35.0	27.5	-0.8	55.8
Accident & emergency costs		16.7	9.3	24.0	22.3	1.1	43.5
Outpatient care costs							
Non-MH OP		122.1	89.6	154.7	142.9	41.8	244.0
Psychiatrist OP		366.9	327.3	406.5	23.1	6.6	39.6
Psychologist OP		34.4	28.1	40.8	313.2	235.5	390.8
Hospitalisation costs							
Non-MH Hospitalization		513.6	254.3	772.9	739.5	27.0	1452.0
MH day care		222.8	127.5	318.1	232.8	-70.4	536.1
MH Hospitalization		1225.0	419.7	2030.3	415.6	-420.4	1251.6

Pharmaceutical costs	859.3	731.8	968.8	909.1	559.8	1193.9
Other direct costs						
Social care	59.5	40.3	78.7	70.8	18.2	123.4
Other private costs	111.2	72.3	150.1	96.8	49.1	144.6
Indirect costs						
Sick leave costs	1554.7	1164.7	1944.6	1117.0	461.5	1772.6
Informal care	7032.9	5318.7	8747.0	7159.9	2104.8	12215.0
Total costs	12616.8	10551.9	14713.4	11674.2	5583.3	17765.2

The contribution of each category of costs to the overall societal costs of BD is presented in **Figure 5.1**. A descriptive summary of the individual cost categories is provided with **Figure 5.2** displaying the average costs per person for individuals cost categories.

Figure 5.1: Distribution of the total costs and health care costs of BD by major sub-categories



Health care costs

Primary care costs:

Primary care costs included visits to/by general practitioners (GP) and general practice nurses. GP costs accounted for 88% of the total primary care costs. The average primary care cost was £253 (SE: £13) per patient, representing 2% of the estimated total average annual societal BD cost per patient.

Community mental healthcare (CMH) costs:

The mean annual CMH cost per patient was £221 (SE: £24). These costs were divided amongst community psychiatric nurse (29%), community psychologists (17%), community psychiatrists (25%), community drop-in centres (5%) and support group visits (25%).

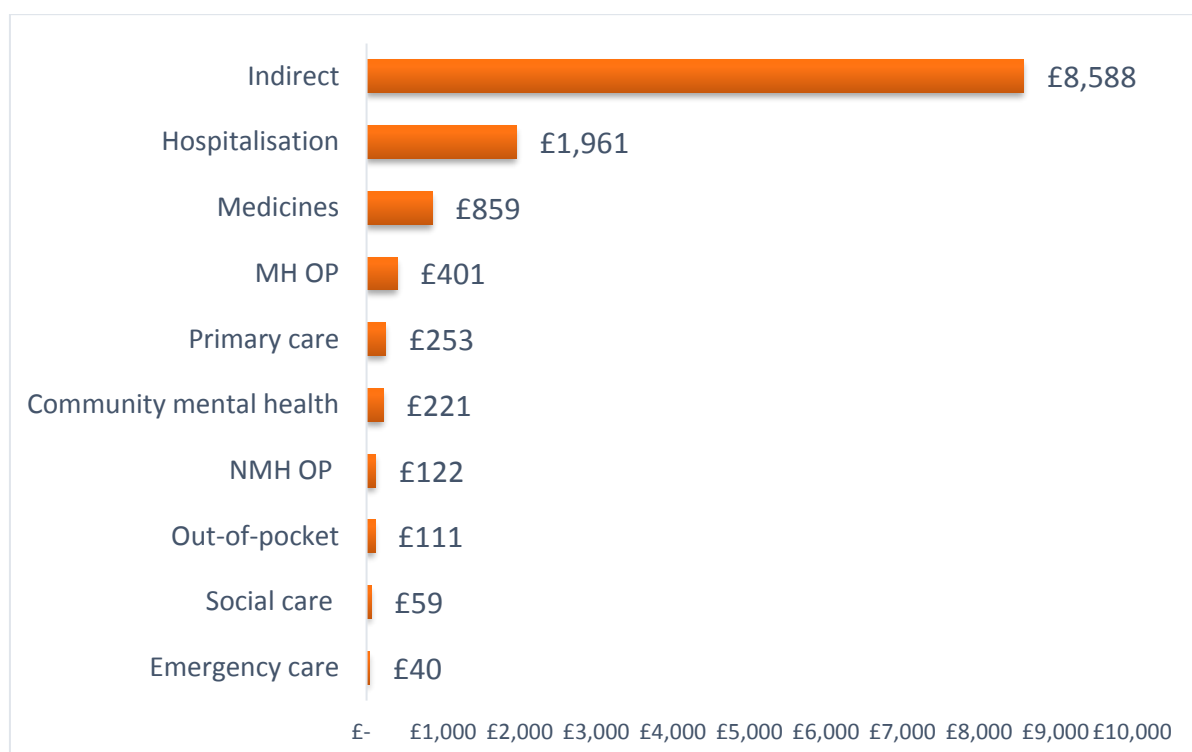
Emergency care:

Visits to A&E and use of ambulance services were categorised under emergency care. The average cost of emergency care per patient was £40 (SE: £7), with 59% (£24) of the costs being associated with the use of ambulance services.

Mental health outpatient care costs:

Outpatient care comprised hospital-based outpatient visits to psychiatrists, psychologists. The average cost for per BD patient using outpatient services was approximately £401 (SE: £23). Psychiatrist visits accounted for 91% of total outpatient costs, followed by psychologist visits (9%).

Figure 5.2: Annual costs of BD to society and the public sector (£, 2010/11 prices, per person with BD)



*Indirect costs include productivity loss and informal care; MH: mental health; NMH: non- mental health

Non-mental health outpatient care costs:

Outpatient visits as a result of non-mental health illnesses cost £122 (SE: £17) per BD patient.

Hospitalisation costs:

Of the total average annual societal costs, hospitalisation cost was the second leading sub-category and accounted for 15.5% of the total annual costs. These costs were analysed separately for mental health (MH) inpatient costs (63%), mental health day-

patient costs (11%), and other, non-mental health inpatient and day-patient costs (27%). Mean mental health-related inpatient hospitalisation costs were estimated at £1225 (SE: 408), mental-health day-patient costs at £223 (SE: £48), and non-mental health hospitalisation costs at £514 (SE: £131) per BD patient.

Psychiatric medication costs:

The estimated annual mean psychiatric medication costs were £859 (SE: £65) per patient. This was the second largest direct healthcare cost sub-category contributing to 6.8% of the overall societal cost of BD.

Other costs

Social care costs:

These costs included services offered by social workers, community support workers, home care workers, housing workers and voluntary workers. In total, mean social care costs per BD patient were estimated to be £60 (SE: £10).

Out-of-pocket patient costs:

Patient costs were out-of-pocket expenditures spent on services offered by private health service providers (e.g. dieticians, chiropodists, private physiotherapists etc.)

and travel costs to the outpatient mental health facilities. The average annual out-of-pocket costs per BD patient were £111 (SE: £20).

Indirect costs:

Indirect costs included costs due to lost productivity as a result of sickness absence, and informal care provided by family members and friends. Informal care costs amounted to £7033 (82%) and sickness absence to £1555 (18%) of the total average indirect costs per patient (£8587, SE: £891). Indirect costs were the largest single cost factor accounting for 68.1% of the total cost of BD in the UK.

Total costs

Overall, the average total costs incurred by a BD patient were estimated at £12,617 (SE: £1085).

5.4.3 Prevalence estimates

A total of 913 citations were retrieved following the initial search, out of which, 26 studies were included for abstract review. Eventually, seven studies were deemed as eligible and were included in a random-effects meta-analysis to derive a pooled estimated for the 12-month period prevalence of BD relevant to the UK (see **Figure 5.3**).

Figure 5.3: Study flow diagram

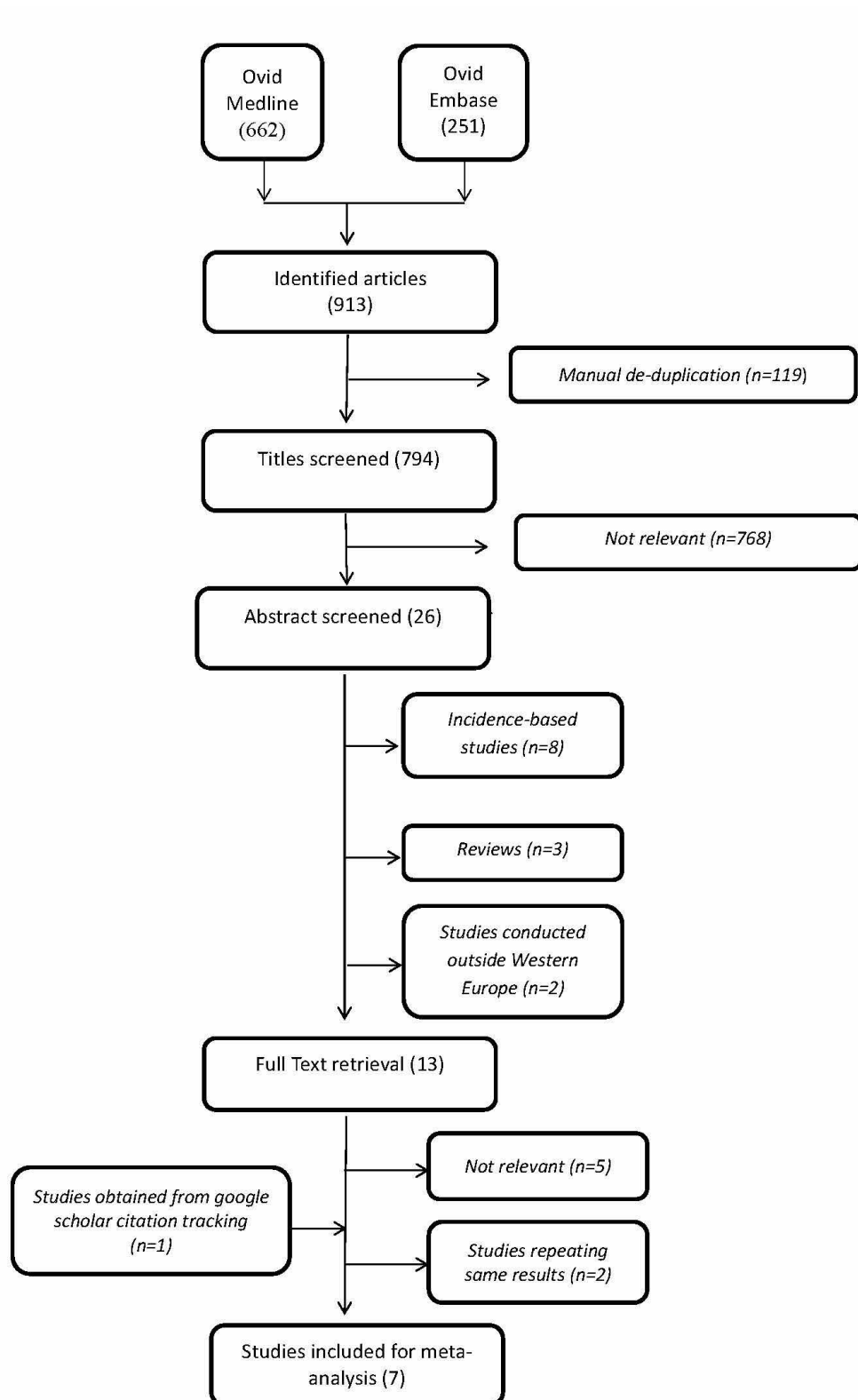
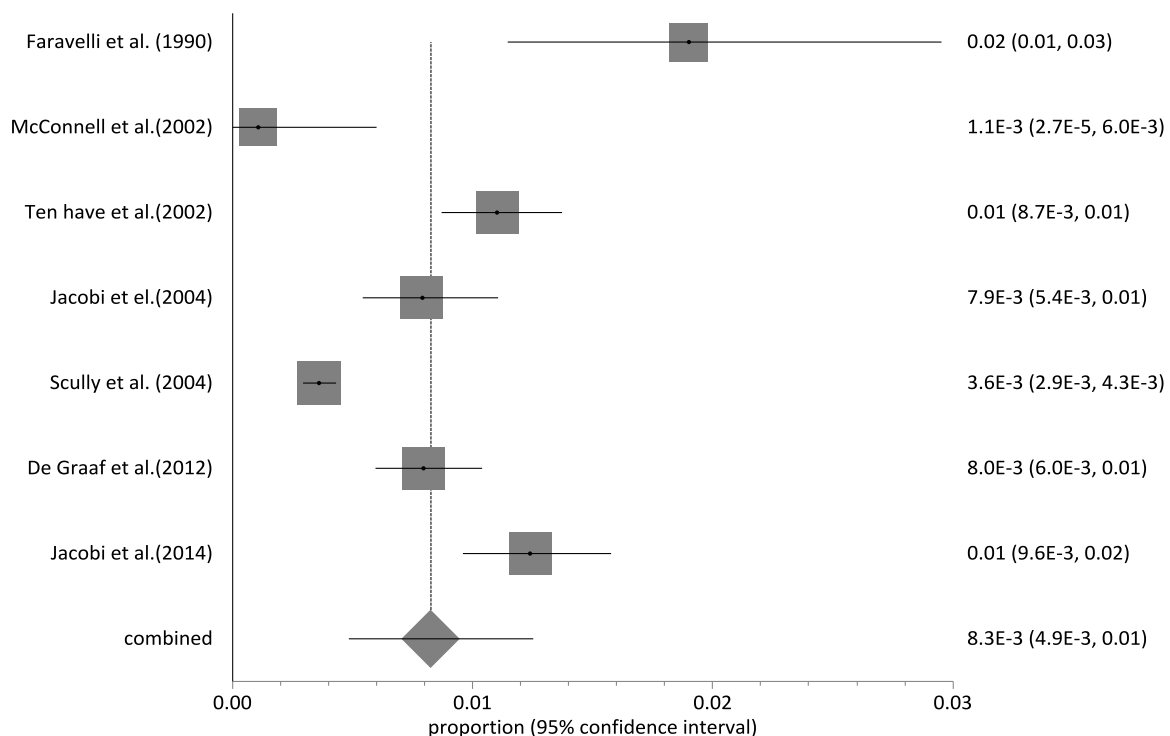


Table 5.5: Population-based prevalence studies

Study	Country, year of study	Type of sample, age group	Study population	Diagnostic criteria	Instrument	Diagnosis	Prevalence % (SE)	Type of prevalence
Faravelli et al. (1990)	Italy, Florence, NR	Community, >15 years	Total=1000 F=537 M=463	DSM-III, DSM-III-R	Structured interview	BDI, BDII, Cyclothymia	Total= 1.9 (0.3) F= 2.8 (0.7) M= 0.9 (0.3)	12 months prevalence
McConnell et al.(2002)	UK, Derry (1993-94)	Community, 18-64	Total=923 F=496 M=421	DSM-III-R, ICD-10	SCAN	BD	Total= 0.2 (0.1)	12 months prevalence
Ten have et al.(2002)	Netherlands (1996)	National, general population, 18-64	Total=7076 F=3488 M=3588	DSM-III-R or ICD-10	CIDI	BD	Total= 1.1 (0.1) F= 1.1 (0.2) M= 1.1 (0.2)	12 months prevalence
Scully et al.(2004)	Ireland, Monaghan (1996)	Community, >18 years	Total=29542 F=14563 M=14979	DSM-III-R	SCID	BD	Total= 0.4 (0.3) F= 0.4 (0.05) M= 0.3 (0.04)	12 months prevalence
Jacobi et el.(2004)	Germany (1998-99)	General population, 18-65	Total=4181 F=2103 M=2078	DSM-IV	M-CIDI	BD	Total= 0.8 (0.1) F= 1.1 (0.2) M= 0.6 (0.1)	12 months prevalence
De Graaf et al.(2012)	Netherlands (2007-2009)	National, general population, 18-64	Total=6646 F=3303 M=3343	DSM-IV	CIDI-3	BD	Total= 0.8 (0.2) F= 1.0 (0.3) M= 0.7 (0.2)	12 months prevalence
Jacobi et al.(2014)	Germany (2008-2011)	National, general population, 18-79	Total=5318 F=2778 M=2540	DSM-IV	M-CIDI	BD	Total= 1.5 (0.2) F= 1.7 (0.3) M= 1.3 (0.3)	12 months prevalence

The characteristics of included studies are provided in **Table 5.5**. Prevalence of BD ranged between 0.15% and 3% in individual studies. The meta-analytical pooled annual prevalence of BD (**Figure 5.4**) according to the DSM-III, DSM-III-R and DSM-IV diagnostic criteria was 0.8% (95% CI: 0.43% to 1.3%), with high heterogeneity ($I^2 = 94.6\%$). This is a conservative prevalence estimate as DSM-III and DSM-III-R criteria requires a minimum of one week of symptoms related to manic, hypomanic and subsyndromal mania/hypomania in contrast to the four days cut-off employed by DSM-IV classification (202). When applying the prevalence rate of 0.8% to the 2010 adult UK population of 50.9 million (≥ 16 years of age), the total number of adults living with BD in the UK in 2010 was estimated at 407,567 (207).

Figure 5.4: Pooled analysis of BD prevalence in population-based prevalence studies



5.5 Economic burden of BD in the UK

Using the above prevalence and per patient cost estimates, the societal cost of BD was estimated at £5.14 billion for the year 2010/11. Of these, £3.5 billion (68.1%) were due to indirect costs including productivity losses and informal care, £1.57 billion (30.6%) were due to healthcare related costs, £24.2 million (0.4%) were due to social care and £45.3 million (0.9%) incurred as patients' out-of-pocket expenses (see Table 5.6).

Table 5.6: Cost of illness of BD in the UK (2010-11, £)

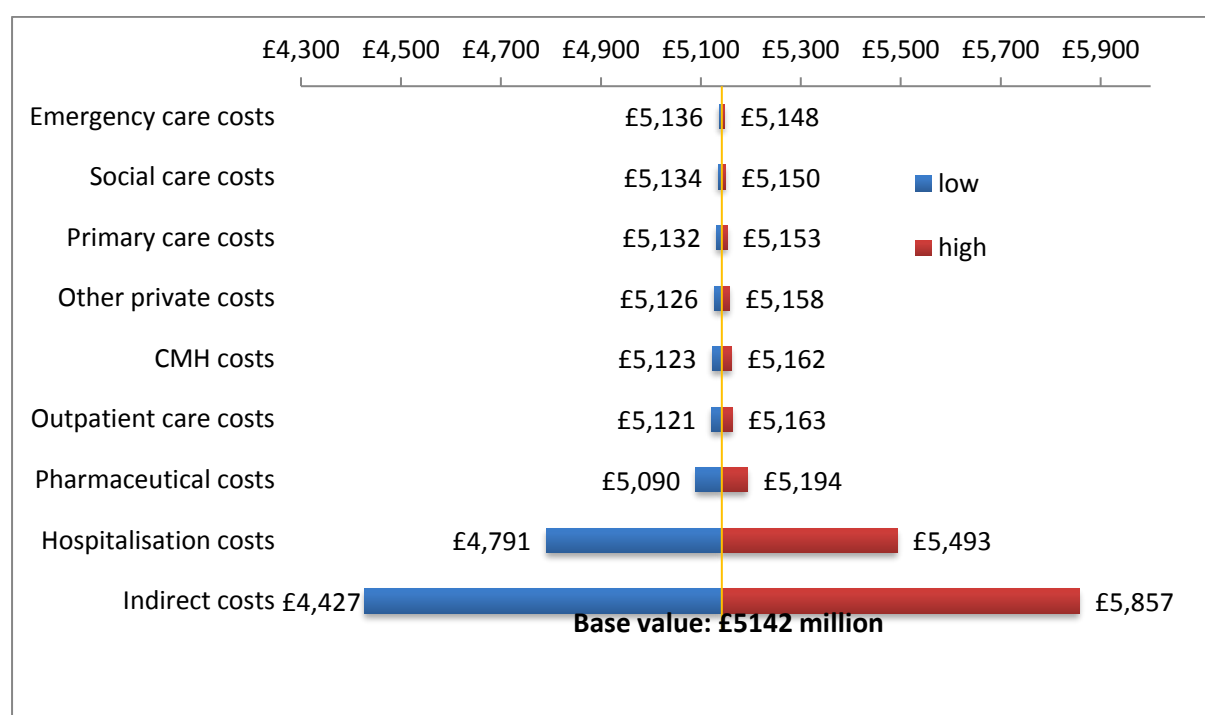
Cost categories	Proportion of total cost	Cost of illness (£ million)
Primary care costs	2.01%	103.2
Community mental health care costs	1.75%	90.0
Emergency care costs	0.32%	16.5
Outpatient care costs	4.15%	213.4
Hospitalisation costs	15.55%	799.4
Pharmaceutical costs	6.81%	350.2
Social care costs	0.47%	24.2
Other private costs	0.88%	45.3
Indirect costs	68.06%	3500.0
Total		5142.2

5.6 Sensitivity analysis

5.6.1 Univariate

As prevalence represents one of the key variables in COI studies, one-way deterministic sensitivity analysis was undertaken using the lower (0.4%) and upper (1.3%) bound of the pooled prevalence estimate. Unsurprisingly, it had the most profound impact on the overall COI with estimates ranging from £2.6 billion (reduction by 50%) to £8.4 billion (increase by 63%). Whilst holding all other variables constant, cost categories were also varied based on their 95% CIs. The results of these one-way sensitivity analyses are presented in Figure 5.5 below.

Figure 5.5: One-way sensitivity analysis of key parameters (in GB£ millions)



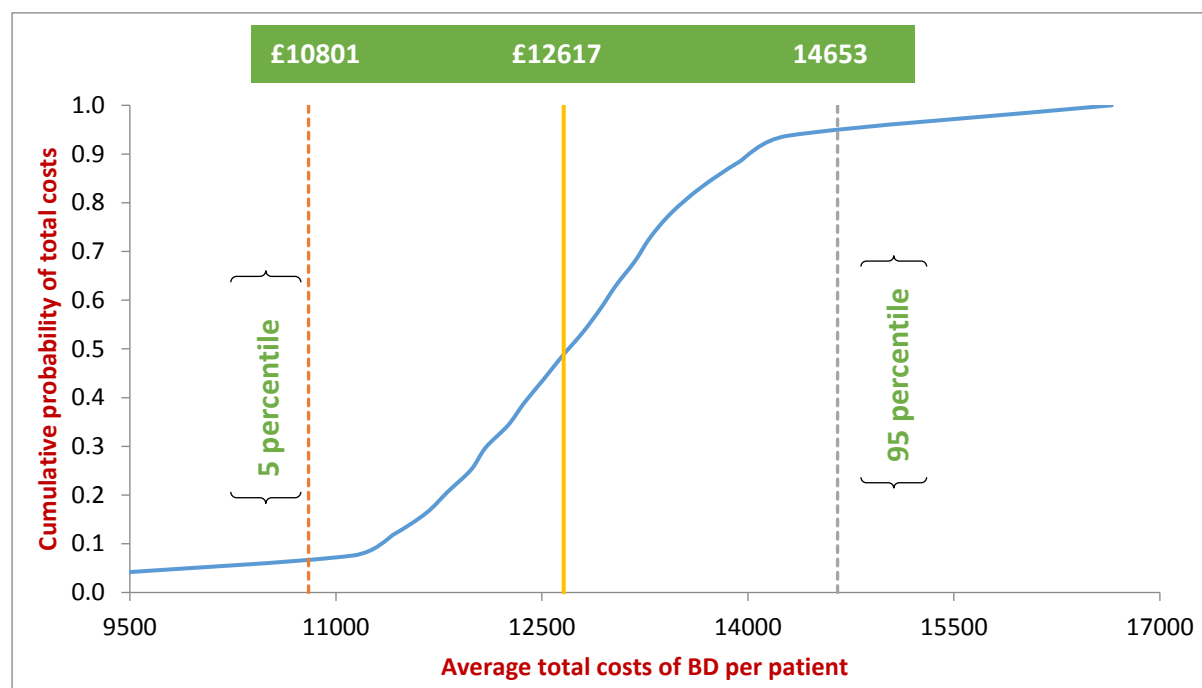
5.6.2 Complete case analysis

As described in **Table 5.2** there were no statistically significant differences between the baseline characteristics of the complete case patients (n=48) and the overall sample (n=91). The complete cases based analysis (**Table 5.4**) estimated the total average annual costs per patient at £11,674 (95% CI: £5583-£17,765). Using these data, the overall total costs of BD in the UK decreased to £4.76 billion.

5.6.3 Probability Sensitivity Analysis (PSA)

Based on the specified distributions of all parameters in the PSA, the 95% confidence interval of the annual total cost per BD patient was estimated being between £10,801 and £14,653 and between £2.71 billion and £8.53 billion for the COI estimate. The cumulative probability distribution of costs per individual is shown in **Figure 5.6**. The estimated median annual total cost per patient was £12646. The observed mean (£12617) annual cost was slightly below the median. This means more than half the individuals with BD will incur costs greater than the mean average annual cost.

Figure.5.6: Probabilistic sensitivity analysis (PSA) results of annuals costs per individual with BD



The Y-axis indicates the cumulative likelihood of observing costs based on a cumulative distribution function. For example, 95% of individuals will have annual costs <£14653

5.7 Determinants of total BD costs

Mean QIDS and ASRM scores over 12 months were significantly associated with the total societal costs. This indicates that a difference between two individuals of one unit in the QIDS score is associated with 7.7%($p=0.012$) increase in the total societal cost for the population and for a single unit difference in the ASRM score, the increase in total costs is 11.6% ($p=0.022$).When healthcare costs were considered, only the association with the severity of mania remained significant at the 5% level, with an unit difference in the mean ASRM score the increase in healthcare costs being 11.5% ($p=0.03$). Indirect costs were more closely related with the severity of depression, an increase in costs of 9.2% ($p=0.036$). Results of GEE analyses are presented in **Table**

5.7.

Table 5.7: Predictors of BD associated costs according to the MI-GEE analysis

VARIABLES	Total costs (1)			Healthcare costs (2)			Indirect costs (3)		
	Exp (coeff)	Exp (SE)	P value	Exp (coeff)	Exp (SE)	P value	Exp (coeff)	Exp (SE)	P- value
BDII	1.041	0.251	0.867	1.176	0.363	0.599	0.931	0.279	0.810
Males	1.016	0.253	0.949	0.880	0.244	0.645	1.164	0.378	0.639
Duration of illness	1.008	0.013	0.570	1.008	0.009	0.421	1.008	0.018	0.638
Mean QIDS score	1.077	0.032	0.012*	1.066	0.041	0.096+	1.092	0.046	0.036*
Mean ASRM score	1.116	0.054	0.022*	1.115	0.056	0.032*	1.119	0.069	0.066+
Time-points									
6th month	0.743	0.159	0.165	0.971	0.222	0.899	0.585	0.202	0.121
9th month	0.813	0.188	0.372	1.087	0.280	0.746	0.615	0.218	0.170
12th month	0.866	0.191	0.513	1.021	0.215	0.923	0.752	0.265	0.419

The coefficients are the results of MI-GEE model using Gamma regression with log-link function

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Time-points: 3rd month as reference category

5.8 Discussion

This chapter presents the up-to-date societal costs for BD in the UK general adult using a prospective pragmatic cohort study design. The chapter has also provided new evidence on the associations between main clinical-demographic factors and healthcare and societal costs for patients with BD. The average annual cost of BD per patient was £12,617 with most of the total costs (68.1%) attributed to productivity losses and informal care, 30.6% to healthcare costs, 0.9% to private out-of-pocket expenses, and 0.5% to social care costs. The average healthcare cost of £3,859 was considerably higher than the Treasury reported average healthcare expenditure of £1,926 per person in the UK showing the increased need of patients with BD for health care services (208).

The estimated prevalence of BD in the UK was 0.8% and the total number of adults (>16 years old) with BD in the UK was estimated to be 407,567 for the year 2010. The overall societal cost for the entire UK population was estimated to be £5.14 billion. One of the key finding from this chapter is that economic costs of BD falls mainly falls outside the healthcare on those who experience the illness and their families. Healthcare expenditure comprises a small proportion of overall costs. This significant result is consistent with the existing literature that have highlighted the substantial burden of lost productivity in individuals with BD (23, 209). Individuals with BD consider employment as one of the major elements in their recovery process (210). It is, therefore, not only important to investigate the cost-effectiveness of long-term management strategies but also interventions that improve the functioning of individuals who are at work.

Secondly, increase in severity of depression or mania resulted in increased total, healthcare and indirect costs. Specifically, costs associated with productivity losses and informal care were significantly associated with depression and healthcare costs with mania. These are important results because they demonstrate a relationship with average symptom level over 12 months; routine monitoring of self-reported health status in BD has a potential to provide a relatively simple method to estimate service performance and improve patient outcomes at the population level (more details in **Section 5.8.2**)

5.8.1 Comparison with other studies estimating the burden of BD in the UK:

COI studies are difficult to compare as they vary in methodology, study setting, and type of costs included. In the UK, Gupta and Guest's retrospective study used a prevalence estimate of 0.5% adopted from the narrative review of literature by Bebbington and Ramana, but it was significantly different from the originally reported population based estimate of 1–1.5% (177, 211). Direct costs in their study accounted for 14% and indirect costs accounted for 86% of total costs in contrast to 30.6% for direct costs and 68.1 % for indirect costs observed in this study (see **Figure 5.7**). This is likely due to the top-down approach they adopted compared with the bottom-up approach used in this study. The increased proportion of indirect costs derived in their study was driven by costs assigned to lost employment (85% of indirect costs) followed by costs due to premature mortality (suicides). The estimates for lost employment were obtained from non-peer-reviewed reports and assumptions. However, the high proportion of individuals who were unemployed and could work (27%) assumed in

Gupta and Guest's study seems to be at odds with much lower rates observed in other relevant studies (8, 212).

Two further issues with their estimates was that they did not include informal care costs and social care costs were estimated based on expert opinion. Informal care costs along with employment losses are particularly important elements in mental healthcare and they place a large burden outside the healthcare sector. Ignoring informal care costs and over-estimating employment losses could have considerable impact on the overall distribution of costs (8). Moreover, their estimates from 1997-98 included health-care resource use of patients who suffered from unipolar psychosis and schizoaffective disorder and this may not be reflective of current practice.

McCrone et al (2008) in their report modelled societal costs of BD in England by combining healthcare costs derived from a sample of 103 BDI patients participating in a randomized controlled trial (RCT) of cognitive behavioural therapy in England, informal care costs adopted from a psychosis study, and productivity losses adjusted from Gupta and Guest's study (2002) along with prevalence data from the US (8, 177, 213). Even though, the broad distribution of costs with regard to proportion of direct (31 %) and indirect (69%) costs in McCrone et al's study is consistent with the present study, their proportion of hospitalisation costs (14%) which was estimated by using inflation, and correction factor (i.e. for BD-II and related disorders) is much lower than the current study and other similar studies (178, 214, 215). A possible explanation for this is that in assessing the service costs they combined bottom-up annualised costs from 3-months healthcare resource use for BDI patients obtained from a RCT with top-

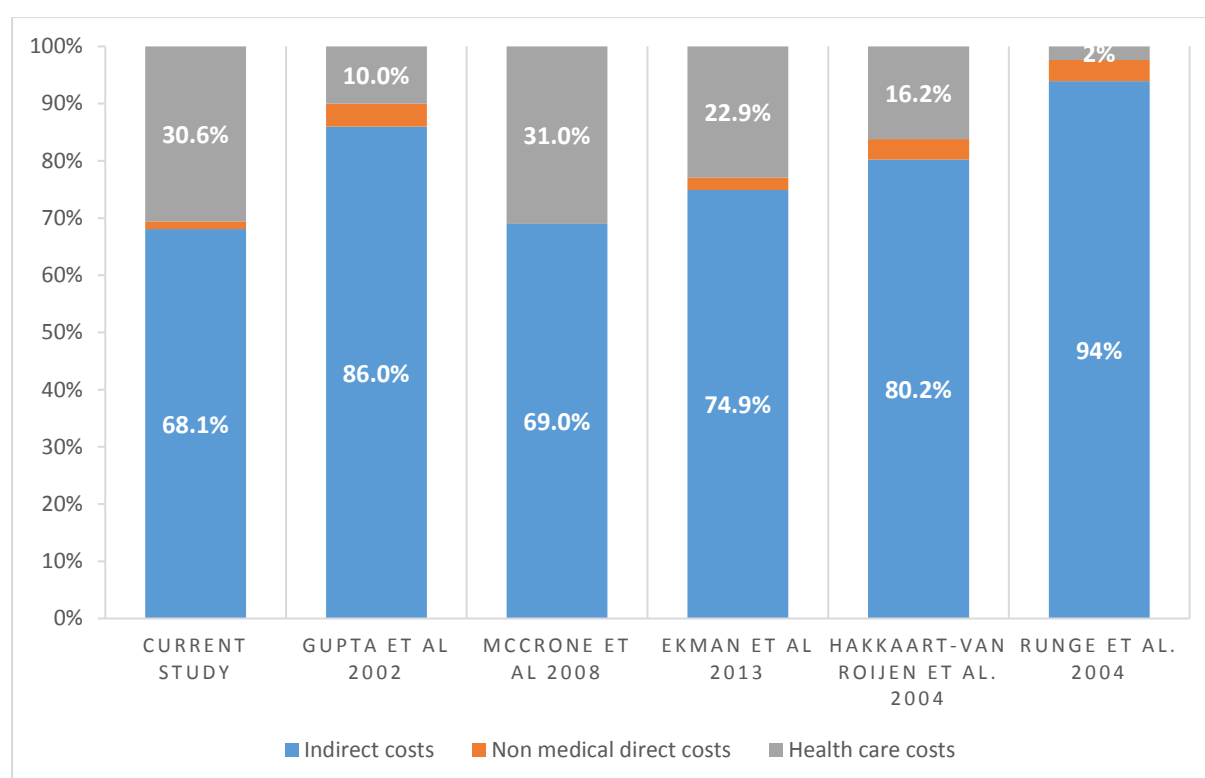
down estimates of hospitalisation (Hospital Episode statistics data) and residential care for BD related conditions. Moreover, for the majority of resource utilisation categories, their estimates were derived from a RCT that was undertaken in 1998-99 with stringent inclusion and exclusion criteria (excluding patients with recent history of active episodes) (213). Such samples may not be representative of individuals with BD in the community.

The most recent research assessing costs of BD in the UK was a top-down study undertaken by Young et al (2011), that included only healthcare costs (178). They reported that the total hospitalisation cost was £207 million corresponding to 60% of all healthcare costs; this translated to £2,370 per BD patient. In comparison, in the current study, hospitalisation related costs was approximately £799 million accounting for 61% of all health care costs related to BD. average hospitalisation cost per patient was £1,961. The 20% higher average hospitalisation costs reported in Young et al (2011) could be explained by the inclusion of admissions falling under the ICD-10 code 30.0 which includes manic/hypomanic episodes on its own. Interestingly, it was observed that the costs attributed to medicines was 7.2 % of total healthcare costs, while the current study reported almost 3 times higher proportional cost at 22% of healthcare costs. The authors acknowledge that their medication cost estimates represent an under-estimation of overall costs attributed to medicines due to lack of prescribing data.

More importantly, Young et al reported a total healthcare cost of £342 million which is much lower than the current study's estimate of £1.31 billion. The prevalence estimate

(0.14%) used in this study was based on the number of patients who contacted their GPs in a defined period; this may have underestimated the prevalence in comparison to community-based prevalence data, as there may still be BD patients who did not access psychiatric care. Additionally, the outpatient and community health services costs calculated in both Young et al (2011) and Gupta and Guest (2002) were not BD patient specific and based on expert opinion.

Figure 5.7: Distribution of annual societal costs of BD in different studies



In the European context, two studies were conducted assessed the societal cost of BD in Germany and Sweden using government healthcare databases (see **Figure 5.7**). Ekman et al (2013) reported that hospitalisations represented 55.0 % of the healthcare costs, and the total annual cost per patient adjusted for the year 2011 and

converted to GBP was £22096 (215). The annual costs per patient in Germany was £8764 with hospitalisations attributing for 62% of healthcare costs (216). Finally, using patient level data, Hakkaart-Von Roijen et al (2004) reported an annual cost of £3638 per person with BD in Netherlands (217) . Overall distribution of costs found in all these studies indicated similar pattern, with indirect costs driving the burden of BD (215, 217).

5.8.2 The relationship between clinical factors and economic costs

Exploration of factors influencing variation in costs was carried out using multivariate regression analysis. Increasing severity of depression or mania resulted in increased total, healthcare and indirect costs. Specifically, indirect costs were significantly associated with depression and direct costs with mania; these are interesting results that are in line with previous studies as mania is often associated with increased hospitalizations, while individuals with predominantly depressive episodes tend to accrue varied resources over a longer period. Begley et al. (2001) conducted an incidence-based COI study in the US where they demonstrated that increasing severity of depressive or manic phases resulted in higher societal cost of illness. Similarly, Perlis and colleagues showed that in a cohort of 1,460 BD patients followed-up for two years, the number of residual manic or depressive symptoms as measured by the Clinical Monitoring Form significantly predicts recurrences, and thus total costs incurred (218, 219).

However, the present study used TrueColours, a method for routine data collection using patient reported symptoms that is regularly and frequently updated. Clinical improvement as demonstrated by decreased scores on the QIDS and ASRM even by a single point has statistically significant and financially important impact on direct, indirect and overall societal costs. This observed association implies a critical clinical message; more effective care should reduce not only the frequency, severity and duration of manic and depressive episodes (as measured by QIDS/ ASRM scores) but also the broader societal costs.

Self-reported outcomes of psychiatric patients have often been regarded as imprecise, subjective and therefore received less attention so far (220). This study suggests that in the case of the bipolar population, self-reported symptom level closely correlates with the wider costs and benefits and could be a measurable outcome also for service success. These results are substantiated by existing research that has not only indicated reliability and validity of self-reported outcomes measures of well-being and disease status in mental illness, but also provided initial evidence on the utility of such self-reported measures in patient-centred care, shared decision-making and performance management of services (221, 222)

5.8.3 Strengths and limitations

The analyses presented in this chapter has several methodological strengths. Firstly, the best method to measure the economic costs is by prospectively collating

information on real expenditure and consumption of health care resources within a cohort of subjects with BD (18).

In this chapter, real-world, self-reported population-based data of BD patients who were followed-up over 12 months was used. This enabled a prospective assessment of clinical status and resource utilisation over the course of concomitant enrolment, and the recording and estimation of societal costs for a cohort of patients at four separate occasions without the need for annualising the cost estimates. This is important as annualised cost estimates may not reflect the true costs associated with severe mental illness like BD that follows an unpredictable episodic course. Furthermore, the bottom-up approach allowed the estimation of all relevant costs including medical, direct non-medical (i.e. social care, travel costs and private expenses), and indirect costs (i.e. informal care costs and productivity losses) incurred by individuals with BD. It also allowed the use of statistical techniques such as GEE modelling for the in-depth exploration of costs attributable to the disease in question.

Secondly, due to the lack of epidemiological studies of BD in the UK, prevalence estimates were derived from a random-effects meta-analysis of population based epidemiological studies of BD in Europe. This yielded a mean prevalence of 0.8% (95% CI = 0.43% to 1.3%), which is consistent with a recently published study that reported the prevalence of BD in the UK using Bayesian meta-regression modelling to be 0.8% (95% CI: 0.7 % to 0.9%) in females and 0.6% (95% CI: 0.5 % to 0.7%) in males (5). Finally, several univariate sensitivity analyses and multivariate PSA were conducted to ascertain the validity of the study results.

However, the study also has limitations and the results need to be interpreted taking these into consideration. Firstly, the relatively small sample of 91 individuals (aged ≥ 16 years) used in this study was recruited from psychiatric clinics across the Oxfordshire and Buckinghamshire counties in England, therefore, the patients and their care may not be fully representative of the entire UK BD population. Nevertheless, the mean age of participants (45.5 years), proportion of males (36.3%), and average primary contacts over a year (10.5, SE:0.90) observed in the current cohort was similar to two recent studies that reported prescribing trends between 1995 and 2009 in 4700 patients and rates of self-harm between 1995 and 2013 in 6671 patients with BD registered in over 470 general practises (comprising over 10 million patients) across the UK (223, 224). Further, research has indicated that COI studies using healthcare databases or claims based data tend to report larger sample sizes, with a range between 40-297,000 individuals (18).

Secondly, as expected in longitudinal studies, missing data was encountered. Although complete case analysis would eliminate any uncertainty caused by missing values, the resulting estimates may be inefficient and often lead to biased results. Moreover, BD patients who fail to complete their follow-up are often those with more severe disease, and therefore the complete case analysis is likely to underestimate costs. For example, the mean QIDS-SR16 score at baseline for individuals who had missing data at any time-point was 8.9, two points higher than the mean QIDS-SR16 for those who provided scores at all five time-points (**see Table 5.3**). Multiple imputation of missing variables assuming a missing at random (MAR) mechanism was utilised to limit the loss of too many observations or variables in this analysis. Despite

the fact that the MAR assumption cannot be verified, MI is the suggested procedure to deal with missing data (225).

Finally, indirect costs were estimated using HCA that is likely to overestimate costs in comparison to the alternative Friction Costs Approach (83). On the other hand, due to the lack of data, productivity losses for unemployed individuals whose unemployment was attributed to BD and premature mortality due to BD were not included. The latter factors may have contributed to some under-estimation of the overall societal costs.

5.9 Conclusion

This study provides robust evidence that BD is a disease of substantial economic burden in the UK. The preponderance of the burden falls on the society as productivity losses and informal care; while hospitalisations represent the next most significant cost burden. The total annual costs for BD in the UK were estimated to be £5.14 billion for year 2010/11, translating to £1.57 billion costs to the healthcare sector in the UK or £1.31 billion in England. This corresponds to approximately 1.1% of the total NHS healthcare expenditure that was £121.45 billion in 2010/11 (226).

Healthcare cost and productivity losses were significantly associated with self-reported levels of symptoms which suggest that such regularly collected self-reported outcomes are also of direct use to health service planners/commissioners.

As described in **Chapters 2 and 3**, it is important to emphasise that the cost of BD outlined in this chapter is to demonstrate the importance of the burden of BD from a public mental health perspective. Nevertheless, as **Chapter 4** highlighted, further cost-effectiveness evidence on alternative management strategies in BD is needed to inform objective judgements that need to be taken at a health-care system level to improve population health. In this context, **Chapter 7** will provide evidence on an economic evaluation of a complex intervention that includes a novel psychoeducation program for individuals with bipolar disorder who use a remote self-monitoring tool of mood.

However, given the methodological challenges with economic evaluations in mental health highlighted in **Chapters 3 and 4**, the next **Chapter 6** will first provide evidence on the appropriateness of using EQ-5D-3L in individuals with BD. EQ-5D-3L is the preferred generic preference-based measure recommended by NICE for conducting economic evaluation in healthcare. The findings based on the next chapter will facilitate a better understanding of an economic evaluation of interventions targeted at individuals living with BD.

Chapter 6: Using preference based measures in bipolar disorder: is the EQ-5D-3L appropriate?

6.1 Introduction

Chapter 4 described the sheer dearth of economic evidence on psychological interventions in the UK; **Chapter 5** highlighted the extraordinary societal burden of BD in the UK, emphasising the need for investing in identifying cost-effective interventions that make the best use of limited societal resources. The critical need of evidence on cost-effectiveness highlighted by **Chapters 4 & 5** requires reliable measurement of health outcomes in individuals with BD.

In routine clinical practise, mental health practitioners have traditionally used disease-specific symptom measures to monitor and evaluate outcomes such as relapse rates and hospitalisations. However, there has been a growing interest in generic patient-reported outcome measures in mental health that captures the health-related quality of life (HRQoL) and overall wellbeing (222, 227). Existing evidence indicates that as these measures are potential indicators of health care quality, they can be utilised across disease areas in several countries including the UK to support clinical, public health, and policy-level decisions (228).

In the UK, the National Institute for Health & Care Excellence (NICE) recommends the use of the generic preference-based measure (PBM) EQ-5D-3L for conducting

economic evaluation and majority of NICE clinical guidelines across all disease areas are based on the cost-utility analysis (4). However, there is insufficient evidence with regards to its appropriateness and the EQ-5D-3L so far has often not been preferred in psychiatric studies, thus precluding the estimation of utilities required for priority setting and resource allocation (229, 230).

As described in **Chapter 3**, the purpose of EQ-5D-3L is to facilitate estimation of Quality adjusted life years (QALYs) which could then be used for economic evaluations. Given the ambiguity around the appropriateness of EQ-5D-3L and the relative reluctance of its use in mental health clinical practise, its psychometric properties need to be assessed to confirm that the EQ-5D-3L is appropriate to measure HRQoL in individuals with BD. This methodological chapter has two parts; the first section investigates the appropriateness of EQ-5D-3L using series of recommended analyses. Given the reluctance of use of EQ-5D-3L, the second section also explores the feasibility of mapping BD-specific mood instruments onto EQ-5D-3L, in the absence of EQ-5D-3L data. The chapter begins by providing a background to the EQ-5D-3L and its use in mental health.

6.2 EQ-5D-3L and mental health

The EQ-5D-3L is the most widely used generic PBM of HRQoL. As **Chapter 3** outlined, EQ-5D-3L can be used for deriving utility weights. However, there are several instruments and elicitation methods to estimate health state utilities, thus, resulting in different QALY estimates that are challenging to compare. Given the need for

consistency across economic evaluations, NICE currently recommends EQ-5D-3L as a tool for quantifying utility and conducting economic appraisals (121).

Despite the importance of EQ-5D-3L, there is debate as to whether or not these generic PBMs capture all important aspects of HRQoL in some disease conditions such as alcohol dependency, limb reconstruction, visual disorders, mental illness and hearing impairment (231).

Specifically, in mental illness, there are concerns that generic PBMs have been primarily designed for acute medical conditions or problems predominantly affecting physical health (232). Mental health practitioners are reluctant to use EQ-5D-3L as it is suggested that EQ-5D-3L fails to cover the impact of changes in mental health problems on the HRQoL (88). The case for the use of generic PBMs is further limited by the distinct paucity of research examining the use of EQ-5D-3L in severe mental illnesses such as BD. Although, there are studies that have reported that the EQ-5D-3L demonstrates construct validity and responsiveness for common mental illnesses such as depression, the results for severe mental illnesses including schizophrenia and psychosis are less convincing (233).

A recent health technology assessment report indicated that in BD, only four studies were identified that provided data on the known group validity and convergent validity of the EQ-5D-3L instrument. The evidence was mixed, with three tests supporting the validity and three psychometric tests failing to provide evidence across the four studies. However, there were no studies assessing the responsiveness of EQ-5D-3L

(ability to detect changes over time). Further, these studies did not find the EQ-5D-3L reliable or valid in patients who experience manic or hypomanic episodes (30).

In spite of the popularity of EQ-5D-3L, it is common that in mental health research, disease-specific measures are the dominant method of routine outcome measurement (234). Failure to include generic PBMs, however, makes it impossible to derive utilities for conducting an economic evaluation. As a second best, there are statistical methods known as “mapping” recommended by NICE that could be utilised to predict health state utilities from disease specific scales provided there is overlap in content between the measures of interest (235). However, at the same time use of an instrument that is not sensitive enough to reflect changes in clinical status would mean that the resulting economic evaluation can be misleading and in the worst-case scenario may lead to erroneous policy decisions.

To address these methodological challenges associated with outcome evaluation in BD, the aims of the present study are two-fold: a) to comprehensively appraise the appropriateness of the EQ-5D-3L in patients with BD and b) explore the feasibility of “mapping” EQ-5D-3L from disease-specific mood instruments used in BD including the Quick Inventory of Depressive Symptoms–Self Report scale (QIDS-SR₁₆) and the Altman Self-Rating Mania scale (ASRM).

6.3 Methods

6.3.1 Participants and study design

The data for the current study was obtained from a pragmatic randomised controlled trial of facilitated integrated mood management (FIMM) versus manualised integrated mood management (MIMM) in BD (OXTEXT-6 trial). Full details of the study are reported elsewhere (236). Briefly, the study looked at evaluating effectiveness of two different types of integrated delivery of care in the management of BD. They were a) individualised psychoeducation in the form of a manual facilitated by mental health practitioners and b) individualised psychoeducation manual self-administered by patients. Both groups also participated in weekly mood monitoring through SMS- or web-based remote mood monitoring system called “True Colours” (TC) (for further details see www.truecolours.nhs.uk) (35).

Participants for the OXTEXT-6 trial were recruited from patients with confirmed diagnosis of BD receiving mental health services at clinics run by the Oxford Health NHS Foundation Trust (OHFT). Ethical approval was granted from the local Research Ethics Committee (Oxfordshire REC A, Reference: 10/H0604/13).

Inclusion criteria were: individuals a) aged over 16 years with a current diagnosis of bipolar disorder I (BDI) or bipolar disorder II (BDII) as per the DSM-IV-TR criteria and not in a current mood episode, b) willing to provide informed consent, and c) reliably

participating in remote mood monitoring (at least 4 responses in the 4 weeks prior to invitation to participate).

Overall, data were obtained from a sample of 121 UK patients with BD over 12 months recruited from April 2011 until December 2012. The baseline questionnaire included data on the EQ-5D-3L, baseline clinical scores, socio-demographic variables along with other study related variables. Baseline data were complete for all the clinical outcomes (i.e. QIDS-SR₁₆, and ASRM) and the EQ-5D-3L.

Participants also provided weekly clinical symptom scores (QIDS-SR₁₆ and ASRM) and monthly EQ-5D-3L responses by answering text or email prompts at TC monitoring system over a 12 months follow-up period. Symptom scores were extracted from the system for the week in which participants were randomised (W0), and for the subsequent 52 weeks (i.e. W1 to W52). EQ-5D-3L responses were extracted from the month in which participants enrolled (M0) over the subsequent 12 months of enrolment (i.e. M0-M12). This means that in an ideal situation, each individual would have sent 52 pairs of clinical symptom scores (QIDS-SR₁₆ and ASRM), and 12 monthly EQ-5D-3L measurements post-baseline. Two individuals withdrew before the follow-up could start, and only baseline data was available for them.

For this study, clinical measures +/-7 days of monthly EQ-5D-3L responses were extracted. The number of repeated observations (SMS- or web-based) per patient ranged from a minimum of two to a maximum of 13 with an average of 10.6 repeated observations per patient. For instance, out of a possible 1547 EQ-5D-3L responses,

1257 observations were available for analyses indicating an overall 18.7% missingness. Participants from both treatment groups were pooled for the methodological analyses presented here.

6.3.2 Outcome instruments

Clinical severity of depression was assessed by self-reported depressive symptoms using the Quick Inventory of Depressive Symptoms–Self Report (QIDS-SR16) scale over a period of 12 months. QIDS-SR16 is a validated scale comprising of 16 items, covering the nine criterion symptom domains used for characterising a major depressive episode. Each domain is ranked 0-3 that are then added to yield an overall score ranging from 0-27; higher scores corresponding to increased severity of depression (180). Further, a cut-off score of greater than 10 indicates an active depressive episode (236). Mood elevation was measured by the Altman Self-Rating Mania (ASRM) scale which consists of five items scored 0-4 with higher scores indicating increased severity of mania. In the literature, a cut-off point of less than or equal to 5 is indicative of ‘being not in a manic state’, whereas a score of greater than 5 indicates a ‘state of mania’ (179).

HRQoL was measured with the EQ-5D-3L. The EQ-5D-3L consists of two parts - a health status instrument containing five items representing the dimensions of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression and a visual analogue scale (EQ-VAS) measuring self-rated health on a scale between 0 (worse imaginable health) to 100 (best imaginable health state). For the main instrument, items are marked on a three-point Likert scale: 1=no problem; 2=some

problems; and 3=extreme problems, yielding a possible 243 health states. Each state is scored using a preference-based weight derived from the UK time-trade-off (TTO) tariff. The tariff set was elicited from a sample of 3395 members of the UK general population. The states can range from severe problems (33333) to no problems (11111) on all five items, which translate to a utility range between –0.594 and 1 (full health) respectively, and where death equals zero (27).

6.4 Statistical methods

6.4.1 Appropriateness of the EQ-5D-3L

According to Fitzpatrick et al, there are several criteria for assessing PROMs including appropriateness, reliability, validity, responsiveness, dimensionality, interpretability, acceptability, and feasibility. Although appropriateness is an essential characteristic of a PROM, no standard psychometric tests are available to directly evaluate this criterion. Nevertheless, appropriateness can be estimated indirectly by measuring reliability, validity, and responsiveness (112, 237).

Broadly speaking, there are two alternative paradigms of assessing validity of instruments including the 'Classical Test Theory' (CTT) framework and the item response theory (IRT). CTT assumes all items are of equal importance in measuring the underlying latent construct and the distance between different item levels are same across all items. CTT primarily relies on item-total correlations (or the item-rest) and provides a measure of internal consistency.

On the other hand, IRT uses mathematical non-linear functions to describe the association between an individual's level on a trait (for example: HRQoL) and the probability of responding positively to a specific level on an item (238). In this study, appropriateness of EQ-5D-3L in relation to the clinical scales used in BD was premised on CTT framework and assessed in line with NICE technical recommendations (237). EQ-5D-3L profile of the current patient sample was summarised using descriptive statistics and contrasted with the general population; the general population EQ-5D-3L data (n=7294) was obtained from the representative 2012 Health Survey for England (239).

6.4.2 Reliability

Reliability is the ability of an instrument to provide consistent and reproducible measurement results free from random error. In the context of patients completing the same PROM on multiple occasions, the instrument should provide repeatable scores with a high degree of agreement across time-points in individuals with apparently no change in their health status (113).

This type of reliability is also called as test-retest reliability and was assessed using two-way fixed effects model of Intra-Class Correlation (ICC) (240). ICCs were evaluated comparing baseline and follow-up measurements of the EQ-5D-3L and the EQ-VAS scores in individuals who reported no change in their health status as measured by the clinical measures (i.e. QIDS-SR16 and ASRM scores). The QIDS-SR16 and ASRM scores were used as clinical anchors to split the 'unchanged' group

into two sub-groups with individuals who stayed either in remission or active disease. Patients in remission had a QIDS-SR16 total score ≤ 10 and an ASRM < 6 . On the other hand, patients were deemed to have a relapse or active disease if they reported > 10 on QIDS-SR16 or ≥ 6 on ASRM scores (236). The ICC quantifies the strength of agreement between repeated measures by estimating the proportion of total variance of an outcome that is attributable to the between individual variance.

variance due to rated subjects (between subjects)

$$\text{ICC} = \frac{\text{variance due to rated subjects (between subjects)}}{(\text{variance due to rated subjects} + \text{variance due to repeated time-points (within)} + \text{residual variance})}$$

Higher ICC indicates a greater degree of homogeneity of responses on outcome variable within individuals and that the variability in measurements is instead associated with the variance between individuals. Therefore, ICC close to 1, implies a high level of reliability; on the contrary, ICC close to 0 indicates that the variance due to random error and repeated measurements within individuals is high resulting in low reliability. Further, the adequacy of ICCs was assessed following Cicchetti and Sparrow's (1981) guideline. Values > 0.75 indicate excellent reliability, 0.6 to 0.74 shows good reliability; 0.4 to 0.59 indicates fair to moderate, and below 0.4 is poor reliability (241).

6.4.3 Validity

Validity is the extent to which an instrument measures what it is supposed to measure. Construct validity is one of the most important types of validity that is amenable to psychometric assessment. This assesses the degree to which the internal structure of a multidimensional instrument measures the underlying construct/traits that it was designed to estimate (112). In the absence of a gold standard HRQoL measure, assessment of validity is a challenging and ongoing task that requires aggregating increasing evidence to establish the appropriateness of an instrument (113). In this study, validity was explored through assessment of convergent validity, and known-group validity (KGV).

Convergent validity is the extent to which the dimensions of an instrument (for example: EQ-5D-3L) relates to other instruments that measure similar or overlapping constructs (113). Convergent validity of the EQ-5D-3L descriptive system was assessed by analysing the association between the response level of EQ-5D items and the dimension scores of the clinical measures used in the study. Similarly, association between the EQ-VAS and the EQ-5D-3L index scores with composite scores of depression as measured by QIDS-SR16, and mania as measured by ASRM scale were also assessed. It was hypothesised that there was an association between the EQ-5D-3L and clinical measures used to measure psychopathology in BD patients.

Specifically, the main assumptions were that there would be a) strong correlation between the EQ-5D-3L dimension 'anxiety/depression' and the nine dimensions of

depression as measured by QIDS-SR16, and b) some kind of association between the two EQ-5D-3L dimensions 'anxiety/depression' and 'usual activities' with the dimensions (i.e. being active, need for sleep, being happy or cheerful) of mania as measured by ASRM. The correlation between these measures were assessed using Spearman's correlation coefficients. As per Cohen (1992), correlation was considered small for Rho values between 0.10 and 0.29, moderate for $0.3 \leq |\text{Rho}| < 0.7$, and a large association for greater than 0.70 (242).

Known group validity is premised on the principle that individuals with different level of severity of disease or different group of patients may expect to be rated differently from each other (113). Scores on the QIDS-SR₁₆ range between 0-27; they can be further categorised as indicating no [0–5], mild/moderate [6-10], moderate [11-15], severe [16-20] and very severe depression [21-27]. Because of the low numbers in some groups and for having a reasonable discrimination between known groups, the patients were broadly collapsed into three groups including [0-5] as no depression, [6-15] as mild/moderate depression and [16-27] as severe/very severe depression (243). As there were repeated measures of outcomes, linear mixed effect regression models adjusted for potential confounders (age, gender, type of BD and concomitant mania scores) were used to compare the EQ-5D-3L utility values across the clinical subgroups. To allow for clustering of data within each individual, a random intercept was introduced for individuals in the regression model.

Further, a similar linear mixed effect regression analysis was carried out for two classes of participants based on their ASRM scores on mania: ASRM scores ≤ 5

indicating a state of 'not in a manic state', while scores >5 implying 'state of mania' after adjusting for possible confounders (i.e. age, gender, type of BD and concomitant depression scores as rated by QIDS-SR16). Standardised effect size (ES) was estimated to assess the magnitude of difference between different sub-groups (more details about ES under the responsiveness section).

6.4.4 Responsiveness

Responsiveness measures the extent to which the EQ-5D-3L measures clinically meaningful changes in HRQoL over time. Responsiveness was assessed by estimating effect sizes (ESs: ratio of the mean change from baseline to the standard deviation of the initial measurement) and standardised response mean (SRM: ratio of the mean change to the standard deviation of that change) in groups defined by anchor-based criteria (244). The responses at baseline, 3, 6, 9, and 12 months were used to estimate the magnitude of the change between the time points.

All patients were classified into three different groups (i.e. better, worse and stable health states) based on the changes in two clinical anchors between any two consecutive time-points, with analyses restricted to individuals who reported improvement or deterioration in their health status over time. The depression score as measured by the QIDS-SR16 and the mania scores as measured by the ASRM were used as the two clinical anchors. In line with existing recommendations and literature, a relevant change was defined as greater than ± 0.5 standard deviations (baseline) corresponding to the minimal clinically meaningful difference for patients (245-248).

Individuals within the ± 0.5 standard deviations of baseline clinical score represented the unchanged, or stable health category, while positive values greater than this interval was classified as 'worse health status' and negative values lower than this interval represented 'better health status'. The main assumptions were that the EQ-5D-3L index and EQ VAS scores would change over time with changes in health as measured by the clinical measures used in this study. As per the guidance on interpretation of the magnitude of change, <0.20 is indicative of trivial change, ≥ 0.20 to <0.50 of a small change, ≥ 0.5 and <0.80 as medium and ≥ 0.80 is viewed as a large change (249).

6.5 Feasibility of mapping EQ-5D-3L from disease-specific instruments in BD

In the current study cohort of patients, three scales QIDS-SR16, ASRM & EQ-5D-3L were used to measure three constructs including depression, mania, and HRQoL respectively. In the absence of EQ-5D-3L data, mapping can be used to derive health state utilities from disease specific questionnaires (121). However, to assess whether disease-specific scales QIDS-SR16, and ASRM could be mapped to EQ-5D-3L, it is paramount that these instruments have some overlapping domains or have no major conceptual differences (235). The accuracy of mapping is reliant on the extent of theoretical and conceptual similarities. Feasibility of mapping was explored through exploratory and confirmatory factor analyses and reported as per the “The MAPS Reporting Statement” (250).

6.5.1 Exploratory factor analyses (EFA) and confirmatory factor analysis (CFA)

Both these techniques are two different but complimentary approaches commonly employed to assess the extent of uni-dimensionality or conceptual similarity. EFA is especially useful in situations where the dimensionality of a construct measured by instrument/instruments is unknown in a given population. EFA taps into the covariance matrix (between items or scales) and extracts unobserved factors from data without specifying the number of factors or determining which items load onto specific factors(251). This analysis enables identification of items with strong correlations among all items within the same domain (i.e. HRQoL) and weaker correlations with items in different domains (112). Four different EFAs were undertaken to explore the conceptual similarities between QIDS-SR₁₆ and ASRM, QIDS-SR₁₆ and EQ-5D-3L, ASRM and EQ-5D-3L and all three scales together.

A maximum likelihood based EFA was analysed. The optimal number of factors was determined by Horn's parallel analysis and confirmed using a parallel analysis plot. In this technique, Monte Carlo simulations are run to extract eigenvalues from random data matrices. Factors are retained whose eigenvalues based on real data is larger than the value obtained from simulations. Following the EFA, factor rotation with an oblique Promax rotation was used, allowing for correlation between potential factors. If the resulting correlations were lesser than 0.32 (10% overlap in variance among factors) then the model was re-estimated using an orthogonal Varimax rotation. Items with factor loadings >0.4 were considered adequate, >0.55 good and >0.70 excellent. The greater the loading, the more of the variance in the observed item is explained by the latent construct. Items are considered poor a) if their factor loading is <0.4, b) they

cross-load on different factors with loadings above 0.4, but none above 0.5 on a single factor, c) items that have high uniqueness >0.7 (error variance that is not accounted by all extracted factors) or lower communalities (amount of variance shared among a set of items that can be explained by a set of factors) (252).

On the other hand, **confirmatory factor analysis** (CFA) is a type of structural equation modelling and is one of the most powerful statistical tools for confirming whether a single latent construct is able to explain the observed variability in multi-dimensional items. Instead of merely exploring the patterns of correlation structure, CFA offers a formal method of testing the hypothesis that items have overlapping domains and measure the same underlying construct (113). Confirmation of conceptual similarities is fundamental for deriving an appropriate algorithm for “mapping” between the different instruments (235).

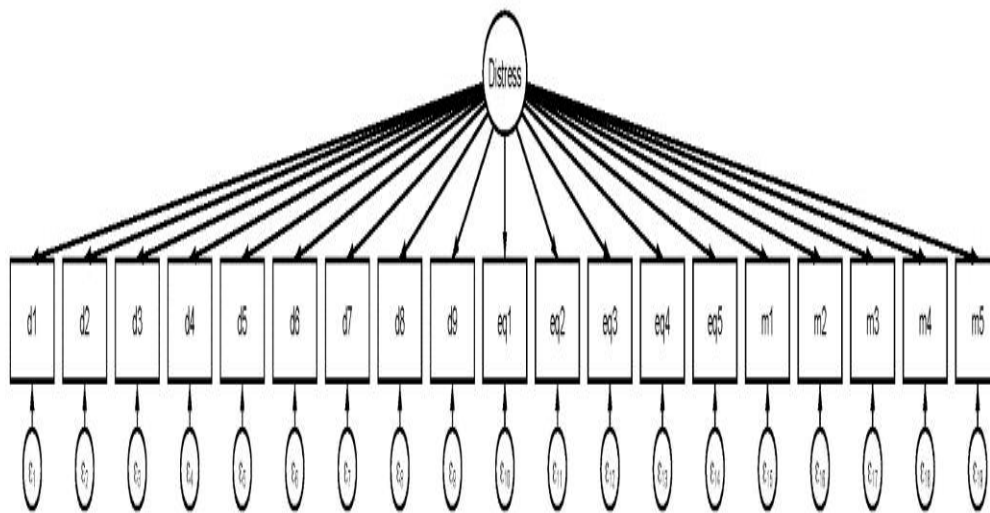
CFA was carried out to validate the models estimated through EFA. As all the three primary outcome measures were ordinal and skewed, weighted least squares means and variance adjusted (WLSMV) estimator that used polychoric correlations between items was used to derive standardised factor loadings and confirm the factor structure. A one-factor model was assessed to explore uni-dimensionality or conceptual similarity of scales and then more factors (two followed by three correlated factors) were added one at a time as guided by the a priori EFA. In addition, bi-factor models were tested to confirm the dimensionality of all the three scales. A bi-factor orthogonal model comprises of a) an overarching single general factor (for example:

psychological distress or wellbeing) that accounts for the commonality among all the items, and b) specific factors (i.e. depression, HRQoL, and mania) that represent instrument-specific item response variances that are not accounted for by the general factor (253). This model provides a useful single confirmatory framework for accounting the overlap and separability of the three included instrument domains.

In total, five different models were considered to determine the extent to which the items of QIDS-SR₁₆ and ASRM overlap with the target dimensions of EQ-5D-3L (**Figures 6.1-6.4**) provides a simplified presentation of the main models): 1) unidimensional (psychological distress), 2) two first-ordered correlated factors (negative affect, mania), 3) three first-ordered correlated factors (depression, HRQoL, mania), 4) bi-factor orthogonal model with a general psychological distress factor (assumption) and individual domain factors (i.e. depression, HRQoL, mania), and 5) an improvised bi-factor orthogonal model with a general psychological distress factor (assumption) and individual domains of depression and HRQoL.

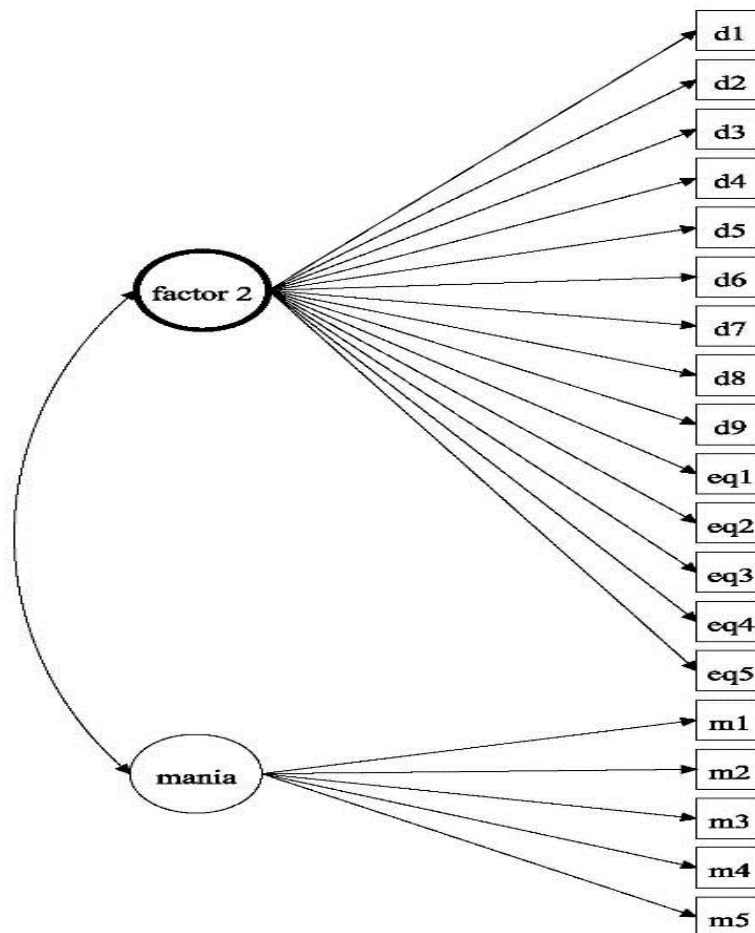
The validity of the CFA to obtain a good fit between the target model and the observed data was confirmed using 1) multiple fit indices including a) a non-significant chi-square value ($p > .05$), b) Root Mean Square Error of Approximation (RMSEA) values < 0.06 , c) the Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI) > 0.95 , d) weighted root mean square residual (WRMR) < 0.90 , and e) explained common variance (ECV) for bi-factor models. ECV is an index of unidimensionality and a score of > 0.70 indicates that the common variance between items can be regarded as essentially unidimensional (251, 254).

Figure 6.1: Unidimensional model with one latent construct (psychological distress)



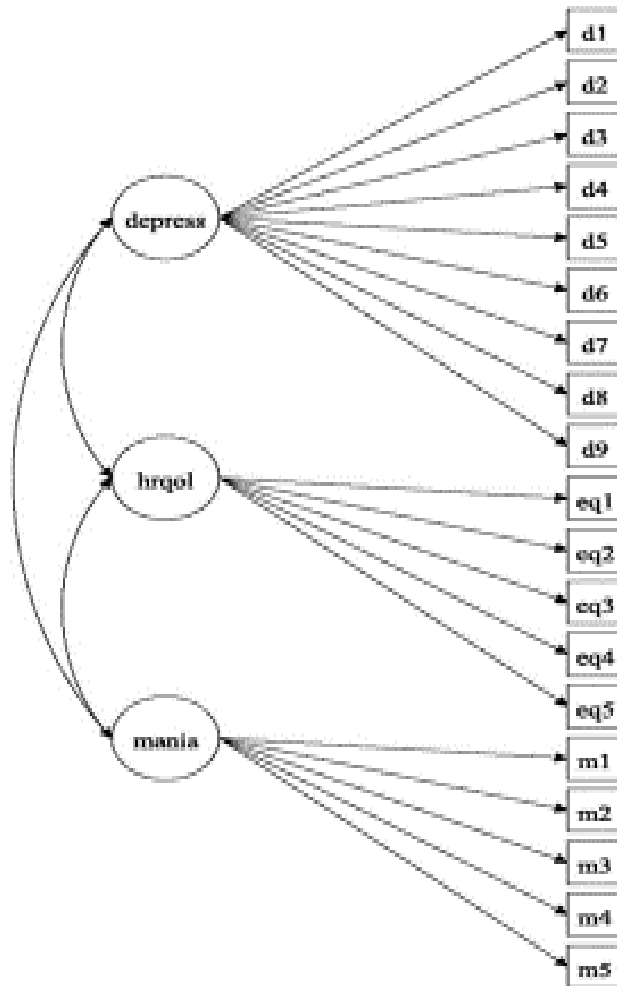
*d1-d9 stand for depression items (QIDS-SR16), eq1-eq5 stand for EQ-5D-3L items, m1-m5 stand for mania items (ASRM)

Figure 6.2: Two first-ordered correlated factors



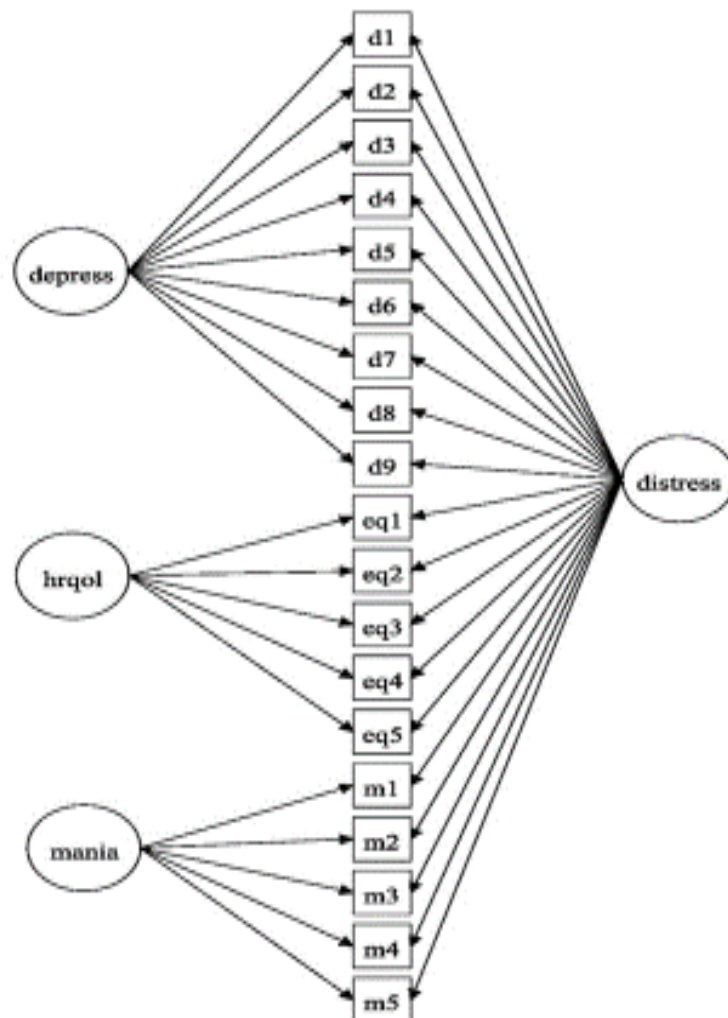
*Two correlated factors: factor 2 representing overlapping depression and HrQOL and the second factor mania

Figure 6.3: Three first-ordered correlated factors (depression, HRQoL, mania)



*d1-d9 stand for depression items (QIDS-SR16), eq1-eq5 stand for EQ-5D-3L items, m1-m5 stand for mania items (ASRM)

Figure 6.4: Bi-factor orthogonal model with a general psychological distress factor and individual domain factors (i.e. depression, HRQoL, mania)



As multiple tests of hypotheses were applied, appropriate Bonferroni correction was used to estimate the level of significance. For instance, for test-retest reliability, the level of significance was corrected to account for ten pair-wise comparisons ($\alpha=0.05/10= 0.005$). Further, all analyses were undertaken using a robust Huber-White sandwich estimate of variance that took into account clustering of responses within individuals. Statistical analyses were performed using the STATA version 14 and MPLUS 7.4 (255, 256).

6.6 Results

The baseline sociodemographic characteristics of the patients are displayed in **Table 6.1**. The mean age of the analysed sample was 44.3 years (SE: 1.24). The majority of participants were female (72.7%), white-British (92.6%), and suffering from BDI (65.3%).

Table 6.1: Baseline sociodemographic characteristics of the study

	Total (121)	FIMM (60)	MIMM (61)	Difference (FIMM-MIMM)	P value
Women (%)	72.7	70	75.4	-5.4	0.504
White-British (%)	92.6	90	95.1	-5.1	0.287
Type of BD (%)					
BDI	65.3	65.0	65.6	-0.6	0.947
BDII	34.7	35.0	34.4	1.1	0.947
Age (mean, SE)	44.3 (1.24)	45.4 (1.78)	43.2 (1.72)	2.17	0.380
Baseline clinical outcomes (mean, SE)					
Baseline Altman score	3.2 (0.36)	3.00 (0.51)	3.36 (0.51)	-0.36	0.621
Baseline QIDS-SR₁₆ score	7.42 (0.50)	7.61 (0.72)	7.23 (0.70)	0.38	0.707
Baseline Health economic outcomes (mean, SE)					
EQ-5D index	0.87 (0.02)	0.87 (0.02)	0.86 (0.02)	0.01	0.835
VAS score	70.6 (1.76)	68.7 (2.52)	72.5 (2.46)	-3.76	0.288

6.6.1 Descriptive statistics of the EQ-5D item distribution

In this study sample, 75 out of 243 possible EQ-5D health states were observed and 20 health states contributed to >90% of observations (please refer Table 6.2). The most frequently observed health state was 11111 indicating no problems at all (33.1% of the total 1257 responses), followed by 11112 (12.3%) and 11212 (8.9%).

Table 6.2: The 20 most frequently (> 90% of 1257) reported EQ-5D-3L health states					
EQ-5D-3L-index	Health states	Frequency	% of Total	Cumulative %	Average VAS
1	11111	485	38.58	38.58	80
0.848	11112	155	12.33	50.91	69
0.812	11212	112	8.91	59.82	61
0.796	11121	71	5.65	65.47	74
0.725	11122	44	3.5	68.97	67
0.689	11222	35	2.78	71.76	57
0.620	21222	29	2.31	74.07	53
0.883	11211	28	2.23	76.29	71
0.691	21221	25	1.99	78.28	65
0.708	12212	21	1.67	79.95	42
0.516	22222	21	1.67	81.62	46
0.727	21121	20	1.59	83.21	72
0.378	11213	19	1.51	84.73	42
0.587	22221	15	1.19	85.92	66
0.255	11223	13	1.03	86.95	46
0.656	21122	13	1.03	87.99	62
0.760	11221	11	0.88	88.86	71
-0.239	22333	10	0.8	89.66	25
0.082	22223	8	0.64	90.29	31
0.414	11113	6	0.48	90.77	30

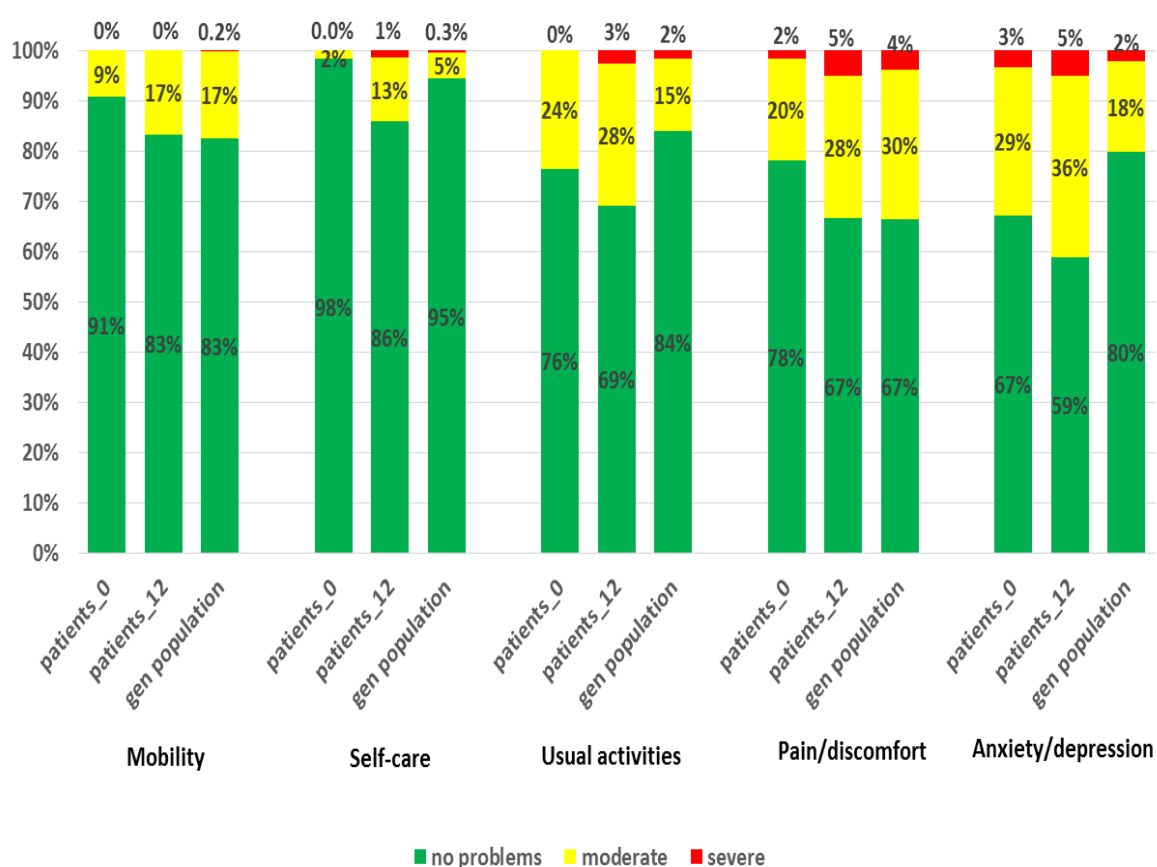
Health states: 1-no problems, 2 – moderate problems, 3 – severe problems

The mean EQ-VAS value for 11111 was 80, which was 20 points below the upper boundary of best health state. The poorest health was 32333 (utility index score: -0.484) with the mean EQ-VAS value of 10, which was 10 points above the lower bound

of worst health state (anchored at 0 for the EQ-VAS). On closer inspection, 48.7%, 24.4%, 27.7%, 32.7% and 31.9% of participants reported full health on the EQ-5D (state of 11111) at the baseline, T3, T6, T9 and T12 time-points respectively. In comparison, 57.1%, 47.9%, 49.6%, 53.8%, and 43.7% of individuals were in remission as defined by the QIDS-SR16 and ASRM threshold scores at the baseline, T3, T6, T9 and T12 time-points respectively.

For all the EQ-5D-3L individual dimensions, responses were predominantly reported as no problems or some problems, indicating good or moderate health status (see **Figure 6.3**). At baseline, in the patient sample, most frequently reported problems were in the item 'anxiety/ depression' (32%), 24% in the item 'usual activities', 22% in the item 'pain/discomfort', followed by 'mobility' (9%) and 'self-care' (2%). In comparison, at the 12th month time-point, although patients reported higher proportion of problems in all the EQ-5D-3L dimensions, only problems in the item 'self-care' was significantly higher than at baseline (14% vs. 2%, $p < 0.001$). Responses were also compared with the general population; at both time-points, patients with BD reported significantly more problems in the EQ-5D dimensions 'anxiety/ depression' and 'usual activities' ($p < 0.05$, see **Figure 6.3**). At the baseline, patients also reported significantly fewer problems in the EQ-5D dimensions 'mobility' and 'pain/discomfort' than the individuals in the general population ($p < 0.01$, please see Figure 6.5).

Figure 6.5: Distribution of responses to items of EQ-5D in patient sample at baseline and 12 months in comparison to the general population



6.6.2 Test-retest reliability

Test-retest reliability was assessed based on five repeated measurements on participants, three months apart including baseline, 3, 6, 9 and 12-month time points; ICCs were estimated for five adjacent paired outcomes. For instance, of 119 patients with data at baseline, 90 had responses at the third month time-point (please see Table 6.3). Of the 90 patients who had data at both time-points, 20 individuals had change in their health status as objectively measured by clinical scales and therefore were not

included in the test-retest reliability estimation. Test-retest reliability was estimated for 71 individuals who were classified in the ‘unchanged’ category.

Table 6.3: Test-retest Reliability of the EQ VAS score and EQ-5D-3L index scores anchored by no change in the BD health states

Time-points	Number of participants in unchanged state		Intraclass correlation coefficient			
			EQ-VAS		EQ-5D-3L	
	remission	active	remission	active	remission	active
t0-t3 (3 m)	45	26	0.20	0.33	0.73*	0.22
t3-t6 (3 m)	42	18	0.32	0.52	0.67*	0.58
t6-t9 (3 m)	45	18	0.44*	0.34	0.64*	0.20
t9-t12 (3 m)	42	14	0.74*	0.77*	0.32	0.13

*significant at $P < 0.005$

Of those individuals whose status remained unchanged (stayed in the same health state) at the third month time-point, 45 participants were under remission and 26 individuals had active (relapse) disease. The reliability was good and significant in individuals under remission ($ICC=0.73$, $p < 0.005$).

Overall, in individuals under remission, EQ-5D-3L index scores demonstrated good reliability in three out of four pairwise comparisons. On the other hand, except for the comparison at t9-t12 ($ICC=0.74$, $p < 0.005$), poor to moderate reliability was observed in the EQ-VAS for most of the pairwise comparisons. In patients with active disease,

the ICCs for the EQ-5D-3L were poor and statistically insignificant, while EQ-VAS demonstrated excellent reliability at the t9-t12 comparison.

6.6.3 Convergent validity

Mean EQ-5D-3L index and EQ-VAS scores were found to be moderately correlated with all the dimensions of the QIDS-SR16 measuring different aspects of depression (ranging between -0.27 and -0.56 for the EQ-5D-3L index and between -0.24 and -0.56 for the EQ-VAS score) (please refer to Table 6.4). The strongest correlation was found between the EQ-5D-3L index score and the overall QIDS-SR16 score (-0.64).

Table 6.4: Correlation matrix between the QIDS-SR₁₆ (depression), EQ-5D-3L, and VAS

Items	EQ-5D-3L	Mobility	Self-care	Usual activities	Pain/discomfort	Depression/anxiety	VAS
QIDS-SR₁₆	-0.64	0.23	0.34	0.58	0.28	0.69	-0.63
Insomnia/ hypersomnia	-0.41	0.18	0.23	0.36	0.25	0.36	-0.36
Appetite/weight change	-0.27	0.13	0.21	0.25	0.16	0.25	-0.24
Psychomotor slowing/agitation	-0.50	0.24	0.29	0.51	0.26	0.51	-0.49
Sad mood	-0.51	0.08	0.26	0.45	0.14	0.67	-0.55
general interest	-0.51	0.16	0.30	0.51	0.16	0.59	-0.54
Energy/ fatigability	-0.54	0.23	0.31	0.51	0.26	0.53	-0.51
Self-outlook	-0.54	0.21	0.34	0.49	0.22	0.60	-0.56
Concentration	-0.56	0.18	0.28	0.52	0.21	0.63	-0.53
Suicidal ideation	-0.46	0.15	0.30	0.43	0.19	0.51	-0.48

All pairwise correlation coefficients were statistically significant at P value<0.05; highlighted coefficients indicate lowest and highest coefficients

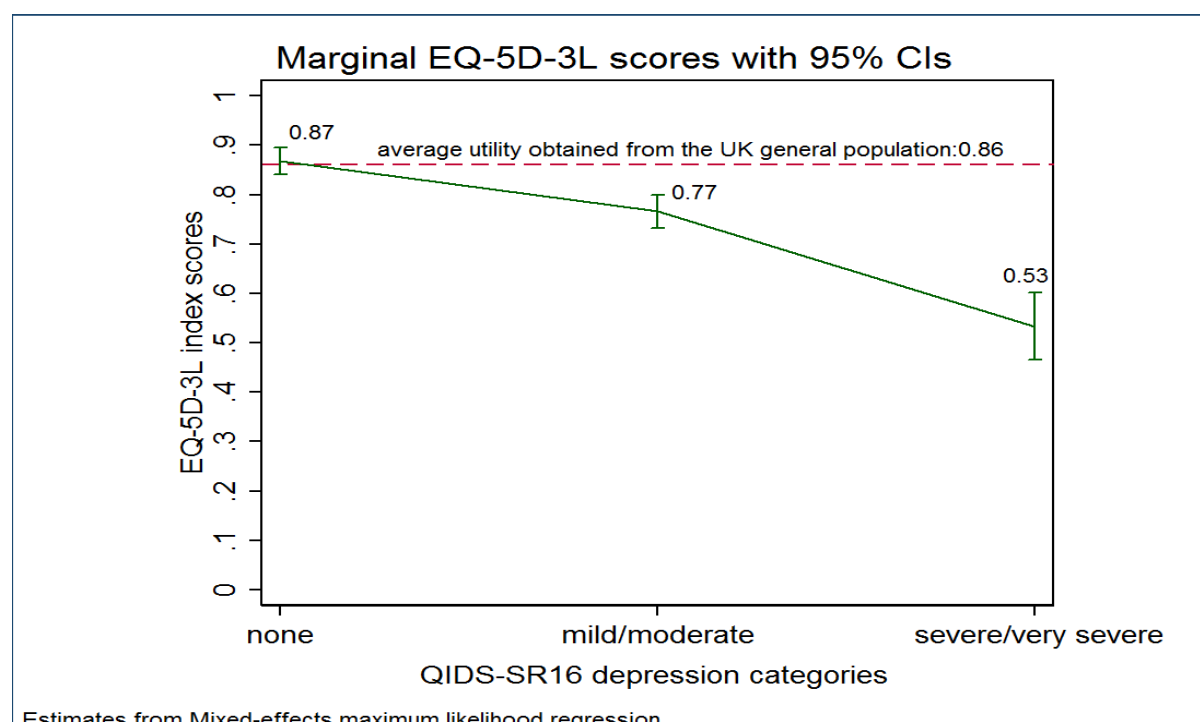
The Spearman rank correlations between the QIDS-SR₁₆ items and the EQ-5D-3L items tended to be small to moderate (range 0.08–0.69). The strongest correlation (0.67) was found between the QIDS-SR₁₆ item “sad mood” and the depression/anxiety dimension of the EQ-5D-3L. Mobility and pain/discomfort dimensions of the EQ-5D-3L had the weakest relationships with the QIDS-SR₁₆ items. The convergence between EQ-5D-3L/EQ-VAS and ASRM scores was extremely poor with the largest significant correlation of -0.216 between the ‘depressions/anxiety’ item of the EQ-5D-3L and ‘elevated mood’ item of ASRM (please see **Appendix 8**).

The correlation between the ASRM and the QIDS-SR₁₆ items were low (<0.2) with overall correlation between the ASRM total score and the QIDS-SR₁₆ being -0.07 (results are not reported).

6.6.4 Known-group validity

After adjusting for potential confounders including age, type of BD, gender, concurrent ASRM score, EQ-5D-3L index scores showed statistically significant difference across different levels of depression severity. **Figure 6.6** describes the EQ-5D-3L scores in different categories of depression. For individuals in the ‘not depressed’ category, the predicted mean EQ-5D-3L score was 0.867 (95% CI: 0.840-0.894).

Figure 6.6: Known-group validity of EQ-5D-3L in different levels of depression



The score for individuals belonging to the ‘mild/moderate’ depression category was 0.766 (95% CI: 0.732-0.799), and that for individuals in ‘severe’ depression category was 0.515 (95%CI: 0.446-0.584). While, the standardized effect size estimated between ‘none’ and ‘mild/moderate’ severity groups is moderate (0.69), the difference between the mild/moderate and severe/very severe depression groups is a very large effect (1.57), implying that there is evidence that the EQ-5D-3L can reliably differentiate between different levels of depression severity (see **Table 6.5**).

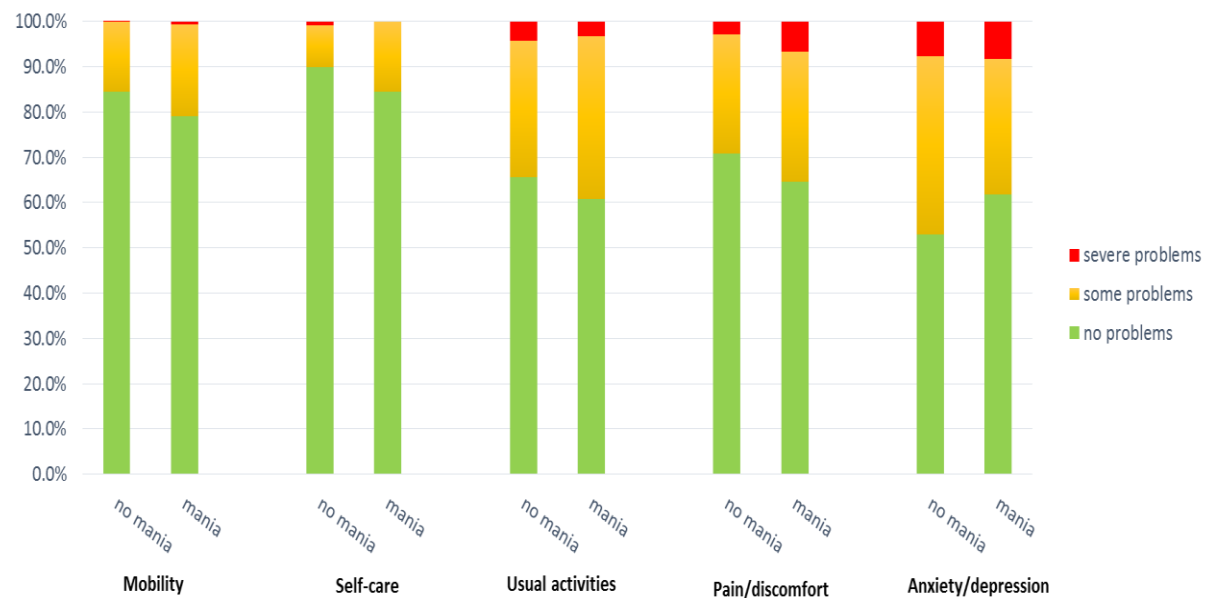
Table 6.5: Known-group validity of the EQ-5D-3L in different levels of depression and mania

Depression	Observations (%)	Mean EQ-5D-3L utility (SE)	95% CI		ES
none	610 (49.2%)	0.868 (0.013)	0.840	0.894	
mild/moderate	480 (38.7%)	0.766 (0.017)	0.733	0.799	0.69
severe	150 (12.1%)	0.532 (0.054)	0.465	0.600	1.57
Mania					
none	1059 (85.4%)	0.786 (0.014)	0.758	0.813	
mania	181 (14.5%)	0.802 (0.026)	0.749	0.853	-0.08
ES: Effect size between adjacent severity groups					

The initial regression model unadjusted for any confounders found a significant difference in mean utility score between the non-manic state in contrast with the manic phase (0.777 Vs 0.833, $p=0.036$), with scores higher in the manic state. However, no significant difference was observed in mean EQ-5D-3L index scores across different levels of mania after controlling for age, type of BD, gender, and depression categories (0.798 vs 0.784, $p=0.111$). EQ-5D-3L reflects differences in depression but apparently not in mania (**Table 6.5**).

Figure 6.7 depicts the distribution of EQ-5D-3L responses in each of its dimension across the two states, i.e., non-manic and manic health states.

Figure 6.7: Distribution of responses on individual EQ-5D-3L dimensions for mania sub-categories



In general, the distribution of scores across three dimensions of EQ-5D-3L including mobility, usual activities, and pain/discomfort was similar in both the severity groups of mania. However, there were two significant differences; firstly, individuals who were not in manic health state reported a lower proportion of problems (some problems or severe problems) in the self-care dimension than those with mania (10% vs 16%, $p=0.03$). Importantly, the proportion of problems experienced on the anxiety and depression dimension decreased with increasing level of mania; the manic sub-group experienced nine percent lower problems than those in the non-manic health state (38% vs 47%, $p=0.02$).

Table 6.6: Responsiveness of the EQ-5D-3L index and the EQ VAS to changes in health status as anchored by change in QIDS-SR16 scores

Better health state							Worse health state					
Period 1: T0 to T3	N	mean (T0)	mean (T3)	mean change	ES	SRM	N	mean (T0)	mean (T3)	mean change	ES	SRM
EQ-5D-3L	28	0.807	0.768	-0.039	-0.17	-0.12	20	0.897	0.653	-0.244**	-1.95	-0.92
EQ VAS	28	61.32	69.57	8.3	0.35	0.37	20	69.85	51.25	-18.6**	-1.25	-0.80
Period 2: T3 TO T6	N	mean (T3)	mean (T6)	mean change	ES	SRM		mean (T3)	mean (T6)	mean change	ES	SRM
EQ-5D-3L	22	0.624	0.814	0.190**	0.76	0.76	18	0.858	0.686	-0.172*	-0.67	-0.64
EQ VAS	22	53.4	68.9	15.5**	0.94	1.11	18	75.4	56.8	-18.7**	-0.98	-0.82
Period 3: T6 TO T9		mean (T6)	mean (T9)	mean change	ES	SRM		mean (T6)	mean (T9)	mean change	ES	SRM
EQ-5D-3L	14	0.616	0.877	0.261*	0.71	0.65	18	0.737	0.592	-0.146	-0.5	-0.47
EQ VAS	14	54.8	73.1	18.3*	0.61	0.77	18	69	54.2	-14.8**	-1.23	-0.90
Period 4: T9 TO T12		mean (T9)	mean (T12)	mean change	ES	SRM		mean (T9)	mean (T12)	mean change	ES	SRM
EQ-5D-3L	22	0.602	0.767	0.165*	0.47	0.60	9	0.798	0.721	-0.077	-0.41	-0.24
EQ VAS	22	55.7	66.0	10.3**	0.57	0.82	9	75.6	62.8	-12.7*	-1.13	-0.86

QIDS-SR16: Quick Inventory of Depressive Symptoms–Self Report (QIDS-SR₁₆) scale; * $p \leq 0.05$, ** $p \leq 0.005$. A change in the QIDS-SR16 score was defined by more than 0 ± 0.5 SD (baseline).

6.6.5 Responsiveness

The responsiveness of EQ-5D-3L and EQ-VAS was assessed by estimating effect sizes (ESs) and standardised response means (SRMs). ESs and SRMs ≥ 0.80 represented large magnitude of change. **Table 6.6**, describes average scores and the mean change scores derived from the difference between neighbouring repeated assessments at the baseline, three, six, nine and twelve months in subgroups anchored by QIDS-SR16. For example, of those who provided measures at both baseline and third-month time-point ($n=90$), the majority of individuals (45%) remained in the 'stable' group, while 28 reported improvement in their health status and 20 deterioration in their health status. In those who reported improvement in their health status, there were no significant differences in both the EQ-5D-3L and the EQ-VAS mean change scores along with low SRM scores of -0.12 and 0.37 respectively.

Overall, in the patients reporting improved health, the EQ-5D-3L index indicated a statistically significant moderate ES and SRMs in most of the period wise comparisons, implying patients reported higher utility in the predicted direction. At the same time, the EQ-VAS scores were more responsive than the EQ-5D-3L with significantly large SRMs at two time-point comparisons (see Fig 6). For patients reporting deterioration in their health, decrement in utility or EQ-VAS scores was observed in the expected direction (**Fig 6.8**). ESs and SRMs were larger for the EQ-VAS scores than the EQ-5D-3L at all time points except at the first pairwise comparison (T3-T0), where EQ-5D-3L was more responsive than the EQ-VAS with greater ES and SRM.

Figure 6.8: Responsiveness as measured by SRM in the EQ-5D-3L index and the EQ-VAS to changes in health status as anchored by change in QIDS-SR16 scores

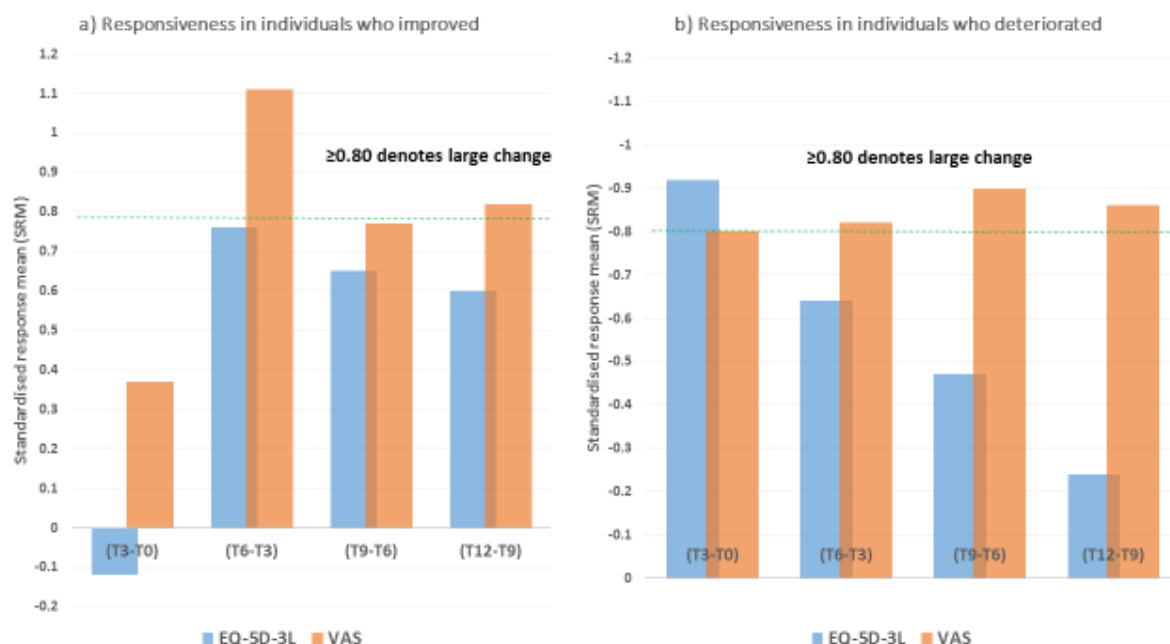


Table 6.7, shows the change in the mean EQ-5D-3L and the EQ-VAS scores in subgroups reporting improved or deteriorated ASRM scores. In individuals who reported an improvement in ASRM scores, implying a decrease in severity of mania, both the EQ-5D-3L and the EQ-VAS scores declined. Although the pattern of change in the mean EQ-5D-3L and the EQ-VAS scores was the opposite of the predicted direction, this change was moderate and significant for both scores only in two out of four comparisons, with similar ES and SRMs. In the group indicating a worsening of ASRM scores, implying individuals with increasing severity of mania, the EQ-5D-3L and the EQ-VAS scores showed an increase. However, this increase in the EQ-5D-3L scores was significant only at the comparison between T9 and T6, with a moderate ES and SRM. Similarly, the EQ-VAS scores showed a significant increase only at the comparison between T12 and T9, with a moderate ES and SRM.

Table 6.7: Responsiveness of the EQ-5D-3L index and the EQ VAS to changes in health status as anchored by change in ASRM scores

Better health state							Worse health state					
Period 1: T0 to T3	N	mean (T0)	mean (T3)	mean change	ES	SRM	N	mean (T0)	mean (T3)	mean change	ES	SRM
EQ-5D-3L	33	0.883	0.762	-0.121*	-0.71	-0.47	15	0.792	0.714	-0.078	-0.28	-0.24
EQ VAS	33	73.2	67.5	-5.7	-0.29	-0.21	15	68.7	61.9	-6.9	-0.33	-0.40
Period 2: T3 TO T6		mean (T3)	mean (T6)					mean (T3)	mean (T6)			
EQ-5D-3L	16	0.760	0.583	-0.177*	-0.52	-0.62	22	0.732	0.839	0.107	0.44	-0.64
EQ VAS	16	68.6	57	-11.6*	-0.57	-0.66	22	69.1	73.8	4.7	0.24	0.27
Period 3: T6 TO T9		mean (T6)	mean (T9)					mean (T6)	mean (T9)			
EQ-5D-3L	23	0.808	0.744	-0.064	-0.23	-0.21	15	0.626	0.843	0.217*	0.59	0.56
EQ VAS	23	74.9	66.8	-8.1*	-0.56	-0.47	15	58.66	70	11.3	0.48	0.47
Period 4: T9 TO T12		mean (T9)	mean (T12)					mean (T9)	mean (T12)			
EQ-5D-3L	12	0.713	0.713	0.00	0	0	19	0.806	0.924	0.118	0.46	0.41
EQ VAS	12	73.8	71.3	-2.5	-0.14	-0.15	19	71.1	80.6	9.5*	0.72	0.69

ASRM: Altman Self-Rating Mania Scale; * $p \leq 0.05$, ** $p \leq 0.005$. A change in the ASRM score was defined by more than 0 ± 0.5 SD (baseline).

6.6.6 Exploratory factor analysis (EFA)

The underlying dimensional structure of QIDS-SR16, ASRM AND EQ-5D-3L were investigated together. The Horn's parallel analysis (**Appendix Figure 1**) indicated that there are three factors underlying the 19 items and have an eigenvalue of greater than one, explaining 83% of the total variance.

Table 6.8 presents the item loadings on each factor after the varimax rotation. Each factor was composed of two to twelve items. The three-factor solution on the 19 items indicated that almost all had significant loadings of above 0.4 on at least one of the factors. The first factor included 12 items: all the original items of the QIDS-SR₁₆ and three items of the EQ-5D-3L. EQ-5D-3L dimension 'depression/anxiety' loaded exclusively onto the first construct, but dimensions 'self-care', and 'usual activities' cross-loaded onto the third factor.

'Appetite/weight changes' (factor loading: 0.372) and 'sleeping pattern items from the QIDS-SR₁₆ loaded poorly onto the first factor and had a high unique variance (>0.70), indicating that these items are poor in explaining factor I. Factor II included five single loading items exclusively reflecting mania as measured by ASRM, implying no overlap with other two constructs. Factor III included EQ-5D-3L dimensions: mobility and pain/discomfort that had strong single loadings and two cross loading items 'self-care', and 'usual activities'. The 'usual activities' had a higher substantive cross-loading (>0.5) onto the first factor than the third factor. However, the 'self-care' item did not load >0.5 on either of the factors, implying that it is a poor item.

Table 6.8: Overview of factor loadings and uniqueness based on an exploratory factor analysis with varimax rotation

Items	Item description	Factor loading	Uniqueness	Eigen-values
CONSTRUCT I				
QSAD	Feeling sad	0.880	0.216	6.28
QINTERST	General interest	0.837	0.272	
QSELFVU	View of Myself	0.821	0.298	
QCONCEN	Concentration/decision making	0.794	0.350	
EQ5D_5	Anxiety/depression	0.735	0.430	
QENERGI	Energy levels	0.725	0.381	
QSUICID	Suicidal ideation	0.712	0.473	
QPSYCHMO	Feeling slowed down/restlessness	0.665	0.485	
EQ5D_3	Usual activities	0.572	0.447	
QSLEEP	Sleeping pattern	0.430	0.740	
EQ5D_2	Self-care	0.354	0.630	3.13
QWT	Appetite/weight gain	0.337	0.844	
CONSTRUCT II				
MALT2	Increased self-esteem	0.872	0.232	3.13
MALT1	Elevated mood	0.843	0.264	
MALT5	Psychomotor agitation (activity levels socially, sexually, at home, work or school)	0.794	0.363	
MALT4	Pressured speech	0.772	0.399	
MALT3	Decreased need of sleep	0.556	0.665	
CONSTRUCT III				
EQ5D_1	mobility	0.744	0.433	1.15
EQ5D_4	pain/discomfort	0.657	0.534	
EQ5D_2	self-care	0.495	0.630	
EQ5D_3*	usual activities	0.475	0.447	

* Item is more strongly associated with construct I, bold factor loadings represent items that are poor observed measures of latent constructs I and III

6.6.7 Confirmatory factor analysis (CFA)

A series of CFAs were also estimated. **Table 6.9** provides the results of model fit indices that describe the fit of different item factor models to the polychoric correlation matrix among the items. The single unidimensional model was the worst fitting model indicating a single factor could not describe the multi-dimensional concept among all the included items. The standard two- and three-factor models fared better than the single factor model, but did not reflect the association among target items as well as the bi-factor models. In contrast, the bi-factor model demonstrated closer fit to the observed data and therefore were examined in greater detail below.

Table 6.9: Confirmatory factor analysis, goodness-of-fit indices of five models

Model	Items	χ^2	df	CFI ^a	TLI ^a	WRMR ^b	RMSEA ^c	RMSEA lower CI	RMSEA upper CI
Unidimensional	19	2842.1	152	0.682	0.643	5.915	0.119	0.115	0.123
Correlated 2-factor	19	665.84	151	0.939	0.931	1.763	0.052	0.048	0.056
Correlated 3-factor	19	643.53	149	0.942	0.933	1.674	0.051	0.047	0.056
Bi-factor factor	3- 19	385.95	133	0.971	0.962	1.191	0.039	0.034	0.044

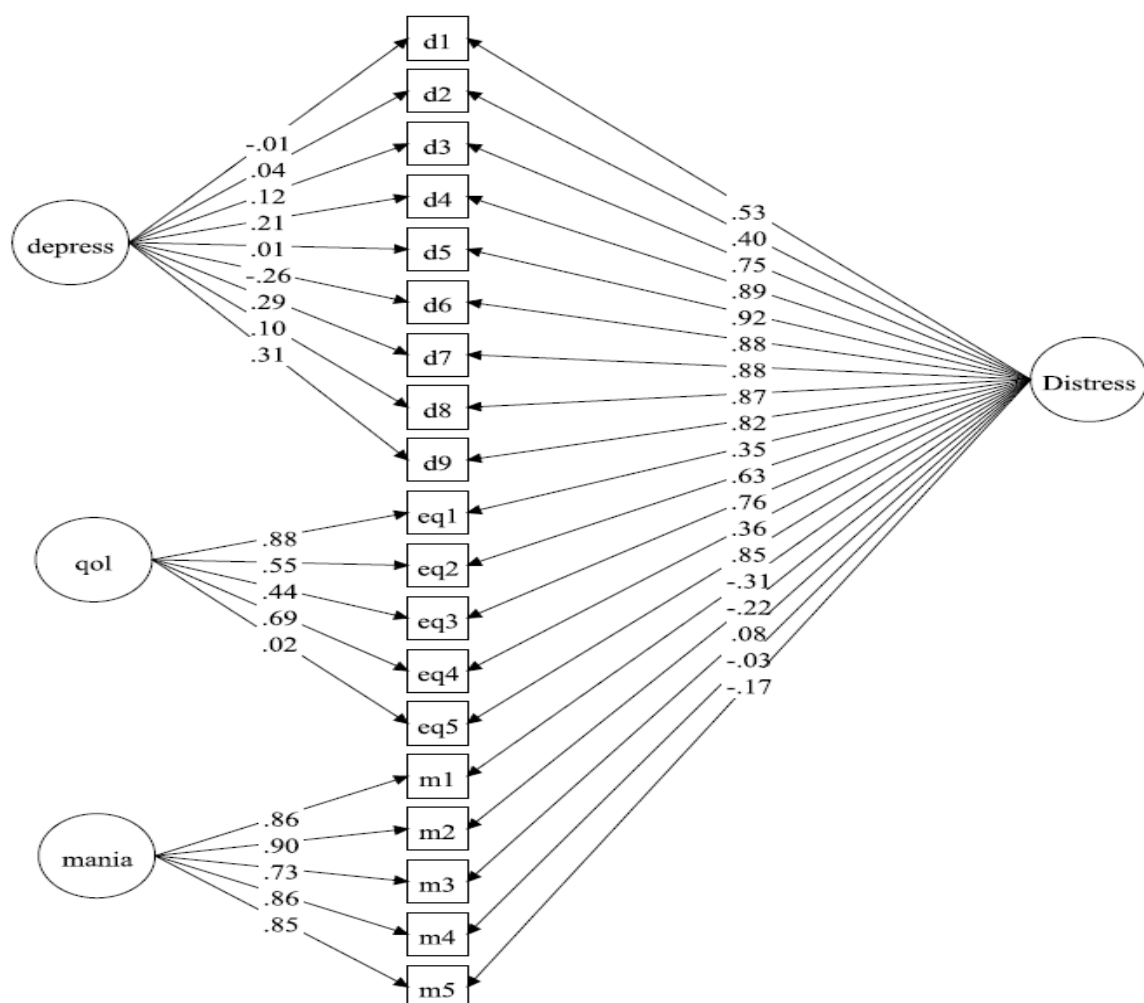
χ^2 chi-square fit statistic; df degrees of freedom; CFI comparative fit index; TLI Tucker-Lewis Index; WRMR weighted root mean square residual RMSEA Root Mean Square Error of Approximation; CI confidence interval

a: higher values indicate better model fit; b: <0.9 indicate better model fit; c: lower values indicate better model fit

In the bi-factor model with three specific factors (**Figure 6.9**), 9 observed variables (d1 to d9) were used as indicators of the latent factor depression, five items (m1-m5) represented the factor mania, and five observed variables (eq1-eq5) reflected the

latent construct health-related quality of life. The general psychological distress factor explained the covariance among items of all the three scales while the domain specific constructs explained the variance of the items within the three specific factors. The observed indicators or items were presented in boxes, and latent constructs or factors in circles. Relationships between variables are indicated by lines. A line with a single arrow represents a hypothesized direct relationship between two variables, with the arrow-tip pointing towards the direction of variable influenced by another variable.

Figure 6.9: A bi-factor CFA model with a general psychological distress factor along with three factors representing three included instruments



All the depression items loaded poorly on the 'depression' specific factor but had significant loadings ranging between 0.40 and 0.92 on the general distress factor. EQ-5D-3L items mobility (eq1) and pain/discomfort (eq4) had significant strong loadings on its own HRQoL domain, but poorer loading on the general factor. Items 'self-care', and 'usual activities' loaded reasonably on both specific and the general factor. The 'depression/anxiety' item (eq5), however, failed to load significantly on its own specific factor, but had one of the highest loadings on the general factor. This indicates that the eq5 item measured completely different aspects of the latent HRQoL in contrast to other items of the same instrument (eq1-eq4); It had more similarity with the items on the depression scale.

The estimated common variance (ECV) for this model was 58%, implying that the common variance in all the items belonging to three scales was not sufficiently explained by the general factor of psychological distress. Further, mania items display no overlap with the general factor and taps into aspects beyond that measured by the other two scales. Taken together, all depression items and three items from EQ-5D-3L had overlap and stronger associations with the general psychological distress factor than with their respective specific domains, implying that these items demonstrated conceptual similarity.

6.7 Discussion

Despite the widespread use of the generic preference-based measure EQ-5D-3L in measuring quality-adjusted-life-years for economic evaluations, its usefulness in

severe mental illness has been challenged, and the evidence supporting its appropriateness is limited. Further, it is not uncommon that only disease-specific instruments are used to assess outcomes in mental illness, thus precluding the possibility of using cost-utility analyses to guide resource allocation among competing disease priorities in publicly funded health-care systems (30). To address the above-stated issues, this study explored a) the appropriateness of EQ-5D-3L in patients with Bipolar Disorder by employing recommended methods of validation, and b) the feasibility of ‘mapping’ disease-specific instruments used in Bipolar disorder onto EQ-5D-3L by utilising exploratory and confirmatory factor analysis methods.

This study has demonstrated that the EQ-5D-3L is useful in measuring HRQoL in BD patients who predominantly experience depressive symptoms, but is not sensitive enough to detect changes in individuals with varying levels of mania, therefore not reflecting the whole spectrum of the disorder. Results of both explanatory and confirmatory factor analyses indicated that multidimensional measures of depression, HRQoL and mania cannot be considered sufficiently unidimensional to allow mapping estimation of health-related utility.

6.7.1 Appropriateness of the self-rated EQ-5D-3L in people with Bipolar disorders

The assessment of validity of EQ-5D-3L in BD was challenging due to the complex spectrum of symptomatology; the evidence on appropriateness was mixed. The feasibility (completion rates) at baseline was 100% for all the three instruments.

6.7.2 Reliability

Reliability is an important measurement property as it indicates the amount of errors of the measurement. Reliability was quantified by ICCs which is a ratio of variability between subjects to the total variability of all measurement in the sample. Thus, higher ICCs reflects the lower variation of measurements within individuals (238). While, in the individuals reporting remission the EQ-5D-3L showed good to moderate levels of test–retest reliability across three out of four time-periods, it failed to demonstrate any significant reliability in individuals who remained in active disease state across the comparison time-points. On the other hand, the EQ-VAS scores indicated similar moderate to good reliability in individuals under remission across two time-periods, and also demonstrated ‘excellent’ reliability among patients with active disease in the fourth time-period comparison (t9-t12). This relative difference in reliability could be due to the small number of individuals in the active disease group.

Further, the lengthy average time interval between the repeated measures in this study could have resulted in a change in people’s health status or the spectrum of patient’s mania/depression symptomatology, thereby lowering reliability. So, despite staying in the same active state, the symptomatology could be different resulting in more varied responses to a multi-dimensional EQ-5D-3L instrument than to a uni-dimensional EQ-VAS. Thus, the EQ-5D-3L seems to be moderately reliable, specifically in patients under remission.

6.7.3 Construct validity as reflected by convergent and Known-group validity (KGV)

Correlations between the EQ-5D-3L and disease-specific measures indicated non-existent to moderate Spearman rank correlation coefficients, with the relationship between the EQ-5D-3L and the ASRM being 0.092, and that of between the EQ-5D-3L and QIDS-SR16 being -0.64 . Importantly, correlations with both the disease-specific measures were mainly driven by the anxiety/depression and to a lesser extent by the 'usual activities' domain. This association is unsurprising as the 'anxiety/depression' item has a direct relationship to one of the fundamental diagnostic criteria of BD.

The linear mixed level analyses revealed that EQ-5D-3L discriminates well between the broad three different severities levels of depression as measured by the QIDS-SR16; higher levels of depression sub-categories had lower EQ-5D scores. Although the number of responses in each sub-group was smaller in the initially classified five categories of depression, the KGV was also explored in the five sub-groups. The results were significantly different ($P < 0.001$) in each group with the EQ-5D scores of 0.88 for the euthymic category (no depression), 0.81 for mild depression, 0.69 for moderate depression, 0.52 for severe depression and 0.40 for those with severe depression, as assessed by the QIDS-SR16. The utility value for euthymic patients in this study was similar to the average of the UK general population (0.86) (257).

However, EQ-5D-3L failed to differentiate between different levels of mania as measured by the ASRM, with the group without mania scoring 0.786 against 0.802 for individuals with mania ($P=0.859$). Further analyses of what might be driving individuals in manic health state overestimating their HRQoL indicated that people with mania significantly reported lower depressive symptoms than those not in manic health state. Interestingly, these adjusted EQ-5D-3L scores in the manic state were still lower than that observed in the general population, implying the presence of residual depressive or mixed symptoms. Considering the KGV of the EQ-5D-3L in both the depressive and manic states together, there is limited evidence that the EQ-5D-3L has discriminative ability across the entire spectrum of BD symptomatology. Nevertheless, the EQ-5D-3L has demonstrated good construct validity in BD patients predominantly with depression, and the findings are consistent with results from other research that have substantiated the construct validity of EQ-5D-3L in depression/anxiety (258).

6.7.4 Responsiveness

Two measures of responsiveness, Effect size (ES) and standardised response mean (SRM) were used to assess the ability of the EQ-5D-3L and the EQ-VAS to detect clinically significant within-individual variability in a patient's condition. SRM is considered a better measure of responsiveness as it is reliant on estimation of within-subject variation (113).

The condition-specific QIDS-SR16 and ASRM instruments were chosen as clinical anchors. As there is no minimal clinically important difference (MCID) reported for these two scales, categorizations were based on the 0.5 SD criteria which correspond

to the MCID observed across several condition-specific PROM (259) . On using the QIDS-SR16 scale as a clinical anchor, in the patients reporting improved health, the EQ-5D-3L index was moderately responsive in three out of four period-wise comparisons, while in individuals with worsening symptoms, the EQ-5D-3L demonstrated moderate to large responsiveness on two pair-wise comparisons. This relative difference in performance could be due to the small sample of BD patients who reported worsening or may be that patients value small improvements in health more than decline. This asymmetry of responses has been reported in other disease areas (244).

On the other hand, the results indicated that the EQ-VAS was more responsive (larger SRMs, see **Figure 6.8**) to both improvement and deterioration of depressive symptoms compared with the EQ-5D-3L in most of the pair-wise responsiveness analyses. In particular, the EQ-VAS appeared to be more sensitive to change in individuals whose depressive symptoms worsened with large SRMS across all the pairwise comparisons. This apparent difference in responsiveness between the EQ-5D-3L and the EQ-VAS could be because of general population's relative insensitiveness to change in BD patient's reported EQ-5D-3L domains and thereby result in under-estimation of population preference-based utility weights. Further, the EQ-5D-3L descriptive system is limited by three levels of severity, whereas the EQ-VAS provides a broader coverage of subjectively important HRQoL dimensions and may be better placed to detect marginally perceived improvements at the individual level (260, 261).

Furthermore, the relatively lower responsiveness of the EQ-5D-3L might be as a result of negatively skewed distributions and average ceiling effects >30 % across all time-points; implying that these responses indicate 'no problems' in all domains resulting in the best health state with a maximum utility or EQ-5D-3L index score of one. Ceiling effects are considered acceptable below 15 % of responses, but EQ-5D-3L descriptive system have shown to suffer from large ceiling effects (>15%) across several diseases areas, thus leading to poor responsiveness (262).

The high ceiling effect observed in this study indicate that the EQ-5D-3L descriptive system is not sensitive enough to discriminate between health states in patients with varying disease severity over time; there are options that are too extreme for participants to respond. For example, individuals in this study found it hard to choose the higher options for 'self-care' item: from "no problems with self-care" to reporting "I have some problems in washing or dressing myself" or even an extreme option of "I am unable to wash or dress myself".

On using the ASRM scale as a clinical anchor, it was noticed that the change in the EQ-5D-3L index and the EQ-VAS scores for patients reporting improvement of mania symptoms was significant and moderate in magnitude (SRMs) on two pair-wise comparisons. In individuals who reported worsening of mania symptoms, the SRMs for the EQ-5D-3L index and the EQ-VAS scores were significant and moderate in size only at one occasion. However, all these changes were in the opposite direction, implying that people who were in mania reported higher EQ-5D-3L index and EQ-VAS scores and vice-versa. An appropriate measure of responsiveness requires that in

participants who report an anchor-defined (i.e. ASRM) increase (or decrease) in case severity, the parallel change in the EQ-5D-3L and the EQ-VAS scores should be negative (or positive) (263). For example, individuals transitioning from mania to 'no mania' should elicit positive increments in their EQ-5D-3L scores.

These findings suggest that in patients with BD, the EQ-5D-3L and the EQ-VAS may be differentially responsive to depression and mania with clinically moderate responsiveness in individuals who reported change in depressive symptomatology. The unexpected direction of the EQ-5D-3L and the EQVAS responses in mania could be due to the poor sensitivity of the instrument to differentiate different levels of mania; the EQ-5D-3L includes questions that captures depression, but does not include any dimension that measures mania. Even though individuals with manic symptoms reported similar or higher EQ-5D-3L scores than those without mania, the index scores were still impaired and lower than that observed in the general population; a finding that was reported previously by few other studies (119).

Overall, this indicates that the EQ-5D-3L and the EQ-VAS scores are responsive to grades of clinical deterioration or improvement in depression, with the EQ-VAS outperforming the EQ-5D-3L. However, both instruments have problems in detecting BD patients with improvement or deterioration in manic symptoms.

6.7.5 Feasibility of ‘mapping’ disease-specific instruments used in Bipolar disorder onto EQ-5D-3L

When applying EFA to explore conceptual similarity among the three instruments, it indicated that these instruments were multidimensional in nature covering three broad areas related to BD (i.e., depression, HRQoL, and mania). Importantly, the “anxiety/depression” item of the EQ-5D-3L loaded exclusively onto to the depression construct and the “usual activities” item had a higher loading under the depression construct than the HRQoL construct. The mania construct only had items from ASRM and did not have any significant loadings from items of other scales. Although CFA provided support for the three-factor model over competing unidimensional and two-factor solutions, the data fitted best into a bi-factor model. The bi-factor model with three specific factors and a global psychological distress factor confirmed that as there is no conceptual equivalence between the observed items, thus mapping is not possible using EQ-5D-3L and disease-specific instruments measuring depression and mania.

6.7.6 Evidence from other studies that explored appropriateness of the EQ-5D-3L in BD

The study results presented here are in line with previous work on the appropriateness of EQ-5D-3L measuring HRQL in BD patients (please refer to **Appendix 10**) With regards to Known-group validity, two of the studies found lower EQ-5D-3L scores in individuals with BD than the general population (217, 264). Hayhurst et al (2006) analysed baseline data of an 18-month randomised controlled trial study in the UK

including patients with Bipolar disorder. They included disease-specific measures completed in the same week of EQ-5D-3L completion. They found significant differences in EQ-5D-3L scores across varying severity of depression as measured by Beck Depression Index (levels of concomitant mania were excluded), with scores in euthymic subgroup (n=76): 1.00, residual Symptoms subgroup (n=55):0.85, sub-syndromal depressed group (n=40): 0.81 and depressed subgroup (n=33) 0.41. They could not find any differences between different states of mania (265).

Similarly, Vojta et al (2001) found differences in EQ-5D-3L EQ-VAS scores among Internal State Scale (ISS) defined mood states; EQ-5D-3L EQ-VAS scores showed a rank order with scores similar for euthymic and manic/hypomanic states (EQ-VAS_{euthymia}=75.9 Vs EQ-VAS_{mania}=68.4, P=NS) but significantly lower for mixed and depressed states (EQ-VAS_{mixed}=46.9 Vs EQ-VAS_{depressed}=57, P=NS) (87). In contrast, Saarni and colleagues (2010), did not find the EQ-5D-3L capable of discriminating between different mood states (266). With regards to convergent validity, all the studies reported moderate to large correlation between depression-specific scales and EQ5D-3L, but small or no association with mania. None of the studies reported on reliability and responsiveness.

6.7.7 Strengths and limitations:

Assessing validity, and appropriateness of instruments used to measure health outcomes and quality of life is a complex task. In HRQoL research, there is no single analysis or test that could unambiguously prove the validity of instruments. Instead,

the process of validation consists of incremental accumulation of evidence that the instruments are sensitive and that they measure what they are expected to measure (113). The results of this study add to the limited body of evidence exploring the appropriateness of EQ-5D-3L with respect to a particular disease.

This study was the second to assess the appropriateness of EQ-5D-3L on a group of patients with a DSM-IV diagnosis of BD using similar criteria proposed by Brazier et al including reliability, known-group validity, convergent validity and responsiveness (237). Additionally, EFA and CFA were used in this paper to explore the feasibility of ‘mapping’ disease-specific instruments in BD onto the generic PBM EQ-5D-3L in line with guidance from the MAPS (MApping onto Preference-based measures reporting Standards) statement (250).

Nonetheless, there were also limitations to this study that need to be considered. The data came from a randomised controlled trial of participants who were engaged with their secondary care providers on an outpatient basis, this could limit the generalisability of the validation results to primary care patients or individuals in the community who may have a wider spectrum of disease symptoms. For example, The QIDS-SR16 and ASRM scores were low indicating that the included patients were relatively stable with few manic or depressive symptoms (**Table 6.1**). The relationship between EQ-5D-3L and mania symptoms could have been different if more people were included with increased severity of mania/hypomania. The study sample was relatively small (n=119), but with 13 repeated measurements over time, the total sample size of all the three measurements evaluated in this study increased to 1257

out of a possible 1547 responses. Although the overall missingness was 18.7 %, the use of repeated measures mixed modelling using maximum likelihood ensured an optimal way of handling missing data (267).

Due to the lack of gold standard HRQoL measure for BD patients, for estimating test-retest reliability of the EQ-5D-3L, and the EQ-VAS, the stable group was classified based on clinical criteria employed in the main trial (236). Those individuals who remained in stable health state were further classified into a) individuals under remission (QIDS-SR16 total score >10 and ASRM ≥ 6 at both time points) and b) individuals in active BD state was based on patients having a QIDS-SR16 total score >10 or ASRM ≥ 6 at both time points. Thus, patients with the active disease could be due to changes either in mania or depression. This clinical anchor of change may have wrongly classified some individuals with divergent changes in either mania or depression into the active disease sub-group and impacted the results. However, mania/hypomania and depression phases are hallmarks of BD disease, and it is challenging to tease out the HRQoL estimates of these stages separately or pool them together (119).

Further, the choice of cut-off points for clinical anchors based on ± 0.5 standard deviations (corresponding to the minimal clinically meaningful difference for patients) used in responsiveness estimation was pragmatically motivated by the need to define valid change categories with adequate numbers for analysis. The use of different values may have led to different results. Given the small sample size and small number of individuals who experience a change in their mental health symptoms, it is not

uncommon, to find studies exploring the responsiveness of EQ-5D-3L in mental illness by including all individuals who complete the instruments at both time-points without using a clinical anchor (30, 233).

The main patient reported outcome data was captured primarily through SMS. Although collection of EQ-5D-3L using SMS technology is not an approved method for collecting EQ-5D data, studies have demonstrated that electronic measures are valid for data collection in trials and routine health care settings(268). Further, research has also indicated the importance of clinical outcomes data collection in mental health including management of depression and bipolar disorder (35, 269).

6.8 Conclusion

In conclusion, this chapter found that there is some support that the EQ-5D-3L is reliable, valid and responsive in individuals with primarily depressive symptoms of BD. However, there is no evidence to support its validity to discriminate between different levels of mania with poor responsiveness in individual who reported change in manic symptoms. Given the mixed results, more evidence is required to use EQ-5D-3L for conducting economic evaluation confidently across the BD spectrum.

Nevertheless, in patients primarily experiencing depressive symptomatology, EQ-5D-3L can be utilised along with depression-specific measures to inform individuals, families and their care providers about the effects of treatment on patient's overall functioning, well-being and day-to-day life. The mapping between disease-specific

instruments (QIDS-SR16 and ASRM) used in BD and the EQ-5D-3L is not possible as there is a lack of conceptual similarity among the three scales.

Given the complexity of the Bipolar disorder, its HRQoL implications was not adequately captured by the five dimensions and three levels of the EQ-5D. This is an important issue that needs to be explored further. **However, what implication does these results have on the economic case of BD?**

The results of the construct validity imply that depressive symptoms in patients with BD are more related to the dimensions “anxiety/depression” and “usual activities” of the EQ-5D-3L, and they predominantly explained loss of HRQoL. Further, BD patients with predominantly manic symptoms are less likely to respond appropriately to EQ-5D-3L. This could be challenging for trials or interventions that aim to reduce hospitalisations or improve outcomes in bipolar manic episodes by targeting reductions in acute symptoms. Use of EQ-5D-3L as an outcome measure in such circumstances may fail to detect such changes and could yield erroneous cost-effectiveness results.

The latest guideline unequivocally recommend psychoeducation as the first line of psychological treatment in the long-term management of BD (7, 270). The next chapter draws on the previous chapters to conduct a cost-utility analysis (CUA) of a novel structured psycho-education (five psycho-education sessions with a facilitator) intervention for individuals who use remote mood monitoring in a UK healthcare setting. This within-trial economic evaluation uses data from the OXTEXT-6 sub-study.

As outlined in **Table 6.1**, the participant in the OXTEXT6 trial were predominately euthymic at the baseline. Further, most of the patients had predominantly depressive symptoms. Nevertheless, as BD represents a continuum of both depressive and manic symptoms, the results of any economic evaluation using the EQ-5D-3L in BD should be treated with caution. Health economists and policy makers should take this limitation on board and until the appropriateness of EQ-5D-3L or similar measure is established, outcome measurement should include contextually validated clinical instruments along with a suitable multi-attribute utility scale.

Chapter 7: An economic evaluation of structured psychoeducation program in BD patients using electronic mood monitoring

7.1 Introduction

Chapter 2 provided the clinical background to this thesis and highlighted that bipolar disorder (BD) is a severe mental illness and a significant cause of morbidity. **Chapter 3** and **Chapter 4** provided the methodological background to this thesis. **Chapter 4** reported the results of a literature review which identified a lack of evidence relating to the cost-effectiveness of BD management strategies, in particular psychological interventions. Based on available evidence, this chapter also highlighted the challenges associated with conducting economic evaluations in BD, noting that evidence generated without using utility measures may limit the ability of decision makers to prioritise health interventions within and beyond mental health.

Chapter 5 presented the cost aspects of BD; the results of a cost-of-illness analysis demonstrated that BD places an enormous burden on society. The total societal costs associated with BD is largely driven by costs incurring outside health and social care and are attributable to productivity losses. The chapter also explored relationships between demographic and clinical characteristics and costs associated with individuals with BD. It identified that worsening of self-reported disease symptoms was associated with both higher healthcare costs and productivity losses.

The economic case of BD cannot be made without choosing the appropriate outcome instrument to measure health benefits. **Chapter 6** looked at the outcome aspects of BD. NICE guidelines for economic evaluations recommend the estimation of health benefits based on generic preference-based measures of HRQoL such as EQ-5D-3L(128). This chapter has shown that appropriateness of EQ-5D-3L is questionable across the spectrum of BD symptomatology, especially in individuals presenting with manic symptoms.

The chapters so far have highlighted the economic impact of BD in the UK, and aspects related to the measurement of health outcomes in BD, but there is a critical need for a better understanding of effectiveness and cost-effectiveness of interventions that enables individuals and their families to manage the illness. **Chapter 7** draws on the previous chapters to bring together evidence on the costs and outcomes associated with alternative treatment strategies in BD management. In the UK, the published guidelines for treating BD emphasises offering patients psychological interventions, including psychoeducation for the long-term management of BD (270). Nevertheless, as **Chapter 4** demonstrated, evidence on the cost-effectiveness of different psychological approaches are still lacking (7).

There are different types of psychological interventions (such as psychoeducation, cognitive behavioural therapy, family focused therapy, interpersonal and social rhythm therapy etc.) as presented in **Chapter 2**. However, as **Figure 7.1** illustrates, all of them have overlapping elements and common goals of improving therapeutic alliance,

increasing patient awareness, improving adherence, early identification of warning symptoms, monitoring treatment response and planning relapse prevention (271).

Figure 7.1: Illustration of inter-relationship between psychological treatments in BD



As a major mental health disorder, BD requires multi-disciplinary and long-term care that places a significant burden on constrained NHS resources. It is therefore important to identify robust patient-centred interventions which are not only clinically effective, but also offer good value for money. This chapter furthers the evidence base, by conducting an economic evaluation of a complex intervention that included a structured psychoeducation program for individuals with BD using an electronic mood monitoring platform.

The chapter begins by briefly describing the trial, interventions, and outcomes before explaining the economic evaluation methodology used in this study. **Section 7.5** describes the analytical methods, before **Section 7.6** presents the results. The chapter ends with a discussion of these results.

7.2 The OXTEXT-6 study

The OXTEXT-6, a sub-study of the larger OXTEXT cohort research program as detailed in **Section 1.2** sought to assess the feasibility and subsequent application of appropriate psychoeducation packages in adjunct with electronic monitoring tools to identify a relapse prevention strategy for people suffering from Bipolar disorder. In this context, a pragmatic randomised controlled trial (RCT) was undertaken to explore the clinical and cost effectiveness of Facilitated Integrated Mood Management (FIMM) compared with Manualised Integrated Mood Management (MIMM) after participation in the OXTEXT-6 programme. Given that most often reported distal outcomes such as relapse relatively require longer term follow up, this pragmatic trial was designed to test the hypothesis that the FIMM approach would be more effective and cost-effective compared to the MIMM strategy in reducing proximal outcomes as measured by weekly measures of current self-reported symptoms through a remote mood monitoring system (236).

7.2.1 Trial design

The OXTEXT-6 trial (Trial registration number: ISRCTN 67300357), was a two-group pragmatic randomised controlled trial, where FIMM was compared against MIMM in

the management of patients with bipolar disorder in the UK. Participants were randomly recruited between April 2011 and December 2012 from the larger OXTEXT-1 pragmatic longitudinal cohort study of patients with a confirmed diagnosis of BD who attended psychiatric clinics in the Oxford Health NHS Foundation Trust (OHFT) and participated in a user friendly SMS- or web-based remote mood monitoring system called “True Colours” (TC) (www.truecolours.nhs.uk) (272).

Figure 7.2: Example “prompt” sent by the TC system to study participants

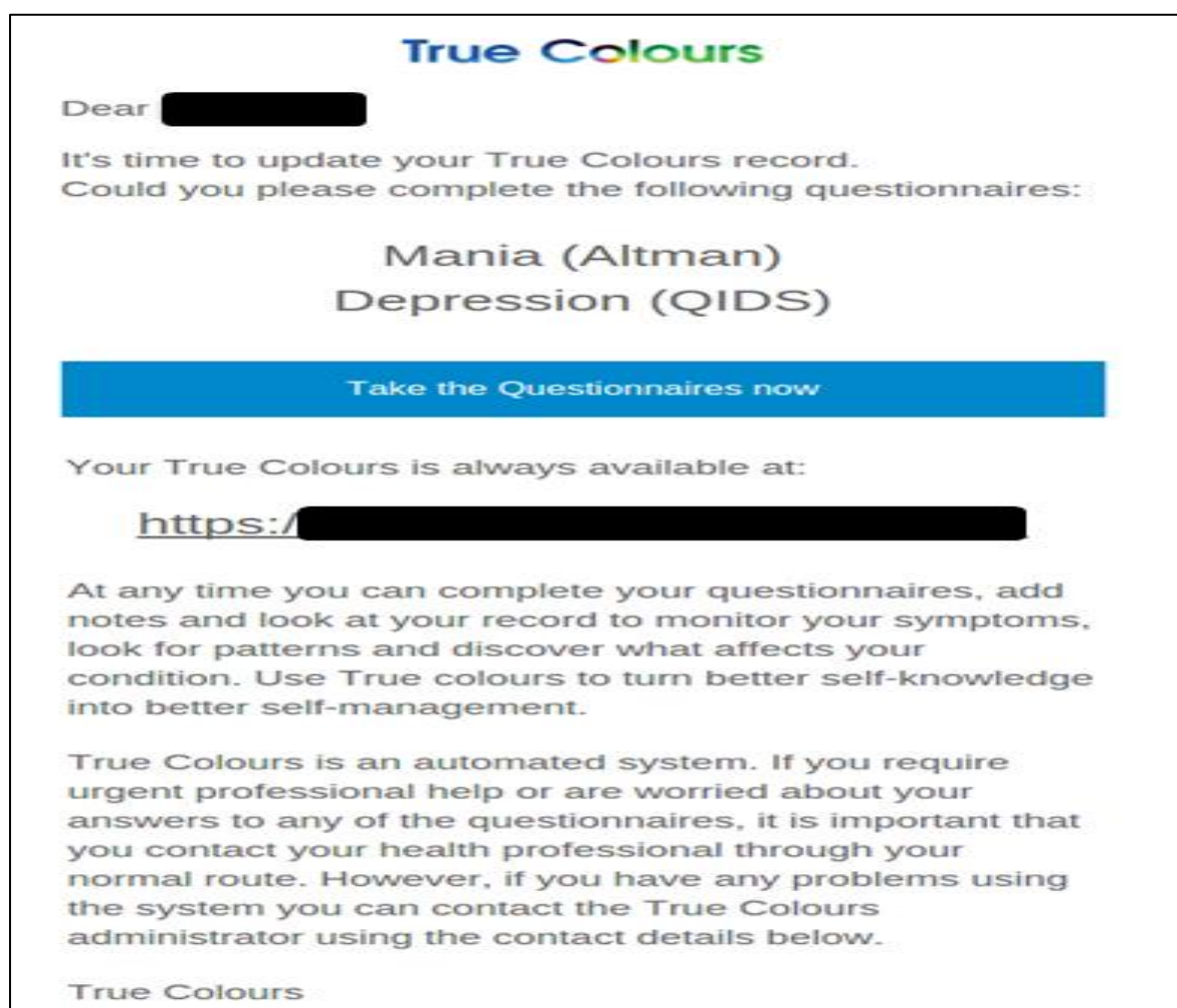


Figure obtained from the main trial article (236)

TC is a system of monitoring similar to paper based self-monitoring tools such as the National Institute of Mental Health Life Chart Methodology (NIMH-LCM) and the Clinical Monitoring Form which are used in clinical practise, but uses commonly available technologies (SMS/‘texting’, E-mail or Internet) to help patients with bipolar disorder to monitor their mood. Patients provide their weekly or more frequent standardised self-report ratings of mood on scheduled prompts (an example of a prompt provided in **Figure 7.2**) from the TC system which can then be automatically charted onto a graph, on a secure online record, enabling patients and their clinicians to oversee mood fluctuations and overall course of illness (35, 69).

Eligible criteria were: a) males and females aged over 16 years old, with a diagnosis of BDI or BDII according to the DSM-IV-TR criteria, but not in a current mood episode, b) current and reliable participation in TC mood monitoring for ≥ 4 weeks (at least out of 4 responses in the 4 weeks prior to invitation to participate), c) willing and able to provide informed consent to participate in the study. Randomisation was stratified by a) illness type including BDI or BDII and b) severity of BD, i.e. whether patients were in recovery or not at time of randomisation. The study received ethical approval by the local Research Ethics Committee (Oxfordshire REC A, Reference: 10/H0604/13).

Further details on the main study design of this two-group pragmatic randomised controlled trial, methods and results have been reported in detail in the main clinical paper (236) This chapter reports an economic evaluation conducted alongside the trial is described next.

7.3 Interventions

The main intervention is called Integrated Mood Management (IMM) which is a complex intervention that draws on established collaborative care models for long-term illnesses including depression (273). Individualised psychoeducation adapted from evidence-based psychosocial treatments for BD forms the core element of IMM (274). The IMM was developed into a manual which could be administered by facilitators (termed as FIMM) or self-administered by patients (termed as MIMM) in combination with weekly mood monitoring through TC. In addition to being on TC, participants allocated to MIMM followed self-administered psychoeducation using a manual and received standard care from their clinicians as per national clinical guidance including pharmacological treatment and referral for other services.

FIMM consisted of standard care from clinicians, and additionally treatment facilitators were expected to provide individualised psychoeducation involving five face-to-face sessions (lasting around 50-minute per individual session) over a period of 12 weeks. Each session covered different chapters of the manual: (1) identifying the relapse signature; (2) reviewing risk and protective factors; (3) daily rhythm and sleep/wake regulation; (4) the role of medications and substance/alcohol abuse; and (5) finalising the mood management plan. Treatment facilitators were either clinical psychologists or researchers working with the target population; they were trained specifically to deliver the intervention within the integrated mood management framework, and received regular support from a specialist in psychosocial interventions for BD and a consultant psychiatrist.

7.4 Outcome measures

7.4.1 Clinical Outcomes

The primary clinical outcome was group differences in self-reported depressive symptoms using the Quick Inventory of Depressive Symptoms–Self Report (QIDS-SR16) scale over a period of 12 months. It consists of 16 items, assessing the nine DSM-IV criteria symptom domains. These items are rated on a four-point Likert scale, ranging between 0 and 3, with an overall score ranging from 0-27; higher scores indicating increased severity of depression (180). QIDS-SR16 is a contextually validated patient-reported rating scale for depression used routinely for monitoring disease status in BD (275, 276). Its psychometric properties have been established in the UK primary care and it may also reliably measure depressive symptoms in individuals with psychosis (243, 277).

Secondary outcomes included mood elevation measured by the Altman Self-Rating Mania (ASRM) scale which consists of five items. The five items include mood, self-confidence, sleep disturbance, speech, and activity level over the past week; items are scored from 0-4, with higher scores indicating increasing severity of mania. ASRM is one of the most widely used self-reported scales measuring symptoms associated with mania and has demonstrated good psychometric properties in individuals with BD (275, 278). In the literature, a cut-off point of less than or equal to 5 is indicative of 'being not in a manic state', whereas a score of greater than 5 indicates a 'state of mania' (179).

Participants provided weekly QIDS-SR16 and ASRM responses by answering text or email prompts at www.truecolours.nhs.uk. Symptom scores were extracted from the TC system for the week in which participants were randomised (W0), and for the subsequent 52 weeks (i.e. W1 to W52). These clinical outcomes were also used as one of the stratification variables for randomised allocation (BD severity: high or low; defined as four consecutive weeks with minimal or no mood symptoms – QIDS-SR16 < 10 and ASRM < 5). Other outcomes included the self-reported measures of knowledge and behaviour relevant to the self-management of BD measured by specifically designed questionnaire the ‘Oxford Bipolar Knowledge Questionnaire’ (OBQ) (236). The OBQ consists of 10 items each scored on a scale between 0-2, higher scores indicating increased knowledge of BD mood management. OBQ was collected at baseline and at 3 months and 12 months post randomisation.

7.4.2 Economic outcomes

Health-related quality of life was measured with the European Quality of life (EuroQol) instrument EQ-5D-3L. The EQ-5D-3L is a five-item questionnaire representing the dimensions of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Items are marked on a three-point Likert scale: 1=no problem; 2=some problems; and 3=extreme problems (115).

The Investigating Choice Experiments Capability Measure for Adults (ICECAP-A) and Oxford CAPabilities questionnaire-Mental Health (OxCAP-MH), both multi-dimensional instruments that draw on Sen’s capability approach were used to assess

a broader measure of wellbeing known as capabilities (279). The ICECAP-A is a self-reported questionnaire consisting of 5 items with overall scores ranging from 0, which reflects an absence of capability (or no capability) to 1, full capability(280).

The OxCAP-MH, a 16-item index is specifically tailored for mental health research. Items are scored on a 1 to 5 Likert scale with higher scores indicating better capabilities (281). Standardised overall scores range from 0-100, 0 reflecting no capabilities and 100 reflecting full capabilities. Participants completed the EQ-5D-3L and ICECAP-A questionnaires at baseline, 3, 6, 9 and 12 months, while OxCAP-MH was completed at baseline and 12 months post-randomisation.

7.5 Economic evaluation

A cost-utility analysis was carried out for the year 2012/2013 from a societal perspective. This included costs due to healthcare (NHS) and social care, personal costs, productivity losses and informal care of patients. NHS costs included the costs for delivering FIMM, primary care contacts, emergency visits, outpatient and inpatient visits, and medications. Personal costs comprised of travel costs to facilities, and costs associated with any private treatment. Information was also collected from each study participant on the time informal carers (e.g. family members and friends) provided unpaid care and support. **Table 7.3** summarises the methods for the economic evaluation used in the base case analysis (the primary analysis).

Informal care costs were estimated using the opportunity cost approach by attaching the average cost derived from the average annual gross UK wage rate for 2012/2013 (average hourly salary) with the number of hours of unpaid work (186, 187). Productivity losses including sick leave was calculated using the human capital approach; the number of sick days was multiplied by the average daily cost of sickness absence in the UK (184, 185).

Table 7.3: Summary of the features employed in the economic evaluation

Aspect of methodology	Strategy used in the base case economic evaluation
Form of economic evaluation	Cost-utility analysis for comparison between FIMM and MIMM
Perspective	Societal
Time horizon	A within-trial analysis
Population	All randomised patients were included
Costs included in analysis	• Primary care • Emergency visits • Inpatient and outpatient visits • Medications • Social care • Private sector costs • Unpaid informal care • Productivity costs • Intervention costs
Utility measurement (primary economic outcome)	EQ-5D-3L (administered at baseline and at three, six, nine and twelve months follow-ups)
Quality-adjusted life-year (QALYs)	QALYs were calculated using the 'area-under-the-curve' approach that assume that patients' utility changes linearly between utility measurements
Adjustment for baseline utility	Regression used to adjust quality-adjusted life-year calculations for differences in baseline utility
Missing data	Multiple imputation

Resource use data over the trial period of 12 months was collected prospectively using self-reported questionnaires (a sample provided in **Appendix 11**) administered at 3, 6, 9 and 12 months follow-ups (covering the prior three-month time periods respectively). Intervention costs of the psychoeducation interventions included the costs of delivering FIMM to service-users to deliver the intervention (time of staff and service users), administrative costs (time of staff required to organise event, follow-up telephone calls) and self-study time for service users in both arms. The FIMM resource use information was primarily gathered from facilitator's diaries and self-reported data (self-study time) as recorded by service users.

Participant-level annual mean costs were calculated by multiplying self-reported resource use with service specific unit costs in British pounds sterling (£). Unit costs were obtained from various recognised sources including the NHS schedule of reference costs (2012-13) for secondary care (282), annual compendium of Costs of Health and Social Care (2013) for primary and social care costs(283), and medicine costs from the British National Formulary (284).

The incremental cost per additional quality-adjusted life-year at 12-month follow-up was the primary economic outcome. Health state values obtained from the EQ-5D-3L at each time-point were converted to utility scores using the standard UK EQ-5D-3L tariff. Quality-adjusted life years (QALYs) were calculated using the 'area-under-the-curve' approach. This method takes into account utility scores at each time point and the duration between them while assuming linear trends between time points (109).

7.6 Statistical methods

The primary economic analysis estimated the incremental cost-effectiveness ratio (ICER), i.e. the difference in mean costs divided by the difference in mean QALYs, between FIMM and MIMM over 12 months.

$$\text{ICER} = (\text{CostFIMM} - \text{CostMIMM}) / (\text{QALYFIMM} - \text{QALYMIMM})$$

The incremental analyses were undertaken using a regression framework. Cost differences between interventions are presented unadjusted and adjusted for pre-specified baseline variables including type of BD, BD severity, gender and age. The models for QALY data are presented unadjusted and adjusted for baseline utility and aforementioned baseline covariates (285). Discounting was not undertaken as the time horizon of the study was only one year.

Uncertainty around the ICERs were estimated using a non-parametric bootstrap with replacement technique to generate a joint distribution of incremental costs and incremental QALYs. 1000 bootstrap replications of the differences in costs and QALYs were generated to plot on the cost-effectiveness plane. Cost-effectiveness results were also summarised within a net-benefit approach by estimating net monetary benefit for each treatment arm and generating cost-effectiveness acceptability curves by using patient specific net monetary benefits for a range of willingness to pay threshold values (ceiling ratios) for an unit improvement in outcome (109).

For the base-case, analysis was conducted as a “complete case analysis” where only patients with available utility scores at all points and available costs were included. A proper application of the intention-to-treat analysis is possible only when data is complete. As a result of missingness observed in trial data, statistical analyses using a ‘true’ intention-to-treat principle (ITT) was not feasible (286, 287). However, in line with recommended guidelines, the intention-to-treat analysis was undertaken after multiple imputation of missing data (further details in **Sections 7.6.2**) (166, 288).

7.6.1 Missing data

There was varying degree of missingness in both the cost and the outcome EQ-5D-3L data due to some participants dropping out, failing to return questionnaires or provide complete questionnaires. Complete 12 months EQ-5D-3L data were available for 66 (55%) patients with data being complete for all participants at the baseline, 81% (n=98) at three months, 83% (n=101) at six months, 74 % (n=90) at nine months and 68% (n=82) at 12 months. Complete cost data at all time-points were available for 69 (57%) participants, with availability at three months being at 81% (n=98), 81% (n=98) at six months, 76% (n=92) at nine months and 65% (78/121) at 12 months. The joint completeness of total costs and total QALY data over the 12-month time-point was 55% (n=66).

7.6.2 Sensitivity analyses

Several assumptions were tested to address uncertainty in the results of the presented economic evaluation.

- a) To consider the effect of missingness in data, multiple imputation using chained equations was used to impute missing cost and outcome data. As both costs and outcomes in the current dataset were skewed, predictive mean matching was used to carry out multiple imputation to generate 100 imputed datasets by assuming that data were missing at random (MAR). This assumption means that, given the observed data, the reason for the missingness is dependent only on observed and does not depend on the unobserved data. The multiple imputation models were run separately for each treatment arm with common covariates being age, gender, type of BD, BD severity and baseline EQ-5D-3L index score (289). In addition, costs and QALYs at previous time points were also included to impute missing costs and QALYs respectively. As in complete case analyses, intention-to-treat (ITT) principle was followed to derive ICERs from the multiple imputed data set. Each of the 100 imputed datasets was analysed separately and combined using Rubin's rules (290).
- b) The impact of potential baseline differences in outcomes was adjusted by using regression analysis (291). Parsimonious adjusted models were run both on complete case analyses and imputed data after controlling for treatment arm, baseline EQ-5D-3L score, age, gender and stratification variables used in the randomisation (type of BD, BD severity).

- c) The impact of extreme outliers of self-reported resource was explored by excluding these and re-running base case analyses. A generalised box plot, that took into account for the positively skewed distribution of cost data in the dataset, was used to identify outliers (292).
- d) Informal care, productivity losses, and personal costs including travel borne by participants were excluded to provide a NHS and social care (third party payer) perspective.

In total 15 different scenarios were explored including exclusion of outliers and multiple imputation, multiple imputation of health and social care costs etc. The results are provided in **Appendix 12**. All analyses were undertaken in STATA, version 14 (293).

7.7 Results

A total of 121 participants were recruited to the OXTEXT-6 trial, with 60 participants in the FIMM arm and 61 participants in the MIMM arm. Disease status at baseline as measured by the main clinical outcomes QIDS-SR16/ASRM scores and the economic outcome EQ-5D-3L utility values did not differ statistically significantly between groups. Furthermore, there were no significant differences in baseline characteristics between the two treatment groups either (**Table 7.2**). Baseline characteristics were also compared with the larger OXTEXT-1 cohort and were found similar to the trial arms.

Table 7.2: Demographic and clinical characteristics of participants

Characteristics of participants	FIMM (60)	MIMM (61)	FIMM versus MIMM Difference (P value)	OXTEXT1-cohort (n=271)
Women (n, %)	42 (70%)	46 (75.4%)	0.05 (0.50)	184 (68%)
White-British (n, %)	54 (90%)	58 (95.1%)		256 (94%)
Type of BD (n, %)				
BDI	39 (65.0%)	40 (65.6%)	0.005 (0.95)	171 (63.1%)
BDII	21 (35.0%)	21 (34.4%)		100 (36.9%)
Employment status, n (%)*				
Employed/self-employed	26 (43.3%)	25 (40.9%)	2.70 (0.28)	136 (50.2)
In education	20 (16.7%)	5 (8.2%)		33 (12.3)
Unemployed/retired	24 (40%)	31 (50.1%)		102 (37.5)
Age at assessment (mean, SE)	45.4 (1.78)	43.2 (1.72)	2.17 (0.38)	40.4 (0.8)
Baseline clinical outcomes (mean, SE)				
Baseline Altman score	3.00 (0.51)	3.36 (0.51)	-0.356 (0.62)	3.5 (0.31)
Baseline QIDS-SR16 score	7.61 (0.72)	7.23 (0.70)	0.377 (0.70)	9.9 (0.40)
Baseline Health economic outcomes (mean, SE)				
EQ-5D index	0.87 (0.02)	0.86 (0.02)	0.006 (0.84)	NA
OXCAPMH index	59.5 (0.88)	59.7 (0.83)	0.172 (0.89)	NA
ICECAP	0.75 (0.02)	0.79 (0.02)	0.040 (0.27)	NA

*The difference in trial arms tested by Pearson chi2 test statistic; FIMM: Facilitated Integrated Mood Management; MIMM: Manualised Integrated Mood Management

7.7.1 Resource use and costs

Table 7.3 presents the average per patient resource use over the 12 months follow-up based on the complete trial data for each resource use category. During the trial, majority of participants used primary care and community mental health care services. Around 70% of participants availed sick leave due to their illness or sought informal care from their family or friends.

Table 7.3: Resource use by participants during the 12 months follow-up period (Complete case analyses)

Resource use	FIMM (n=36)	MIMM (n=33)	% using service
	Mean (SE)	Mean (SE)	
General practitioner contacts	7.6 (1.5)	7.9 (1.1)	97%
General practise nurse contacts	1.9 (0.5)	2.3 (0.5)	65%
Community nurse contacts	0.6 (0.3)	0.2 (0.1)	15%
Community psychiatric nurse contacts	2.7 (1.1)	3.7 (1.5)	28%
Community drop-in centre visits	0.5 (0.3)	0.3 (0.2)	12%
Day centre visits	3.2 (2.0)	0.6 (0.5)	13%
Support group contacts	4.9 (3.8)	2.2 (0.9)	25%
Ambulance call-outs	0.5 (0.3)	0.2 (0.1)	13%
Accident & emergency visits	0.5 (0.2)	0.2 (0.2)	16%
Non- Mental health OP visits	2.9 (0.6)	1.3 (0.4)	51%
Psychiatrist OP visits	5.2 (0.7)	4.5 (0.7)	86%
Psychologist OP visits	2.4 (0.5)	3 (0.7)	61%
Non- Mental health day care (days)	0.9 (0.6)	0.2 (0.1)	15%
Non-MH Hospitalization (days)	0.3 (0.2)	0.1 (0.1)	10%
Mental health day care (days)	0.3 (0.2)	0.2 (0.2)	6%
Mental health inpatient (days)	0.4 (0.3)	3.7 (3.2)	10%

PICU (days)	0.3 (0.3)	0 (0)	1.5%
Social care contacts	2.5 (0.9)	5.6 (2.2)	35%
Other private contacts	7.7 (2.2)	4.9 (2.2)	50%
Travel (miles)	45.4 (10.7)	31.5 (9.1)	99%
Sick leave (days)	39.7 (12.1)	29.6 (10.3)	67%
Informal care (hours)	499 (124)	425 (172)	71%

Table 7.4 presents the average per patient intervention-specific resource use over the 12 weeks intervention period based on the available case data for each resource use category.

Table 7.4: Resource use related to the psychoeducation interventions (FIMM and MIMM)

Intervention for the FIMM arm (n=59)	Number	Mean (SE)	[95% Conf. Interval]
Number of sessions	59	4.4 (1.4)	0 - 5
Average administration time (mins)	59	217 (23)	171 - 262
Average session time (mins)	59	253 (11)	231 - 275
Total FIMM time (mins)	59	469 (28)	413 - 526
Travel for participants (miles)	59	43 (9)	25 - 60
Travel time for facilitators (mins)	59	79 (19)	41 - 117
Self-study time over 12 weeks (hrs)*	56	4.2 (0.41)	3.4 - 5
Follow up calls (mins)	59	76 (7)	62 - 90
Intervention for the MIMM arm			
Self-study time over 12 weeks (hrs)*	49	7.5 (1.5)	4.5-10.5

*This Resource use was reported by participants and the summary is based on available data over the intervention period (12 weeks). The provided summary is for descriptive purposes and the final base-case analyses was conducted on a complete-case basis.

On an average, 82% of participants in the FIMM arm attended all their five sessions. The average total time for delivering five FIMM sessions to each participant was 469 minutes (7.8 hrs); it included the time for actually running the sessions (253 minutes/4.2 hrs) and administration time (217 min/3.6 hrs) required for paperwork or preparation. The sessions were followed up by phone calls for addressing any further enquiries or concerns. The self-study time in the MIMM arm was 7.5 hours in comparison to 4.2 hours in the FIMM arm over the intervention period (12 weeks).

The unit costs used in this chapter along with their sources of information are provided in **Appendix 12**. The estimated mean cost per participant for the delivery of the FIMM intervention was £772 (95% CI: £677-£866). This cost estimate included the travel costs for both participants and facilitators. The costs for FIMM were primarily based on FIMM diaries maintained by 9 facilitators who provided sessions to 60 FIMM participants. A home-work or self-study time based on self-report was also part of the intervention. As expected, FIMM intervention costs was statistically significantly larger than the intervention costs for MIMM (difference: 642; P value=<0005) (see **Table 7.5**).

Table 7.5 details the costs for all resource use categories by treatment group over the 12 months follow-up. Overall, most of the costs were higher in the FIMM arm than MIMM. However, a large difference was noticed in the mental health hospitalisation costs, with costs higher in MIMM than FIMM, with an adjusted mean difference of £1061 (P=0.312). This was largely due to one individual in the MIMM group, who reported length of stay of 106 days. On removing this extreme case, the adjusted cost difference reduced to £183 (P=0.576)

Table 7.5: Estimated mean costs per patient per year (£, 2012-13 values) based on complete data

Type of costs	FIMM, n (33)	MIMM, n (33)	FIMM versus MIMM	Adjusted FIMM versus MIMM*
General practitioner	320 (202 - 438)	332 (244 - 421)	-13(-87-125; 0.77)	-8 (-89 - 123; 0.81)
General practise nurse	25 (12 - 37)	31 (18 - 44)	-6 (-13 - 20; 0.68)	-6 (-13 - 20; 0.72)
Community (comm) nurse	10 (-3 - 24)	3 (0 - 6)	7 (-8 - 26; 0.21)	7 (-6 - 29; 0.21)
Comm psychiatric nurse	104 (17 - 192)	146 (20 - 272)	-42 (-61-171; 0.90)	-44 (-62 - 193; 0.90)
Comm drop-in centre	18 (0 - 36)	9 (-4 - 23)	9 (-17 - 70; 0.23)	10 (-15 - 72; 0.20)
Day centre	111 (-20 - 241)	20 (-16 - 55)	91 (-30 - 264; 0.12)	95 (-26 - 268; 0.10)
Support group	168 (-93 - 429)	75 (14 - 136)	93 (-398 - 155; 0.38)	95 (-403 - 154; 0.38)
Ambulance	48 (-8 - 104)	17 (-7 - 42)	31 (-25 - 104; 0.22)	32 (-24 - 105; 0.21)
Accident & emergency	58 (2 - 113)	28 (-7 - 63)	30 (-36 - 95; 0.25)	32 (-34 - 97; 0.22)
Non-MH OP	277 (154 - 400)	145 (43 - 256)	132 (-27 - 274; 0.10)	140 (-17 - 297; 0.08)
Psychologist OP	126 (66 - 186)	169 (86 - 252)	-43 (-56 - 141; 0.39)	-38 (-59 - 135; 0.44)
Psychiatrist OP	315 (231 - 400)	274 (180 - 369)	41 (-165 - 83; 0.52)	50 (-168 - 68; 0.39)
Non-Mental health day care	512 (-35 - 1059)	100 (-53 - 253)	412 (-168 - 992; 0.16)	405(-180 - 990; 0.17)

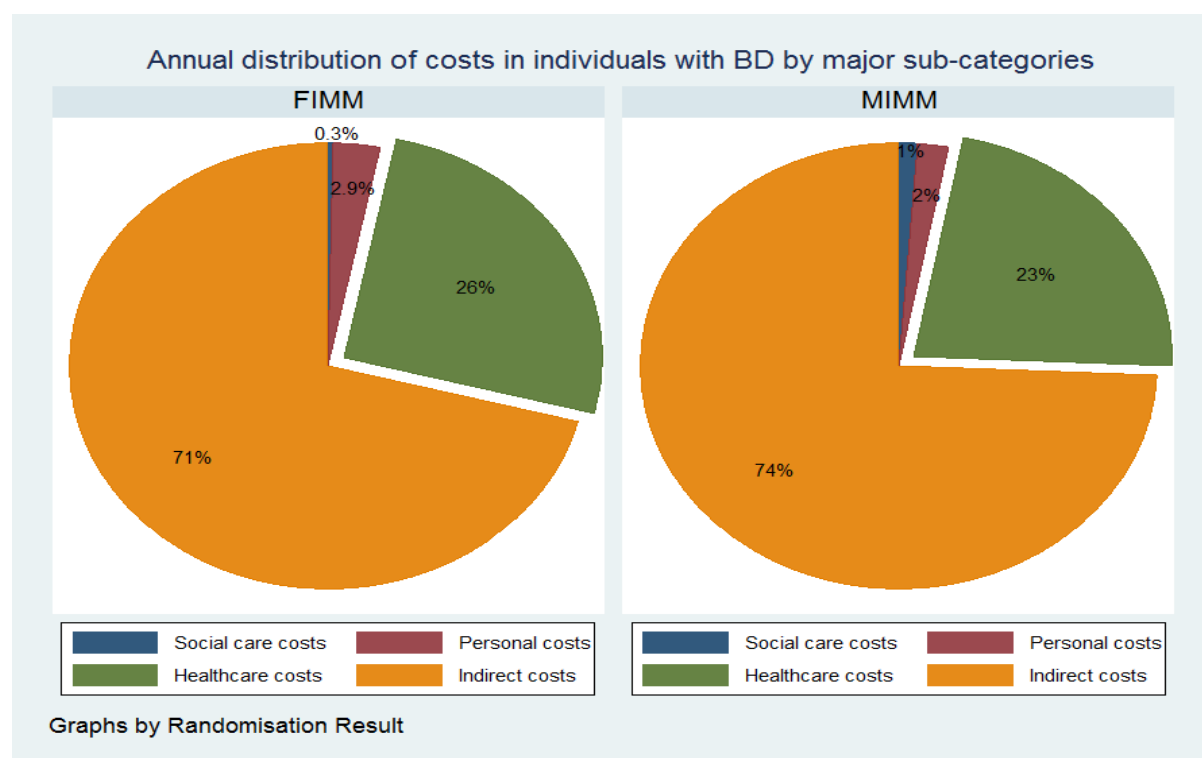
Non-Mental Health Hospitalisation	318 (-87 - 723)	227 (-136 - 589)	91 (-798 - 373; 0.74)	80 (-797 - 381; 0.77)	
Mental health day care	40 (-13 - 92)	20 (-20 - 60)	20 (-92 - 40; 0.44)	22 (-93 - 38; 0.41)	
Mental health inpatient	153 (-39 - 346)	1265 (-1000 - 3530)	-1112 (-3243 - 1020; 0.29)	-1061 (-3201 - 1079; 0.31)	
PICU	173 (-178 - 523)	0 (0 - 0)	173 (-187 - 532; 0.34)	176 (-186 - 538; 0.33)	
Drug costs	282 (111 - 453)	95 (52 - 139)	187 (6 - 367; 0.04)	190 (8 - 372; 0.04)	
Social care	45 (0 - 90)	154 (-20 - 327)	-109 (-61 - 272; 0.21)	-106 (-65 - 276; 0.27)	
Other private	425 (130 - 720)	217 (28 - 405)	208 (-143 - 559; 0.21)	210 (-129 - 570; 0.21)	
Travel	62 (7 - 116)	46 (6 - 86)	16 (-10 - 29; 0.32)	18 (-30 - 9; 0.30)	
Sick leave	3689 (1405 - 5973)	2753 (796 - 4711)	936 (-3775 - 2454; 0.67)	1148 (-3,813 - 2,125; 0.57)	
Informal care	7960 (3940 - 11980)	6780 (1191 - 12368)	1180 (-5492 - 7852; 0.72)	1208 (-5525 - 7941; 0.72)	*
Intervention costs	772 (677 - 866)	130 (65 - 194)	642 (528 - 756 ;< 0005)	641 (526 - 756 ;< 0005)	
Healthcare costs	3819 (2599 - 5039)	3087 (655 - 5518)	732 (-1869 - 3334; 0.58)	808 (-1795 - 3411; 0.537)	
Societal costs	15999	13036	2963 (-6138 - 12064; 0.51)	3248 (-12544 - 6047; 0.488)	

*adjusted for baseline EQ-5D-3L utility

In the medicines category, broadly, there were six classes of drugs that were prescribed over a period of one year including anticonvulsants, antipsychotic, antidepressants, lithium, anxiolytics and hypnotics. At the 12th month time point, large, but statistically insignificant differences were noticed in prescription uptake for anticonvulsants (FIMM: 47.7% Vs MIMM: 31.6%, diff: 16.1%, $P = 0.137$), antidepressants (FIMM: 36.3% Vs MIMM: 26.3%, diff: 10.0%, $P \text{ value} = 0.329$), and lithium (FIMM: 40.1% Vs MIMM: 26.3%, diff: 14.6%, $P \text{ value} = 0.165$). There was no difference in the average number of medicines used at the baseline (2.03 in FIMM Vs 1.70 in MIMM, diff: 0.33, $P = 0.085$). After adjusting for the pre-treatment mean number of medicines as a covariate and treatment as an independent variable, the mean number of medicines decreased in both arms (1.77 in FIMM Vs 1.63 in MIMM, diff: 0.14, $P = 0.190$) at the 12-month time-point. Although the mean number of drugs consumed did not differ in both arms, the annual drug consumption costs were statistically higher in FIMM versus MIMM, with an adjusted mean difference of £190 ($P = 0.04$).

In both groups, informal care and productivity losses, represented the largest burden on individuals with BD (**Figure 7.3**). Costs due to sickness absence was higher in FIMM with adjusted cost-difference of £1148 ($P = 0.57$). Informal care costs were also higher in FIMM than MIMM, with adjusted cost difference of £1208 ($P = 0.72$). The overall pattern of resource use was slightly higher in FIMM than MIMM, with the estimated annual mean costs per participant in the FIMM group at £15999, which was £2963 more than in the MIMM group ($P = 0.509$). This difference was not statistically significant after adjustment for baseline EQ-5D-3L utility values.

Figure 7.3: Distribution of costs in treatment arms



*Personal costs in Figure 7.3 include costs for private treatments and travel costs; indirect costs include productivity losses and informal care costs

7.7.2 Outcomes

Table 7.6 reports data on the main economic outcomes. EQ-5D-3L utility values showed a decline, while self-reported visual analogue scores showed improvements in health status over time in both groups, but the difference was not statistically significant in either cases. After adjusting for baseline EQ-5D-3L- utility values, an insignificant difference of -0.0142 (95% CI -0.04 to 0.07) in the QALYs was noted over 12 months (see Table 7.7). No statistically significant difference was noted in ICECAP-A scores across any time-point. However, OXCAP-MH index showed a statistically significant adjusted difference of -3.8 (95% CI -6.2- -1.4, $P=0.002$), with the MIMM group indicating higher capabilities at the 12 months follow-up.

Table 7.6: Summary statistics and difference in outcomes between interventions at each time-point based on complete case analysis (FIMM, n=33; MIMM, n=33)

Economic outcomes	FIMM (n=33)	MIMM n (33)	Mean difference (CI; p-value)	Adjusted* mean difference (CI; p-value)
EQ-5D-3L index_0	0.87 (0.02)	0.86 (0.02)		
EQ-5D-3L index_3	0.79 (0.04)	0.82 (0.03)	-0.025 (-0.12 - 0.07; 0.59)	-0.026 (-0.11 - 0.06; 0.54)
EQ-5D-3L index_6	0.83 (0.03)	0.81 (0.03)	0.017 (-0.07 - 0.10; 0.69)	0.005 (-0.08 - 0.09; 0.90)
EQ-5D-3L index_9	0.83 (0.03)	0.85 (0.03)	-0.018 (-0.10 - 0.04; 0.66)	-0.018 (-0.09 - 0.06; 0.63)
EQ-5D-3L index_12	0.82 (0.03)	0.84 (0.02)	-0.02 (-0.10 - 0.06; 0.64)	-0.025 (-0.10 - 0.05; 0.53)
ICECAP0	0.75 (0.02)	0.79 (0.02)	-0.040 (-0.11 - 0.02; 0.22)	
ICECAP3	0.74 (0.03)	0.75 (0.03)	-0.011 (-0.09 - 0.07; 0.80)	0.013 (-0.06 - 0.08; 0.72)
ICECAP6	0.77 (0.03)	0.77 (0.03)	-0.001 (-0.08 - 0.07; 0.98)	0.012 (-0.06 - 0.09; 0.74)
ICECAP9	0.76 (0.03)	0.79 (0.03)	-0.023 (-0.10 - 0.05; 0.55)	-0.005 (-0.07 - 0.06; 0.89)
ICECAP12	0.76 (0.03)	0.81 (0.03)	-0.043 (-0.12 - 0.03; 0.27)	-0.035 (-0.10 - 0.03; 0.29)
OXCAPMH index0	59.5 (0.9)	59.7 (0.8)	-0.17 (-2.6 - 2.2; 0.89)	
OXCAPMH index12	58 (1.2)	61.4 (1.1)	-3.39 (-6.6 - 0.15; 0.04)	-3.83 (-6.2 - -1.4; 0.002)

*The difference at 3, 6,9,12 months is adjusted for respective baseline scores

7.7.3 Cost-utility analysis

Base case analysis

Based on the complete case analysis, the FIMM group cost an additional £3028 costs than the MIMM arm and QALY scores were 0.002 points worse over 12 months, indicating that FIMM is strongly dominated by MIMM. Table 7.7 presents data both on unadjusted and adjusted cost/effects differences in aggregated and disaggregated manner. The results of the ITT analysis post multiple imputation is also presented in the same table for comparison.

Figure 7.4 illustrates the scatterplot of 1000 bootstrapped replicates of incremental costs and incremental QALYs on the cost-effectiveness plane. The point estimate is in the north-west quadrant, implying that FIMM on average is costlier and less effective than MIMM. After adjusting for covariates, the likelihood of FIMM being dominated by MIMM increased to 56%. The corresponding cost-effectiveness acceptability curve in Figure 7.5 suggests that at a willingness-to-pay threshold of £20000, there is only a 27% (adjusted probability: 25%) probability of FIMM being cost-effective in comparison to MIMM. Increasing the ceiling ratio (up to £500,000) did not increase the adjusted probability of FIMM being more cost-effective beyond 28%.

Table 7.7: Cost-utility analyses for FIMM versus MIMM

Base case (complete case analysis)	Total costs (£)			Total QALYs over 12 months			Point ICER estimate [quadrant]	Probability that FIMM is:				
Analysis	FIMM	MIMM	Difference, CI, P value	FIMM	MIMM	Difference, CI, P value		More effective	Less costly	Dominant	Dominated	cost-effective at £20 000
A. ITT, no adjustment for covariates	15999 (3234)	13036 (3200)	2963 (-6138 - 12064; 0.510)	0.826 (0.026)	0.828 (0.027)	-0.002 (-0.08 - 0.073; 0.956)	-1481500 Dominated [NW]	45.8%	26.5%	18%	45.4%	27.4%
B. ITT, adjusting for covariates	15996 (3178)	12748 (3319)	3248 (-6047 - 12544; 0.488)	0.821 (0.026)	0.833 (0.027)	-0.013 (-0.07 - 0.04; 0.659)	-249846 Dominated [NW]	30.6%	24%	12%	57.4%	23.2%
Post Multiple Imputation	Total costs (£)			Total QALYs over 12 months			Point ICER estimate [quadrant]	Probability that FIMM is:				
Analysis	FIMM	MIMM	Difference	FIMM	MIMM	Difference		More effective	Less costly	Dominant	Dominated	cost-effective at £20 000
C. ITT, no adjustment for covariates	19550 (3646)	12043 (1959)	7507; (-666 - 15680; 0.071)	0.828 (0.022)	0.835 (0.018)	-0.008 (-0.06- 0.05; 0.783)	Dominated - 938375 [NW]	38%	2%	1.3%	61.3%	2.3%
D. ITT, adjusting for covariates	19956 (2952)	11644 (2788)	8311; (127 - 16495; 0.047)	0.826 (0.017)	0.838 (0.015)	-0.012 (-0.06 - 0.03; 0.596)	Dominated -692583 [NW]	22%	1.9%	0.7%	76.8%	1.9%

Figure 7.4: (a) Cost effectiveness plane of unadjusted bootstrapped net cost and QALY pairs (complete case analysis); (b) Cost effectiveness plane of adjusted bootstrapped net cost and QALY pairs

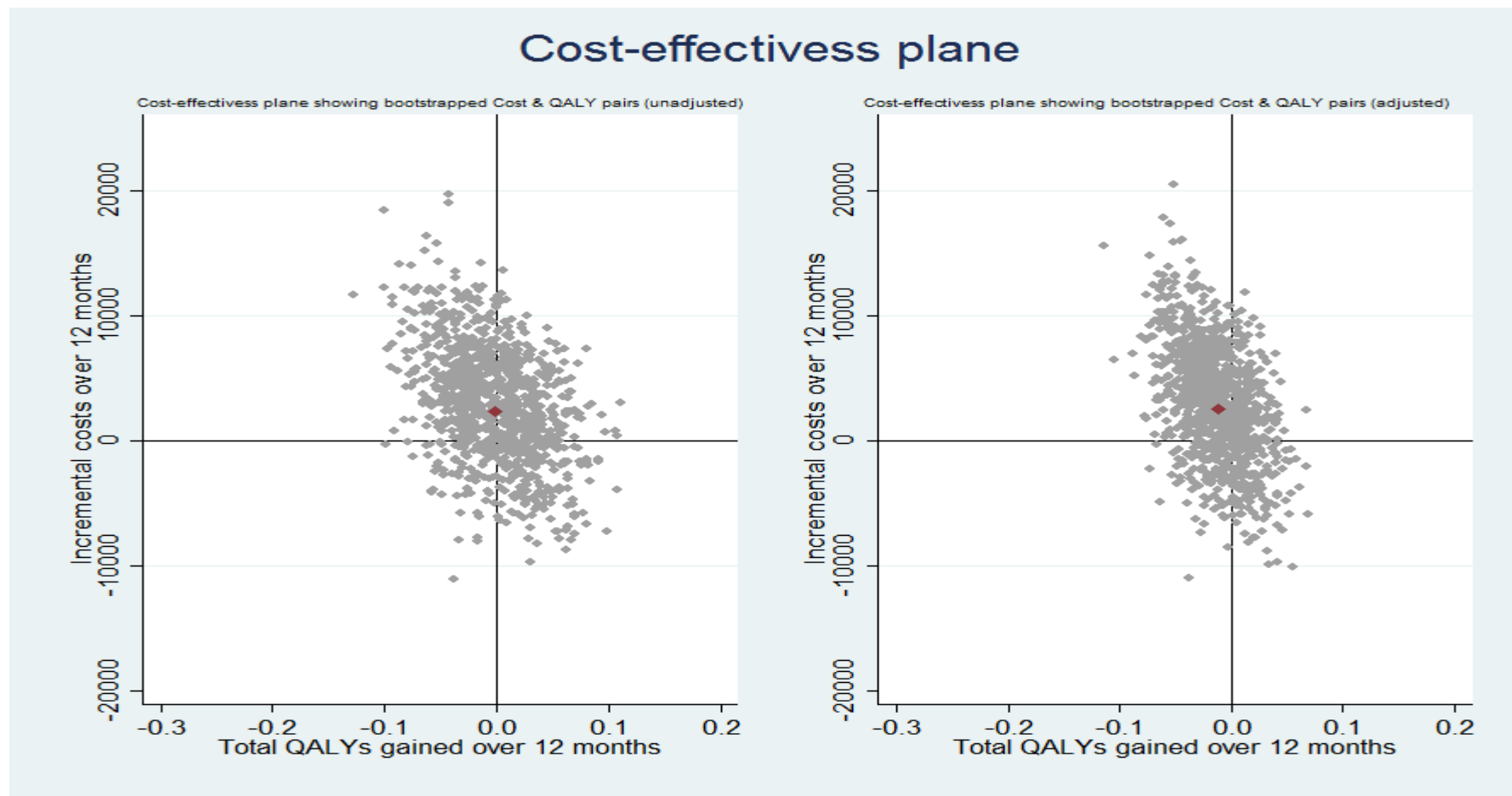
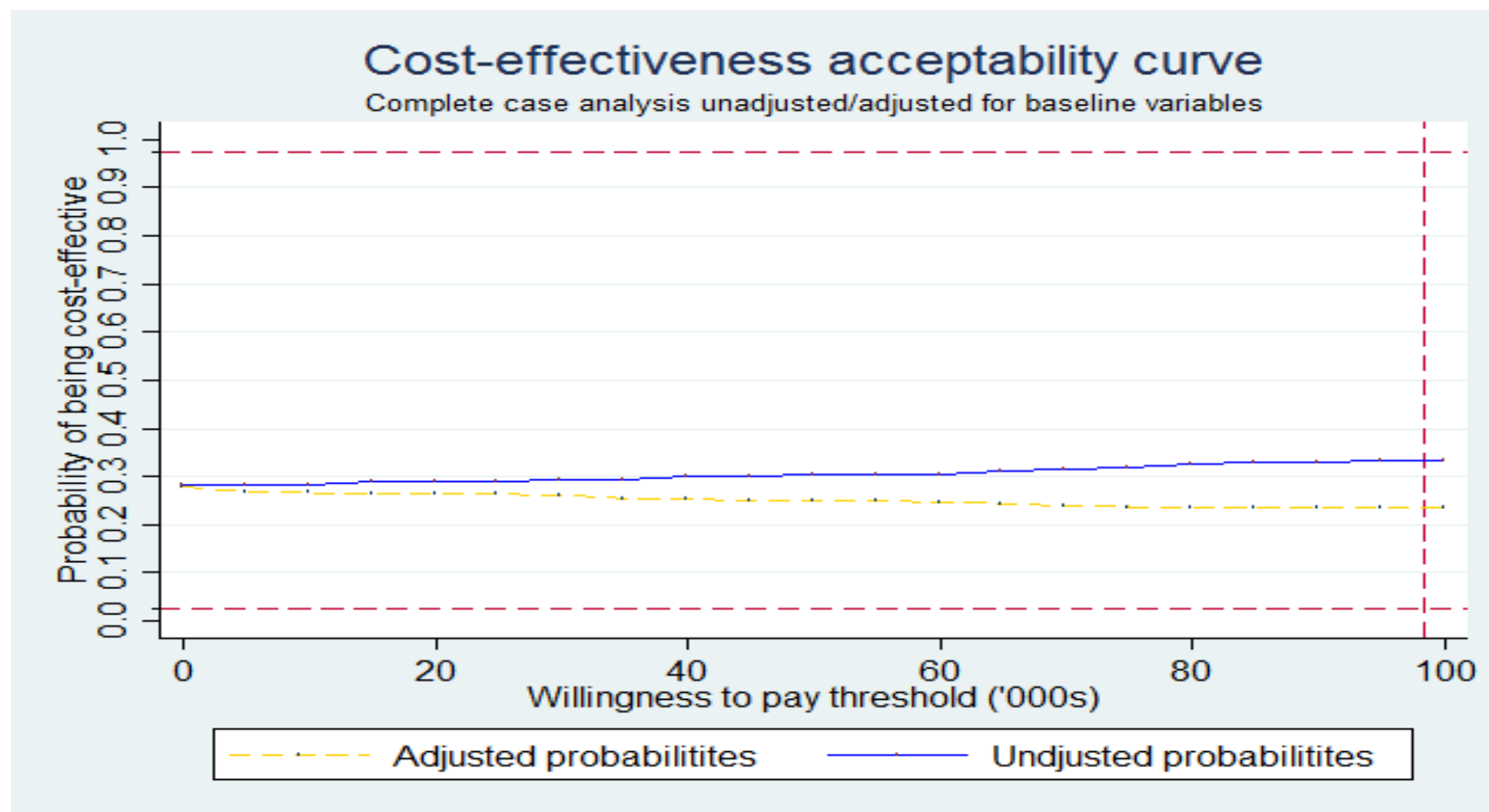


Figure 7.5: Cost effectiveness acceptability curve (complete case analysis), FIMM Vs MIMM



7.7.4 Sensitivity Analyses

In total, 15 different scenarios were explored and the results are presented in a table in Appendix 12.

- a) The base case analyses were undertaken with 45% missingness in both cost and QALYs at the 12-month time-point. In a sensitivity analysis, multiple imputation was undertaken both for missing cost and outcome data. This resulted in an unadjusted incremental cost of £7506 ($P=0.071$) and a differential QALY loss of 0.008 ($P=0.783$), indicating that FIMM is less effective and more costly than MIMM. On plotting 1000 bootstrapped pairs of incremental costs and QALYs (see **Appendix 13**), the chance of FIMM being dominated by MIMM was 61%, while the probability of FIMM being cost-effective at £20,000 was 2% (see **Appendix 14**).
- b) Parsimonious models using regression framework while adjusting for baseline covariates were also analysed and the conclusions were the same as for unadjusted analyses (see **Appendix 15**).
- c) Three individuals were identified as extreme outliers in the dataset. Two of them were from FIMM group; one of them used 192 days of sick leave with 3328 hours of information, while the other availed 280 days of sick leave and 1989 hours of informal care. The third individual was from the MIMM group and utilised 14 days of sick leave and 5200 hours of informal care. On excluding these three individuals and undertaking the same base case scenario, it was found that the incremental societal cost reduced

from £2963 in the base case to £1712 and there was an incremental QALY gain of 0.014 QALYs. The ICER was £122286 per QALY (north-east quadrant of cost-effectiveness plane) with a 32% probability of FIMM being cost-effective at a willingness to pay of £20,000 per QALY. However, repeating the analysis using a parsimonious regression model yielded a negative ICER point estimate in the north-west quadrant (similar to the base case analysis) with a 29 % probability of being cost effective at £20000.

- d) Repeating the analysis with estimated mean costs for NHS and social care (third party payer perspective), a broadly similar pattern of resource use across both interventions were observed with an incremental cost of £1118 (-1582 - 3817; $P=0.411$) and QALY loss of 0.013 QALYs (-0.07 - 0.04; $P= 0.659$) yielding a negative ICER of 486109 with a 25 % probability of being cost effective at £20000.

All the other scenarios (see **Appendix 12**) indicated similar results implying that FIMM was dominated by MIMM.

7.8 Discussion

In the UK, BD affects one percent of the population and accounts for around £5.14 billion annually to the society (294). Although the mainstay of BD management is an appropriate use of pharmaceutical agents, the results of previously published studies and existing guidelines unequivocally advocate adjunctive psychological interventions

to achieve greater improvement in clinical outcomes and quality of life in BD patients than pharmacological treatment alone (33, 271, 295, 296).

This is one of the first studies to explore the cost-effectiveness of a complex intervention including individual psychoeducation and electronic mood monitoring in the management of bipolar disorder in the UK. From a broad societal perspective, FIMM involving five sessions of psychoeducation was no more effective than the manualised integrated mood management and more costly, and is therefore not cost-effective. Overall, there were no statistically significant differences between treatment arms in the use of health and social care resources during the 12 months of assessment except the costs for intervention itself and medicines.

Even though the overall medicine consumption decreased in both arms over the trial period, the proportion of main drugs used in BD including anticonvulsants, lithium and antidepressants were higher in FIMM. Further, a statistically significant increase in medication costs was found in FIMM participants; this could be due a combination of higher medicine burden and also consumption of higher proportion of costlier medicines. For example, 8 out of 9 participants who reported Aripiprazole (a second-generation antipsychotic) use at the end of 12 months' time-point belonged to FIMM. Cost of one year treatment with Aripiprazole is around £2714, but the costs are expected to decrease following the introduction of generic versions of Aripiprazole in the near future and can lead to considerable cost savings (297).

Indirect costs including productivity losses due to sickness absence and informal care accounted for 71% of societal costs in the FIMM group, while in the MIMM group it amounted to 74% of total costs. However, these differences were not significantly different from each other.

Analysis of primary economic outcome, QALYs, further emphasised the lack of evidence in favour of FIMM. Over a period of one year, the individuals under the FIMM arm had lower incremental QALYs than the MIMM arm (Diff:-0.002, P= 0.961). Sensitivity analyses of ICERs resulted in MIMM being dominant in all but one scenario where extreme outliers were excluded. The cost-effectiveness acceptability curves indicated that the probability of FIMM being cost-effective remained less than 35%, regardless of chosen scenario.

7.8.1 Strengths and Limitations

This was a within-trial economic evaluation along with a complex RCT of a complex intervention that included patients using remote mood monitoring. **Chapter 3** and **Chapter 4** highlighted the elements of an economic evaluation and provided key suggestions for conducting a trial-based economic evaluation in the context of a mental illness. Drawing information from previous chapters, the economic evaluation was undertaken as per good reporting practises (166). Multiple scenarios were explored through sensitivity analyses and uncertainty in the cost-effectiveness ratio was analysed using the nonparametric bootstrap approach (298).

The sample size of 121 individuals in the current study was greater than the average number of participants (n=86 in both arms) enrolled in similar trials of psychological treatments in BD (299). Data on resource use were collected every three months from self-reported interviewer administered questionnaires. Despite the comprehensive nature of resource use assessment, missingness of up to 45 % of data was found in costs and outcomes. This is a major challenge for analysis in mental health clinical trials which often report missingness up to 50% of data (299, 300). Although missingness was handled appropriately using multiple imputation, the assumption of missingness at random cannot be verified. The pattern of dropouts or missingness may be related to disease severity, therefore implying missing not at random (MNAR) (288).

Although the trial compared two different types of psychoeducation, this was a complex intervention that included patients using remote mood monitoring. The costs for mood monitoring system was not included as their costs were common in both arms. Ideally, there is a need to separate the costs and effects of the components of a multi-component intervention, but it was not possible in this trial. However, the compliance with the remote mood monitoring was compared by estimating the proportion of weekly QIDS responses sent over 52 weeks and found no difference between the treatment arms (76% Vs 69.9%, diff:7.1%, P=0.168). Nonetheless, the use of remote mood monitoring and an active comparison group may have reduced the differences in the chosen interventions.

As is the case for many non-pharmacological behavioural interventions, it was challenging to identify an appropriate treatment-as-usual control group (301-303). A

variant of usual care, an “enhanced usual care” in the form of manualised integrated mood management was adopted as a control in this trial. MIMM received care as usual along with a manual for integrated mood management. This degree of enhancement was minimal and systematically chosen to mimic the routine care that the participants would anyway receive in the form of patient booklet if they were not involved in a RCT(304). However, at the same time, studies have argued that in pragmatic trials of psychological interventions, other forms of treatment-as-usual control groups such as a wait-list control group or control group with no information should not be employed as they do not control the placebo effects or non-specific effects of a therapeutic relationship (305, 306).

Participants included in this trial were randomly selected from a larger cohort of patients using remote mood monitoring and attending clinics in Oxfordshire and Buckinghamshire; this can affect the external validity of the study. However, systematic mood monitoring has been an integral component in the long term management of BD in the form of mood diaries, clinical monitoring forms and other paper based self-monitoring tools (36, 69). Mobile phone-based self-monitoring is not unique to the field of BD, and have showed promising results in long term disease management of other chronic diseases (307).

The findings of this economic analysis are consistent with the clinical study of this trial, which failed to find any difference in the main clinical outcomes over the 12 months follow up period. This could possibly be explained by the shorter duration of intervention (five sessions) in comparison to the average number of sessions ranging from 13 to 15. Nevertheless, it was also noted in the main trial that the knowledge

level as measured by the OBQ questionnaire showed greater increase in FIMM at the 3 month follow-up point, but this difference in knowledge scores was not sustained and the score converged at the 12 month time point (236). It is hypothesised that psychoeducation may increase the knowledge about management of illness including identifying warning symptoms of relapse, drug adherence and seeking prompt care (308). A recent systematic review of 16 psychological trials found a lack of evidence for individual-based psychoeducation programs but indicated that group psychoeducation appears to be effective in preventing relapse in bipolar disorder. At the same time authors recommended further research on individual psychoeducation as it is currently recommended by the NICE guidelines in the long-term management of BD (7, 299).

However, the mediating mechanism through which psychoeducation influences preference based health related quality of measurements is unclear. To complicate things further, as **Chapter 6** highlighted, it is challenging to know whether the EQ-5D-3L is sensitive enough to capture potential health benefits in individuals with BD. The relative lack of sensitiveness of EQ-5D-3L to clinical change could have diluted the differences between the two arms. The other economic outcome used in this study such as the ICECAP-A was similar across the time-points in both arms. The OXCAP-MH showed minimal difference at the 12th month time-point, but it is not sure whether that difference is clinically relevant. Further, OXCAP-MH is not a preference-based generic measure, thereby limiting its role in economic evaluation (281).

7.8.2 Existing literature

Despite+ the above limitations, to the best of my knowledge this study presents the only evidence of the cost-effectiveness of a facilitated individual-based psychoeducation for a pragmatic sample of individuals with BD in the UK. One similar study in the UK reported a cost-effectiveness study of adjuvant cognitive behavioural therapy (CBT) based on a randomised controlled trial. The study did not estimate QALYs, but reported “bipolar free day” as an outcome; it indicated that adjuvant CBT is a cost-effective treatment option for preventing relapse in patients with BD (145). However, the usual care was poorly defined and it was difficult to conclude with which concomitant pharmacotherapy CBT would work best.

There were two other collaborative intervention studies based in the United States that explored the costs and effects of multicomponent interventions (147, 148). Bauer et al (2006) conducted a costs consequence analysis in patients with BD in a veteran’s health care setting based in US, reported that collaborative care provides 4.5 additional weeks free of Manic episodes at no extra cost. The components of the program included a) Support to providers in the form of reference manual b) Structured group psycho-educational program c) Facilitated feedback by nurse care coordinator. Simon et al (2006) conducted a cost-effectiveness analysis in patients with having any diagnosis of bipolar spectrum disorder (bipolar disorder type I or type II, schizoaffective disorder, or cyclothymia) in a managed care organisation based in US setting. The interventions consisted of a) Care coordination b) structured monthly telephone calls c) Feedback to the mental health treatment team d) Structured group psycho-educational program. They reported that at an incremental cost of \$1251, people who were subjected to collaborative care had 5.5 additional weeks free of

262

severe manic symptoms than the control group. Even though both studies provide economic evidence for collaborative care, the applicability of results is limited to the study settings and type of health care structure.

7.9 Conclusion

There is a high societal burden of BD in the UK; as per NICE guidance, psychological interventions that has been developed specifically for BD should be offered concomitantly with pharmacotherapy as first-line treatment for adults in the longer-term management of the disorder (7, 294). However, the evidence for effective psychological treatments is scarce and that for cost-effectiveness is almost non-existent in the UK setting (270, 309, 310). The current trial was designed to fill the gap in the evidence for BD management in secondary care settings, and to test the hypothesis that FIMM strategy is more effective and cost-effective than the alternative MIMM.

The findings reported in this chapter indicate that using FIMM with five sessions (relative to MIMM) for managing BD is unlikely to be cost-effective. A psychoeducational intervention with an increased number of sessions over a longer time horizon or a trial comparison with a group psychoeducation could be explored for future research. At the same time, there was high compliance with remote mood monitoring and provision of integrated mood management in the form of a manual in addition to usual care was the dominant strategy to improve outcomes over a period of 12 months. Given the increasing importance of self-management in chronic illness,

this line of research may be of considerable importance in the field of public mental health.

The aim of this chapter was to provide economic evidence for the NICE recommended psychoeducation program and inform rational decision-making given scarce health care resources. **Chapter 8** concludes this thesis by considering all the evidence presented in this chapter and previous chapters.

Chapter 8: Discussion and conclusions

8.1 Introduction

The overall aim of this thesis is to contribute to mental health economics by providing economic evidence to relevant policy and methodological questions relating to the management of Bipolar disorder (BD) in the UK, which have been recognised as important in the current literature.

The purpose of this discussion chapter is to present an overview of the research findings followed by a discussion of the challenges observed during this research. Subsequently, contribution to the literature and relevant policy implications are put forward. Finally, directions for future research will be suggested.

8.2 Overview of thesis findings

Bipolar disorder (BD) is one of the leading causes of disability worldwide. BD is chronic in nature with frequent recurrences that could have a debilitating impact as well. It can occur at any age, but typically the diagnosis is made in early adolescence or adulthood, implying that BD has profound implications on this population over several decades. Patients with BD not only have significantly higher rates of medical co-morbidities such as anxiety disorders, attention-deficit hyperactivity disorder, metabolic syndrome, and substance abuse disorder but also experience higher rates of suicide attempts. People with BD die on average 10-15 years earlier than the

general population, with a substantial proportion of these deaths attributed to avoidable risk factors and co-morbid cardiovascular diseases.

The life course of BD is poorly understood, and patients are heterogeneous in their symptoms and signs. The disease has no cure, and its natural history follows a protracted waxing and waning course with frequent relapses, making the management of BD a massive challenge. However, this challenge is compounded by the frequent misdiagnosis of BD as unipolar depression with an extended delay of up to ten years in its diagnosis.

In current times of constraining healthcare budgets, the economic case of a disease plays a crucial role in shaping public debate and concern and leverage policy decisions. It is vital to understand the relative burden of illness on the society so that appropriate investments are made to identify cost-effective interventions and implement health policy decisions to achieve maximum social benefit. This is particularly important in the case for BD, where there is a sheer dearth of evidence-base on the prevalence of disease, its costs, the relevance of outcomes to capture health benefits, and cost-effectiveness of management strategies in the UK.

Chapters 1 to 3 provided the background, rationale, clinical and economic framework required to understand the methods proposed in the following chapters that explored aspects of the health economics of BD that are pivotal for priority setting and decision making. These elements included a systematic review and critical appraisal of the

existing economic evaluations in BD, the economic burden of BD in the UK, including the factors that are associated with the distribution of costs within and beyond the health sector, and appropriateness of the widely used EQ-5D-3L in quantifying health benefits in BD. Drawing from the evidence and suggestions from **Chapters 3-6**, an appropriate method of economic evaluation was used to illustrate the cost-effectiveness of alternative psychoeducational strategies in the long-term management of BD in **Chapter 7**. The results from these chapters are briefly discussed below.

Identification of the most cost-effective interventions in the field of BD represents one of the key considerations into the broader task of informing and assisting the decision-maker in allocating scarce health care resources, but to **what extent does the current evidence go in achieving this role?**

Chapter 4 reviewed and critiqued published economic evaluations of both pharmacological and psychological therapies in the management of BD. Pharmaceutical companies funded 60% of the studies identified in this review. Overall, the trial-based evaluations were of poor quality and primarily were cost-consequences analyses. These studies presented favourable results for collaborative care models and cognitive behavioural therapy (valid in the included study setting) but not pharmaceutical options. The quality of modelling based studies was variable with a great deal of methodological heterogeneity and ambiguity about the correct evaluation methods.

Given the variability of methods and poor quality of identified studies, it is challenging to provide a definitive recommendation for the most cost-effective therapy in the management of BD. Nevertheless, based on the consistency of findings, it can be concluded that Lithium (LI) or its combination with atypical antipsychotics like Quetiapine (QTP) are cost-effective pharmacotherapies in the management of BD.

This chapter also highlighted the ideal characteristics of cost effectiveness studies including

- a) Adoption of societal perspective,
- b) Inclusion of relevant treatment comparators with clear description,
- c) Transparency in costing and outcome parameters,
- d) The need to consider extrapolating results from a trial using models, and having a lifetime horizon for models
- e) The importance of the QALY as the health-related quality of life (HRQoL) outcome measure along with its associated caveats in mental illness and
- f) Appropriate handling of uncertainty.

These aspects were considered in subsequent chapters. However, given the severe health consequences of BD as highlighted in previous chapters, it concerns to find that more economic evaluations have not been funded. NICE recommends psychological therapies for the long-term management of BD, but the apparent lack of high-quality studies, especially for psychological interventions indicates an urgent

need for further empirical research in economic evaluations of psychological interventions (7, 68, 270).

The discussion so far has indicated that despite BD being a severe mental illness with high mortality rates and poor quality of life, there is a dearth of economic evidence on available interventions in the field of BD management; a lack of which impedes the possibility of rolling out interventions widely that could potentially lead to better outcomes at a lower price. This could be due to relative lack of awareness or limited knowledge on the burden of disease or lack of effective interventions. **So, what is the scale of the problem? What are the associated costs and distribution impact across sectors?**

Chapter 5 provided the most comprehensive evidence till date of BD in the UK, indicating that BD is a disease of substantial economic burden. Societal costs of BD were estimated using a prospective cohort study design. The estimated prevalence of BD in the UK was 0.8%; applying it to the 2010 adult UK population of 50.9 million (≥ 16 years of age), the total number of individuals with BD in the UK was estimated to be 407,567.

The majority of the economic burden fell on the society as productivity losses and informal care (68.1%); while hospitalisations represented the next most significant cost burden (15.5%). Around 6.8% of overall costs was attributed to medicines. The total annual costs for BD in the UK were estimated to be £5.14 billion for year 2010/11,

translating to £1.57 billion costs to the healthcare sector in the UK or £1.31 billion in England. The average healthcare cost of £3,859 was considerably higher than the treasury reported average healthcare expenditure of £1,926 per person in the UK showing the increased need of patients with BD for health care services. Further, change in severity of depression or mania resulted in increased total, healthcare and indirect costs. Specifically, costs associated with productivity losses and informal care were significantly associated with depression and healthcare costs with mania. Secondly, increase in severity of depression or mania resulted in increased total, healthcare and indirect costs. These are important results because they demonstrate a relationship with average symptom level over 12 months; routine monitoring of self-reported health status in BD has a potential to provide a relatively simple method to estimate service performance and improve patient outcomes at the population level.

The extraordinary societal burden of BD emphasises the need for increased attention to the illness in society in general and acknowledgement from policy makers with an aim to increase funding to search for cost-effective treatment strategies in the long-term management of BD.

The urgent need of evidence on cost-effectiveness highlighted by **Chapter 5** requires reliable measurement of health outcomes in individuals with BD. In the UK, the National Institute for Health & Care Excellence (NICE) recommends the use of the generic preference-based multi-attribute utility (MAU) instrument EQ-5D-3L for conducting economic evaluation and majority of NICE clinical guidelines across all disease areas are based on cost-utility analysis (121).

However, there is insufficient evidence with regards to its appropriateness in BD and the use of EQ-5D-3L so far has been rather limited in psychiatric studies, thus precluding the estimation of health benefits (i.e. utilities) required for priority setting and resource allocation (30). In situations where EQ-5D-3L is unavailable, NICE recommends mapping (as only a second best) of other available instruments onto EQ-5D-3L (121) Given the ambiguity around the appropriateness of EQ-5D-3L and the relative reluctance of its use in mental health clinical practise, **Is the EQ-5D-3L appropriate to measure HRQoL in individuals with BD? In the absence of EQ-5D-3L data, is it feasible to map BD-specific instrument on EQ-5D-3L?**

Chapter 6 investigated the appropriateness of the EQ-5D-3L to measure in a group of individuals with bipolar disorder and explore whether it is feasible to map EQ-5D-3L from disease-specific mood instruments, when EQ-5D-3L is not available.

121 participants with BD from the OXTEXT-6 study (brief description of the OXTEXT case study is provided in **Chapter 1**) provided 1243 set of responses. These contemporaneous responses were for two measures of disease-specific outcomes including the Quick Inventory of Depressive Symptoms–Self Report (QIDS-SR16) and the Altman Self-Rating Mania Scale (ASRM) and one generic preference-based-measure (PBM) the EQ-5D-3L over the one-year follow-up. The reliability, construct validity and responsiveness of the EQ-5D-3L were assessed in comparison with disease-specific indicators for depression and mania. Mixed method modelling using a maximum likelihood approach was adopted to explore the ability of EQ-5D-3L to

differentiate across varying levels of disease severity. The effect sizes (ESs) and standardised response means (SRMs) were computed to estimate the strength of discrimination between different disease states and assess changes over time. The feasibility of mapping was determined by using exploratory and confirmatory factor analyses and goodness of fit assessed for a series of models.

This study demonstrated that the EQ-5D-3L was useful in measuring HRQoL in BD patients who predominantly experience depressive symptoms, but was not sensitive enough to detect changes in individuals with varying levels of mania, therefore not reflecting the whole spectrum of the disorder. Results of both explanatory and confirmatory factor analyses indicated that multidimensional measures of depression, HRQoL and mania cannot be considered sufficiently unidimensional to allow mapping estimation of health-related utility.

As per NICE guidance, psychological interventions that have been developed specifically for BD should be offered concomitantly with pharmacotherapy as first-line treatment for adults in the longer-term management of the disorder (7, 294). However, the evidence for effective psychological treatments is scarce, and that for cost-effectiveness is almost non-existent in the UK setting (270, 309, 310)(269, 308, 309). Despite this lack of economic evidence, psychoeducation is considered as the first-line psychological treatment strategy in the long-term management of BD (270). Keeping the evidence regarding the appropriateness of EQ-5D-3L derived in **Chapter 6** in context, the final **Chapter 7** answered the question on **whether Psychoeducation in BD is cost-effective?**

A Cost-utility analysis (CUA) was conducted alongside a pragmatic randomised controlled trial (OXTEXT-6) assessing the costs and outcomes of two interventions: facilitated integrated mood management (FIMM, n=60) and manualised integrated mood management (MIMM, n=61). Costs and outcomes were assessed at baseline and at 3, 6, 9 and 12 months follow-ups. Incremental cost per quality adjusted life year (QALY) were estimated over a period of one year from a societal perspective. Cost-utility ratios were estimated based on complete case analysis for the base case analysis. Sensitivity analyses were undertaken for 15 different scenarios including using a multiple imputation approach for missing data. Uncertainty was presented using the cost-effectiveness plane and the cost-effectiveness acceptability curve (CEAC).

The trial included 121 participants, aged 16 years and above, with bipolar disorder who were participating in a remote mood monitoring system. The base case analysis found no statistically significant differences in costs and outcomes between the groups. FIMM was more costly and less effective than MIMM, implying that MIMM dominated FIMM strategy. Uncertainty analysis using CEACs indicated that there is less than 30% likelihood that FIMM is more cost-effective than MIMM. Sensitivity analysis using multiple imputation for missing data generated similar results, with FIMM appearing less effective and more costly than MIMM; there was a 65% probability that FIMM was dominated by MIMM.

The subsequent sections of this chapter outline the challenges surrounding the main objectives of this thesis. The chapter concludes by highlighting the main policy

implications and potential area for future research in the field of health economics of BD

8.3 Challenges of this research

- 1) The main aim of the systematic review was to identify cost-effective pharmacological and psychological interventions in the management of BD. However, it was not possible to arrive at a conclusive evidence on the relative cost-effectiveness of interventions due to heterogeneity across studies related to study design, comparators, and other factors described in the previous section. Appropriate quality checklists were used to review the quality of trial-based and modelling studies. However, all items were weighted equally, and the scoring was inherently subjective. This could have resulted in wrongly classifying some studies into poor quality studies. Nevertheless, the difference in quality scores between individual studies was large, and the conclusion would still have remained the same if different scoring was used.
- 2) Notwithstanding the considerable benefits of using a prospective cohort of individuals in determining COI estimates in **Chapter 5**, there were some challenges in doing so. Perhaps, the greatest challenge was the small sample size of 91 individuals who are part of a cohort of individuals accessing mental healthcare across Oxfordshire and Buckinghamshire counties in England, which could affect the generalisability of results. Due to the lack of data, productivity losses for unemployed individuals whose unemployment was not attributed to BD

and premature mortality due to BD were not included. It was also not possible to include costs due to contact with the criminal justice system and disrupted/lost education. This may have contributed to underestimation of the overall societal costs. At the same time, the treatment seeking nature of participants meant that it may be possible that their resource utilisation is higher than those who do not access care.

Another challenge was the varying levels of missing data. For instance, sickness absence data was missing between 15% at the third-month time-point to 25 % at the 12th-month time-point. The pooled prevalence estimate for the UK was based on the assumption that the prevalence found in included European countries would be similar to that of the UK. Further, the pooled prevalence estimate was not adjusted for the quality of included studies.

Even though the sample size was small, the use of a longitudinal design allowed around 364 responses. Appropriate statistical techniques including multiple imputations, generalised estimating equations and probability sensitivity analyses were used to address some of the shortcomings (311, 312). The results of the post-imputed analyses were also compared with complete-case information and no significant differences were noted between the results of the two methods. The pooled prevalence estimated in this study was similar to the estimates reported by a recently published global burden of disease study (5) The age, gender, type of disease and primary care use observed in this study was similar to cohort of patients (4700- 6671 patients) with BD registered in over 470 general practises across UK (223, 224).

- 3) The analysis of appropriateness of EQ-5D-3L in BD was conducted on the OXTEXT6 dataset. This analysis required concurrent availability of disease-specific measures and the EQ-5D-3L. The OXTEXT-6 dataset included participants from a randomised control trial (RCT) who were engaged with their health care providers. The patients were in remission or a stable state at the baseline. The dataset population may not be fully representative of individuals with varying levels of BD disease symptomatology.

Secondly, a challenge that follows the previous point is that most of respondents were in level 1 or 2 in the Anxiety/ depression dimension of the EQ-5D-3L. That means estimation of convergent validity using a correlation measure based on just two categories may not be an appropriate test. EQ-5D-3L is generic PBM and it was designed to be sensitive to different types of health effects not just mental health. So, in the presence of a correlation there is evidence to support validity, but failure to find a strong correlation may not necessarily mean that the measure is not valid.

A further limitation was that, for the test-retest and responsiveness analyses, the stable group or groups that experienced change (improvement or deterioration) were classified by clinical scores, due to unavailability of a gold standard instrument to measure HRQoL in BD. The resulting groups were small, especially those that indicated change in their mental health symptoms.

The sample of this study may be small, but 121 participants enrolled in the trial provided 13 repeated measurements over time, with the total sample size of all the three contemporaneous measurements evaluated in this study increasing to 1257 out of a possible 1547 responses. These measures provided a wide range of responses over the 12-month follow-up period. Although correlation analyses of PBMs (i.e. measure of association/dissociation) with clinical status measures is not the best way to assess validity, the confirmatory factor analyses presented in **Chapter 6** provided a robust complementary evidence of conceptual similarity by looking at the health state classification of EQ-5D-3L that generates preference-scores. The lack of a gold standard HRQoL means that the creation of sub-groups for conducting test-retest and responsiveness based on clinical scales is not unusual in mental health research, but accordingly the results of such studies should be considered in light of this limitation (30).

- 4) The trial-based evaluation of a standardised psychoeducation intervention did not find any economic evidence to support psychoeducation delivered by facilitators in comparison with psycho-education using a manual (self-study mode). The trial, however indicated that FIMM improved knowledge. It may be that the sample size was not sufficiently powered to detect differences in economic outcomes. Another interesting challenge that follows the evidence from the previous chapter was the apparent lack of sensitivity of the EQ-5D-3L across the spectrum of BD symptomatology. In the extreme scenario, this may have resulted in the dilution of any difference in health benefits in between the two arms of the trial. Further, the limitations of missingness in data for OXTEXT6 and associated generalisability was described in detail in **Chapter 7**.

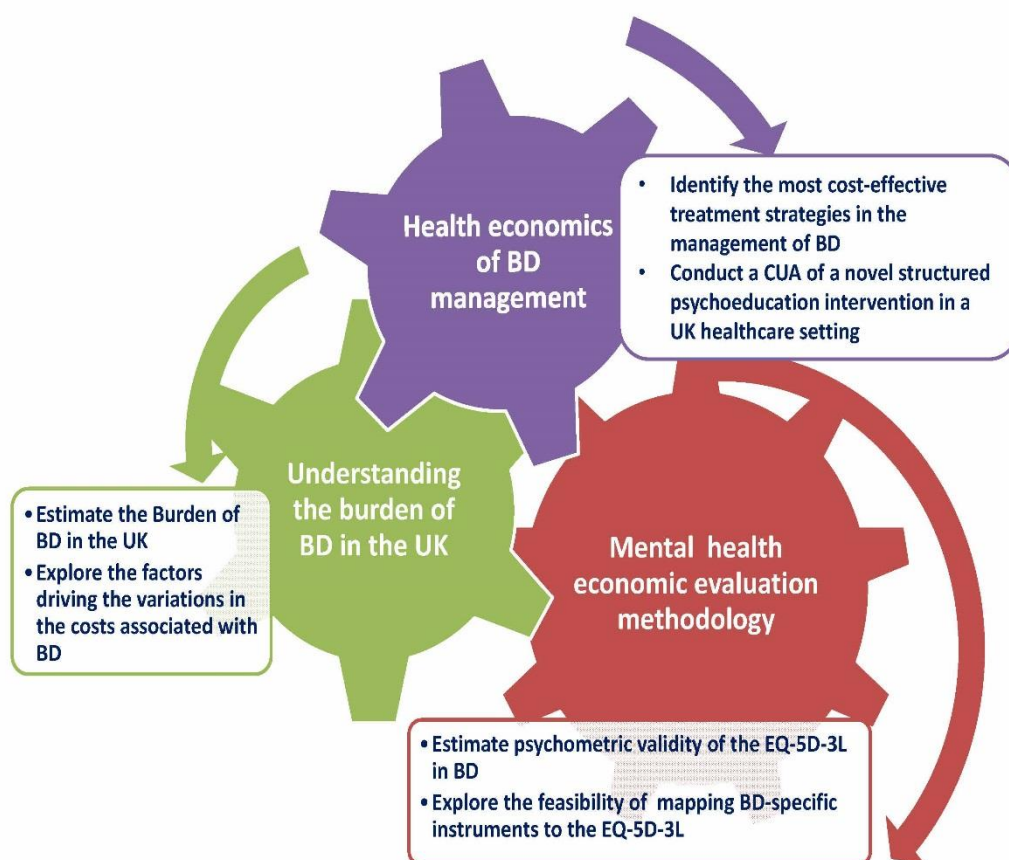
It may be argued that in chronic diseases like BD where costs and benefits of treatment typically accrue over a period longer than the duration of most RCTs, it may not be necessary to estimate sample size or power to test a hypothesis based on economic outcomes. A sample size based on economic endpoints would be necessarily larger and it would be unethical to continue a trial once effectiveness is established based on clinical outcomes (166, 313). However, it is unlikely that a five-session psychoeducation intervention would have resulted in a longer term impact.

The findings of this research presented here need to be considered along with its associated challenges; the next section provides a description of the contribution this thesis makes to the literature.

8.4 Contribution to the literature

This thesis contributes to the existing literature in three research domains: the health economics of BD management, understanding the burden of BD in the UK and mental health economic evaluation methodology (see **Figure 8.1**).

Figure 8.1: Contribution to the literature



The first contribution to the health economics literature is made by a systematic review of economic evaluations for BD management strategies. Studies were critiqued with the aim of identifying the most cost-effective psychological and pharmacological treatment strategies in the management of BD. The results of this review indicate that there is a dearth of high-quality economic evidence for BD management in the UK, especially for psychological interventions in the long-term management of patients. There was a great deal of heterogeneity observed in the included studies with

relatively poor quality of evidence. As a result, it is difficult to provide a conclusive recommendation for the most cost-effective therapy for BD in the UK.

Given the financial crisis in UK health care funding, demonstrating that an intervention is effective is no longer adequate. The CUA presented in **Chapter 7** advances this literature, providing the first evidence of the cost-effectiveness of a novel psychoeducation intervention in a UK healthcare setting. As per existing guidelines, psychoeducation is the first choice of psychological intervention in the long-term management of BD. This CUA indicated that a complex intervention comprising of five psycho-education sessions with a facilitator in addition to remote mood monitoring failed to improve the primary outcome and was not cost-effective in comparison to remote mood monitoring and self-administered psychoeducation using manuals for managing individuals with bipolar disorder. However, both forms of psychoeducation were associated with an increased awareness of disease symptoms that are essential for improved self-management of BD.

There is a relative lack of awareness or limited knowledge on the burden of BD in the UK. This thesis also contributes to an improved understanding of the burden of BD in the UK. The cost-of-illness study in **Chapter 5** provided a comprehensive estimate of the societal burden of BD derived from a pragmatic cohort of individuals with BD in the UK. This study estimates that the annual societal costs for BD in the UK were £5.14 billion for the year 2010/11. Of this, £1.57 billion is attributable to the UK healthcare sector. Healthcare and indirect costs (e.g. productivity losses due to illness, and informal care provided by family and friends) were significantly associated

with self-reported levels of symptoms. This suggests that regularly collected self-reported outcomes are of direct use to health service planners and commissioners.

The third significant contribution is related to mental health economic evaluation methodology. Given the need for consistency across economic evaluations, the National Institute for Health and Care Excellence currently recommends the EQ-5D-3L as a tool for quantifying utility and conducting economic appraisals. However, there is a lack of research examining the use of the EQ-5D-3L instrument in BD. This study has demonstrated that the EQ-5D-3L is useful in measuring HRQoL in BD patients who predominantly experience depressive symptoms, but is not sensitive enough to detect changes in individuals with varying levels of mania, therefore not reflecting the whole spectrum of the disorder. More evidence is therefore required before the EQ-5D-3L instrument can be widely applied in economic evaluations of BD-related interventions. Furthermore, it is not possible to map between disease-specific instruments and the EQ-5D-3L in BD as there is a lack of conceptual similarity between the scales.

Finally, findings from two of the chapter have been published in peer-reviewed journals and this would contribute to wide dissemination of this research's findings (294, 310).

8.5 Policy context and implications

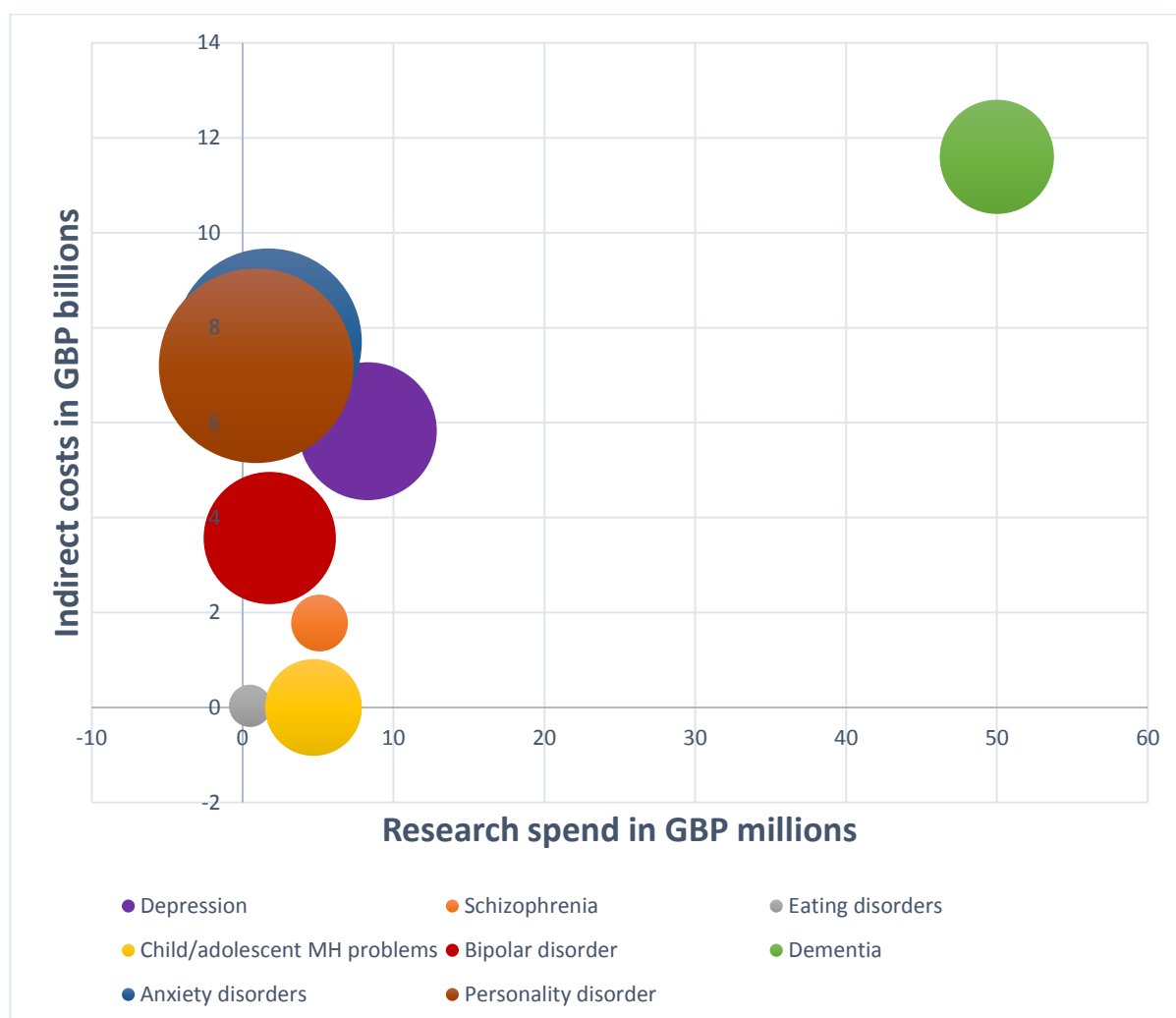
BD could lead to a significant life-long economic burden to the patients, their families and wider society which is likely to increase with an ageing population and escalating healthcare expenditure. However,

“Bipolar disorder has been and still is a relatively neglected condition” (270).

Before, considering the policy implications of this thesis in detail, it would be helpful to briefly understand the landscape of existing mental health priorities. Research investigating the burden of BD has indicated that indirect costs of the disease mainly drive the total costs of BD; this cost distribution is unlike that for schizophrenia or dementia, where much of the total cost falls directly on health and social care system. With a prevalence of 0.21 million people, the total costs of schizophrenia were reported as approximately £4 billion per year in England (based on 2007 estimates) including £2.23 (56%) in health and social service costs and £1.78 (44%) as productivity losses (314).

Figure 8.2 highlights the burden of eight diseases in the UK and associated total mental health research funding allocations by disease category (8, 9, 22). This figure charts the indirect costs of disease on the x-axis, and the y-axis represent mental health research funding in 2013. **The size of each bubble represents prevalence of the disease.**

Figure 8.2: Burden of disease and associated research funding



With 0.81 million individuals living with dementia, the economic burden of dementia was estimated at £26.3 billion (2012-2013 prices) that included a similar distribution of costs as incurred in Schizophrenia (56% in health and social care costs, and 44% as indirect costs) (22). Schizophrenia receives around £5 million and dementia £50 million of research funding per year. On the other hand, 0.4 million living with BD and an associated burden of around £5.14 billion (31% in health and social service costs and 69% as indirect costs) receive a research funding of £1.8 million (9). The burden of BD need to be taken in context of the high rates of its misdiagnosis of up to 40% as

major depressive disorder or personality disorder (315, 316)(314, 315)(314, 315)(314, 315)(315, 316). Taken together, it indicates that the levels of funding for BD are not aligned with the disease burden in the UK, and there is a funding gap that could have a significant impact on health, financial & social outcomes for individuals with BD and the wider society.

Implications:

The findings reported in this thesis have a number of consequences for policymaking. One of the main conclusion is that there is a sheer dearth of economic evidence needed to support implementation of cost-effectiveness strategies for the long-term management of BD in the UK, especially the evidence for psychological therapies is almost non-existent. Given the substantial health related implications of BD, it was surprising to find that more economic evaluations have not been funded.

The results of the cost-of-illness study of BD confirmed the enormous human and societal consequences of BD. Most of the economic burden fell on the society as productivity losses and informal care (68.1%); while hospitalisations represented the next most significant cost burden (15.5%). Around 6.8% of overall costs was attributed to medicines. Policy makers need to pay attention to this distribution of expenditure in BD as medicines represent only a fraction of the total costs, implying that the focus of any future interventions should be related to improving the functionality and productivity of individuals with BD.

The COI study also demonstrated that increase in severity of depression or mania resulted in increased total, healthcare and indirect costs. Specifically, costs associated with productivity losses and informal care were significantly associated with depression and healthcare costs with mania. Those are important results because they demonstrate a relationship with average symptom level over 12 months. Average symptom levels are largely determined by the clinical state between episodes and are not usually regarded as a serious target for management. Instead, all long term treatment studies to this point focus on the prevention of relapse (317).

From a public mental health perspective, these results generate the hypothesis that significant cost savings could result from improved monitoring and control of self-reported symptoms between episodes using a low-cost and relatively simple method of mood monitoring.

In the absence of a gold standard HRQoL instrument, it is a challenging and iterative process to establish appropriateness of an instrument. It is only possible to determine the varying levels of validity, the extent to which evidence fits or does not fit with the responses from this group of patients. EQ-5D-3L is recommended by NICE for evaluating outcomes. Although EQ-5D-3L was appropriate to measure HRQoL in BD patients who predominantly experience depressive symptoms, it was not sensitive enough to detect changes in individuals with varying levels of mania, therefore not reflecting the whole spectrum of the disorder. More evidence is therefore required before the EQ-5D-3L instrument can be widely applied in economic evaluations of BD-related interventions. Regardless, until the appropriateness of EQ-5D is established,

outcome measurement should include contextually validated clinical instruments along with a suitable multi-attribute utility scale.

The findings reported in the CUA indicate that using FIMM with five sessions (relative to MIMM) for managing BD is unlikely to be cost-effective. Given that NICE recommends psychoeducation program, it is important to explore other delivery models or different intensity levels of psychological interventions.

8.6 Future research

Based on the thesis findings, there are several areas where future research could provide further insight. Individuals with BD consider employment as one of the major elements in their recovery process (25). Given the combination of lack of economic evidence on psychological interventions, and the preponderance of burden attributed to productivity losses, it is, therefore important to investigate the cost-effectiveness of interventions that improve the functioning of individuals with BD who are at work.

More evidence on psychological interventions is needed; a psychoeducational intervention with an increased number of sessions over a longer time horizon or a trial comparison with a group psychoeducation could be explored for future research. Further, there was high compliance with remote mood monitoring and provision of integrated mood management in the form of a manual. Given the increasing

importance of self-management in chronic illness, this line of research may be of considerable importance in the field of public mental health.

Further evidence is also required to evaluate whether it is feasible and cost-effective to use online mood monitoring in measuring system-level performance based on patient reported outcome measures in BD.

Whereas, the EQ-5D-3L has shown moderate to good validity in depressive states, the descriptive system of the EQ-5D-3L lacks the disease-specific constructs or sufficient levels of response to an item that are relevant to the current study population. For instance, it does not include features that will encompass responses related to “increased energy” and “elated mood”, which are key determinants of manic traits (318). Therefore, there is a need to explore measures with descriptive systems sensitive to the BD disease phenomenology. This could include but not limit to evaluating condition-specific measure such as Brief QoL.BD, well-being measures such as ICECAP-A and OxCAP-MH, and EQ-5D with five levels for each dimension (improving the instrument’s responsiveness and sensitivity) (88, 281, 319, 320).

More generally, given the lack of reliable economic evidence in several aspects of BD management, it is critical that further studies are conducted to facilitate priority setting and policy development in this challenging area.

8.7 Concluding remarks

Bipolar disorder (BD) is one of the leading causes of disability worldwide and has a detrimental impact on HRQoL, and personal and social functioning. The large premature mortality gap of up to 15 years is unfair and unacceptable. Despite this, there is insufficient knowledge of the costs, HRQoL implications relevant to BD, and also the cost effectiveness of current treatments for BD in the UK.

Through detailed empirical and methodological work, this thesis builds an economic case for BD in the UK by bringing together evidence on the societal burden of BD in the UK, appropriateness of the widely used EQ-5D-3L in BD, and cost-effectiveness of treatment strategies in the management of BD.

The deleterious health and societal consequences of BD emphasise the need for investing in identifying cost-effective interventions that make the best use of scarce societal resources. The lack of economic evidence on psychological interventions and relatively minor expenditure on drugs indicate that policy makers should prioritise their attention on interventions that improve functioning and participation in employment.

Health economists should consider the evidence that despite the preference of NICE on the use of EQ-5D-3L for the estimation of QALYs in the economic evaluation of healthcare, this may not be available or appropriate in BD. More evidence is therefore required before this instrument can be widely applied in economic evaluations of BD-related interventions.

The evidence presented here will be disseminated to health economists, public health professionals, primary care and secondary clinicians, policy-makers and wider audience through future peer-reviewed journal articles, and presentations with a view to communicate, interact, raise the profile of BD and influence mental health policy decisions to improve outcomes in patients with BD.

References:

1. Phillips ML, Kupfer DJ. Bipolar disorder diagnosis: challenges and future directions. *The Lancet*. 2013;381(9878):1663-71.
2. Mitchell P, Hadzi-Pavlovic D, Loo C. Course and Outcome of Bipolar Disorder. In: Manji HK, Zarate Jr CA, editors. *Behavioral Neurobiology of Bipolar Disorder and its Treatment*. Current Topics in Behavioral Neurosciences. 5: Springer Berlin Heidelberg; 2011. p. 1-18.
3. Conus P, Macneil C, McGorry PD. Public health significance of bipolar disorder: implications for early intervention and prevention. *Bipolar Disorders*. 2013;n/a-n/a.
4. IsHak WW, Brown K, Aye SS, Kahloon M, Mobaraki S, Hanna R. Health-related quality of life in bipolar disorder. *Bipolar Disorders*. 2012;14(1):6-18.
5. Ferrari AJ, Stockings E, Khoo JP, Erskine HE, Degenhardt L, Vos T, et al. The prevalence and burden of bipolar disorder: findings from the Global Burden of Disease Study 2013. *Bipolar Disorders*. 2016;18(5):440-50.
6. Geddes JR, Miklowitz DJ. Treatment of bipolar disorder. *The Lancet*. 2013;381(9878):1672-82.
7. NICE. Bipolar disorder (update): the management of bipolar disorder in adults, children and adolescents in primary and secondary care. (Clinical Guideline 185). In: Excellence NIfHaC, editor. London: National Institute for Health and Care Excellence; 2014.
8. McCrone P, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. *Paying the Price: The Cost of Mental Health Care in England to 2026*. London: King's Fund; 2008.
9. Mental Health Research Funding Landscape Report London: MQ Transforming Mental Health; 2015.
10. Harrison PJ, Cipriani A, Harmer CJ, Nobre AC, Saunders K, Goodwin GM, et al. Innovative approaches to bipolar disorder and its treatment. *Annals of the New York Academy of Sciences*. 2016;1366(1):76-89.
11. Shemilt I, Mugford M, Vale L, Marsh K, Donaldson C, Drummond M. Evidence synthesis, economics and public policy. *Research Synthesis Methods*. 2010;1(2):126-35.
12. Soares-Weiser K BV, Beynon S, Dunn G, Barbieri M, Duffy S, Geddes J, Gilbody S, Palmer S, Woolacott N. A systematic review and economic model of the clinical effectiveness and cost-effectiveness of interventions for preventing relapse in people with bipolar disorder.; 2007. Contract No.: 39.
13. Bridle C, Palmer S, Bagnall AM, Darba J, Duffy S, Sculpher M, et al. A rapid and systematic review and economic evaluation of the clinical and cost-effectiveness of newer drugs for treatment of mania associated with bipolar affective disorder. *Health Technology Assessment (Winchester, England)*. 2004;8(19):iii-iv, 1-187.
14. CLEMENT S, SINGH SP, BURNS T. Status of bipolar disorder research. *Bibliometric study*. 2003;182(2):148-52.
15. Parsonage M, Grant C, Stubbs J. Priorities for mental health. *Economic Report for the NHS England Mental Health Taskforce* London: Centre for Mental Health. 2016.
16. Munro J, Osborne S, Dearden L, Pascoe K, Gauthier A, Price M. Hospital treatment and management in relapse of schizophrenia in the UK: associated costs. *The Psychiatrist*. 2011;35(3):95-100.
17. Olié JP, Lévy E. Manic episodes: the direct cost of a three-month period following hospitalisation. *European Psychiatry*. 2002;17(5):278-86.
18. Kleine-Budde K, Touil E, Moock J, Bramesfeld A, Kawohl W, Rössler W. Cost of illness for bipolar disorder: a systematic review of the economic burden. *Bipolar disorders*. 2014;16(4):337-53.
19. Fajutrao L, Locklear J, Prialux J, Heyes A. A systematic review of the evidence of the burden of bipolar disorder in Europe. *Clinical Practice and Epidemiology in Mental Health*. 2009;5(1):3.

20. Byford S, Torgerson DJ, Raftery J. Cost of illness studies. 2000.
21. Rice DP. Cost of illness studies: what is good about them? *Injury Prevention*. 2000;6(3):177-9.
22. Prince M, Knapp M, Guerchet M, McCrone P, Prina M, Comas-Herrera A, et al. *Dementia UK: -overview*. 2014.
23. Jin H, McCrone P. Cost-of-illness studies for bipolar disorder: systematic review of international studies. *PharmacoEconomics*. 2015;33(4):341-53.
24. Grande I, Berk M, Birmaher B, Vieta E. Bipolar disorder. *The Lancet*. 2016;387(10027):1561-72.
25. Bowling A. *Research methods in health: investigating health and health services*: McGraw-Hill Education (UK); 2014.
26. NICE. *Guide to the methods of technology appraisal 2013*. London: National Institute for Health and Care Excellence. 2013.
27. Dolan P. Modeling valuations for EuroQol health states. *Medical care*. 1997;35(11):1095-108.
28. Knapp M, Mangalore R. The trouble with QALYs.... *Epidemiologia e psichiatria sociale*. 2007;16(04):289-93.
29. Brazier J. Is the EQ-5D fit for purpose in mental health? *The British Journal of Psychiatry*. 2010;197(5):348-9.
30. Brazier J, Connell J, Papaioannou D, Mukuria C, Mulhern B, Peasgood T, et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. 2014.
31. Longworth L, Yang Y, Young T, Hernandez Alva M, Mukuria C, Rowen D, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: systematic review, statistical modelling and survey. 2014.
32. Malhi GS, Bassett D, Boyce P, Bryant R, Fitzgerald PB, Fritz K, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. *Australian and New Zealand Journal of Psychiatry*. 2015;49(12):1087-206.
33. Yatham LN, Kennedy SH, Parikh SV, Schaffer A, Beaulieu S, Alda M, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) collaborative update of CANMAT guidelines for the management of patients with bipolar disorder: update 2013. *Bipolar disorders*. 2013;15(1):1-44.
34. Reilly S, Planner C, Gask L, Hann M, Knowles S, Druss B, et al. Collaborative care approaches for people with severe mental illness. *The Cochrane Library*. 2013.
35. Bopp JM, Miklowitz DJ, Goodwin GM, Stevens W, Rendell JM, Geddes JR. The longitudinal course of bipolar disorder as revealed through weekly text messaging: a feasibility study. *Bipolar disorders*. 2010;12(3):327-34.
36. Ortiz A, Grof P. Electronic monitoring of self-reported mood: the return of the subjective? *International Journal of Bipolar Disorders*. 2016;4(1):28.
37. Ciapala J, Mullins B, Forrest A, Molodynski A. Bridging the gap between research and clinical practice—a new initiative. *Progress in Neurology and Psychiatry*. 2016;20(4):9-11.
38. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The lancet*. 2013;380(9859):2197-223.
39. Anderson IM, Haddad PM, Scott J. Bipolar disorder. *Bmj*. 2012;345:e8508.
40. Müller-Oerlinghausen B, Berghöfer A, Bauer M. Bipolar disorder. *The Lancet*. 2002;359(9302):241-7.
41. Kerner B. Genetics of bipolar disorder. *Appl Clin Genet*. 2014;7:33-42.
42. Rosenblat JD, Cha DS, Mansur RB, McIntyre RS. Inflamed moods: a review of the interactions between inflammation and mood disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*. 2014;53:23-34.

43. Bhattacharya A, Derecki NC, Lovenberg TW, Drevets WC. Role of neuro-immunological factors in the pathophysiology of mood disorders. *Psychopharmacology*. 2016;233(9):1623-36.
44. Gonzalez R. The relationship between bipolar disorder and biological rhythms. *The Journal of clinical psychiatry*. 2014;75(4):323-31.
45. Merikangas KR, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Archives of General Psychiatry*. 2011;68(3):241-51.
46. Hirschfeld RM, Lewis L, Vornik LA. Perceptions and impact of bipolar disorder: how far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder. *The Journal of clinical psychiatry*. 2003;64(2):161-74.
47. Matza LS, Rajagopalan KS, Thompson CL, de Lissoyoy G. Misdiagnosed patients with bipolar disorder: comorbidities, treatment patterns, and direct treatment costs. *The Journal of clinical psychiatry*. 2005;66(11):1432-40.
48. McElroy SL, Altshuler LL, Suppes T, Keck Jr PE, Frye MA, Denicoff KD, et al. Axis I psychiatric comorbidity and its relationship to historical illness variables in 288 patients with bipolar disorder. *American Journal of Psychiatry*. 2001;158(3):420-6.
49. Cerullo MA, Strakowski SM. The prevalence and significance of substance use disorders in bipolar type I and II disorder. *Substance Abuse Treatment, Prevention, and Policy*. 2007;2(1):1.
50. Pavlova B, Perlis RH, Alda M, Uher R. Lifetime prevalence of anxiety disorders in people with bipolar disorder: a systematic review and meta-analysis. *The Lancet Psychiatry*. 2015;2(8):710-7.
51. Salloum IM, Thase ME. Impact of substance abuse on the course and treatment of bipolar disorder. *Bipolar disorders*. 2000;2(3p2):269-80.
52. Zimmerman M, Morgan TA. The relationship between borderline personality disorder and bipolar disorder. *Dialogues Clin Neurosci*. 2013;15(2):155-69.
53. Hu L-Y, Shen C-C, Hu Y-W, Chen M-H, Tsai C-F, Chiang H-L, et al. Hyperthyroidism and risk for bipolar disorders: a nationwide population-based study. *PloS one*. 2013;8(8):e73057.
54. Mansur RB, Brietzke E, McIntyre RS. Is there a "metabolic-mood syndrome"? A review of the relationship between obesity and mood disorders. *Neuroscience & Biobehavioral Reviews*. 2015;52:89-104.
55. Judd LL, Akiskal HS, Schettler PJ, et al. A prospective investigation of the natural history of the long-term weekly symptomatic status of bipolar II disorder. *Archives of General Psychiatry*. 2003;60(3):261-9.
56. Keck PE, Jr., McElroy SL, Strakowski SM, West SA, Sax KW, Hawkins JM, et al. 12-month outcome of patients with bipolar disorder following hospitalization for a manic or mixed episode. *The American journal of psychiatry*. 1998;155(5):646-52.
57. Gitlin MJ, Swendsen J, Heller TL, Hammen C. Relapse and impairment in bipolar disorder. *The American journal of psychiatry*. 1995;152(11):1635-40.
58. Judd LI, AHSSPJ, et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Archives of General Psychiatry*. 2002;59(6):530-7.
59. Goldberg JF, Harrow M. Consistency of remission and outcome in bipolar and unipolar mood disorders: a 10-year prospective follow-up. *Journal of affective disorders*. 2004;81(2):123-31.
60. Vázquez GH, Holtzman JN, Lolic M, Ketter TA, Baldessarini RJ. Recurrence rates in bipolar disorder: Systematic comparison of long-term prospective, naturalistic studies versus randomized controlled trials. *European Neuropsychopharmacology*. 2015;25(10):1501-12.
61. Prisciandaro JJ, Tolliver BK. Evidence for the continuous latent structure of mania and depression in out-patients with bipolar disorder: results from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *Psychological Medicine*. 2015;45(12):2595-603.
62. Harris EC, Barraclough B. Suicide as an outcome for mental disorders. A meta-analysis. *The British Journal of Psychiatry*. 1997;170(3):205-28.

63. Crump C, Sundquist K, Winkleby MA, Sundquist J. Comorbidities and mortality in bipolar disorder: A Swedish national cohort study. *JAMA Psychiatry*. 2013;70(9):931-9.
64. Roshanaei-Moghaddam B, Katon W. Premature mortality from general medical illnesses among persons with bipolar disorder: a review. *Psychiatr Serv*. 2009;60(2):147-56.
65. Kessing LV, Vradi E, McIntyre RS, Andersen PK. Causes of decreased life expectancy over the life span in bipolar disorder. *Journal of affective disorders*. 2015;180:142-7.
66. Post RM, Leverich GS, Altshuler LL, Frye MA, Suppes TM, Keck PE, et al. An overview of recent findings of the Stanley Foundation Bipolar Network (Part I). *Bipolar disorders*. 2003;5(5):310-9.
67. Hayes J, Miles J, Walters K, King M, Osborn D. A systematic review and meta-analysis of premature mortality in bipolar affective disorder. *Acta psychiatrica Scandinavica*. 2015;131(6):417-25.
68. Oud M, Mayo-Wilson E, Braidwood R, Schulte P, Jones SH, Morriss R, et al. Psychological interventions for adults with bipolar disorder: systematic review and meta-analysis. *The British Journal of Psychiatry*. 2016;208(3):213-22.
69. Bauer M, Grof P, Rasgon N, Glenn T, Alda M, Priebe S, et al. Mood charting and technology: new approach to monitoring patients with mood disorders. *Current Psychiatry Reviews*. 2006;2(4):423-9.
70. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. *Bmj*. 2002;324(7347):1193-4.
71. Anhøj J, Møldrup C. Feasibility of collecting diary data from asthma patients through mobile phones and SMS (short message service): response rate analysis and focus group evaluation from a pilot study. *Journal of Medical Internet Research*. 2004;6(4):e42.
72. Cho J-H, Lee H-C, Lim D-J, Kwon H-S, Yoon K-H. Mobile communication using a mobile phone with a glucometer for glucose control in Type 2 patients with diabetes: as effective as an Internet-based glucose monitoring system. *Journal of Telemedicine and Telecare*. 2009;15(2):77-82.
73. Naslund JA, Marsch LA, McHugo GJ, Bartels SJ. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. *Journal of Mental Health*. 2015;24(5):321-32.
74. Kleinman L, Lowin A, Flood E, Gandhi G, Edgell E, Revicki D. Costs of Bipolar Disorder. *PharmacoEconomics*. 2003;21(9):601-22.
75. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *The Lancet*. 2013;382(9904):1575-86.
76. Gore FM, Bloem PJN, Patton GC, Ferguson J, Joseph V, Coffey C, et al. Global burden of disease in young people aged 10-24 years: a systematic analysis. *The Lancet*. 2011;377(9783):2093-102.
77. Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*. 2015;386(9995):743-800.
78. Sanchez-Moreno J, Martinez-Aran A, Tabarés-Seisdedos R, Torrent C, Vieta E, Ayuso-Mateos JL. Functioning and Disability in Bipolar Disorder: An Extensive Review. *Psychotherapy and Psychosomatics*. 2009;78(5):285-97.
79. Reed C, Goetz I, Vieta E, Bassi M, Haro JM. Work impairment in bipolar disorder patients – results from a two-year observational study (EMBLEM). *European Psychiatry*. 2010;25(6):338-44.
80. Judd LL, Akiskal HS, Schettler PJ, Endicott J, Leon AC, Solomon DA, et al. Psychosocial disability in the course of bipolar I and II disorders: a prospective, comparative, longitudinal study. *Archives of general psychiatry*. 2005;62(12):1322-30.
81. Gitlin MJ, Mintz J, Sokolski K, Hammen C, Altshuler LL. Subsyndromal depressive symptoms after symptomatic recovery from mania are associated with delayed functional recovery. *The Journal of clinical psychiatry*. 2010;72(5):692-7.

82. Andreou C, Bozikas VP. The predictive significance of neurocognitive factors for functional outcome in bipolar disorder. *Current opinion in psychiatry*. 2013;26(1):54-9.
83. Larg A, Moss J. Cost-of-Illness Studies. *PharmacoEconomics*. 2011;29(8):653-71.
84. Arnold LM, Witzeman KA, Swank ML, McElroy SL, Keck Jr PE. Health-related quality of life using the SF-36 in patients with bipolar disorder compared with patients with chronic back pain and the general population. *Journal of affective disorders*. 2000;57(1):235-9.
85. Zhang H, Wisniewski SR, Bauer MS, Sachs GS, Thase ME, Investigators S-B. Comparisons of perceived quality of life across clinical states in bipolar disorder: data from the first 2000 Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) participants. *Comprehensive psychiatry*. 2006;47(3):161-8.
86. Bonnin C, Sanchez-Moreno J, Martinez-Aran A, Sole B, Reinares M, Rosa A, et al. Subthreshold symptoms in bipolar disorder: impact on neurocognition, quality of life and disability. *Journal of affective disorders*. 2012;136(3):650-9.
87. Vojta C, Kinosian B, Glick H, Altshuler L, Bauer MS. Self-reported quality of life across mood states in bipolar disorder. *Comprehensive psychiatry*. 2001;42(3):190-5.
88. Murray G, Michalak EE. The quality of life construct in bipolar disorder research and practice: past, present, and possible futures. *Bipolar disorders*. 2012;14(8):793-6.
89. Fagiolini A, Forgione R, Maccari M, Cuomo A, Morana B, Dell'Osso MC, et al. Prevalence, chronicity, burden and borders of bipolar disorder. *Journal of affective disorders*. 2013;148(2-3):161-9.
90. Babak Roshanaei-Moghaddam M, Katon W. Premature mortality from general medical illnesses among persons with bipolar disorder: a review. *Psychiatric Services*. 2015.
91. Health NCCfM. Bipolar Disorder: The Assessment and Management of Bipolar Disorder in Adults, Children and Young People in Primary and Secondary Care. Leicester UK: The British Psychological Society. 2014.
92. Knapp M, Lemmi V. The economic case for better mental health. 2014.
93. WHO. Economic aspects of the mental health system: Key messages to health planners and policy-makers. 2006.
94. Hodgson TA, Meiners MR. Cost-of-Illness Methodology: A Guide to Current Practices and Procedures. The Milbank Memorial Fund Quarterly Health and Society. 1982;60(3):429-62.
95. Samuelson P. *Economics*. Columbus, OH: McGraw-Hill; 1948.
96. Jefferson T, Demichelli V, Mugford M. Elementary economic evaluation in health care: BMJ Publications; 2000.
97. Segel JE. Cost-of-Illness Studies—A Primer Research Triangle Park, North Carolina: RTI-UNC Center of Excellence in Health Promotion Economics; 2006.
98. Chisholm D, Stanciole AE, Edejer TTT, Evans DB. economic impact of disease and injury: counting what matters. *BMJ*. 2010;340:583.
99. Tarricone R. Cost-of-illness analysis: What room in health economics? *Health Policy*. 2006;77(1):51-63.
100. Frick KD. Micro-costing quantity data collection methods. *Medical care*. 2009;47(7 Suppl 1):S76.
101. Drummond M. Cost-of-Illness Studies. *PharmacoEconomics*. 1992;2(1):1-4.
102. Greenberg PE. Cost of Illness: An Ongoing Battle Worth Fighting. *PharmacoEconomics*. 2014;32(12):1151-2.
103. Torrance GW, Drummond MF. *Methods for the economic evaluation of health care programmes*: Oxford university press; 2005.
104. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes*. 3rd ed. Oxford: Oxford University Press; 2005.
105. Shearer J, Byford S. The basics of economic evaluation in mental healthcare. *BJPsych Advances*. 2015;21(5):345-53.
106. Larg A, Moss JR. Cost-of-illness studies. *PharmacoEconomics*. 2011;29(8):653-71.
107. Pritchard C, Sculpher M. *Productivity costs: principles and practice in economic evaluation*: Office of Health Economics London; 2000.

108. Shearer J, McCrone P, Romeo R. Economic Evaluation of Mental Health Interventions: A Guide to Costing Approaches. *PharmacoEconomics*. 2016;1-14.
 109. Gray AM, Clarke PM, Wolstenholme JL, Wordsworth S. Applied methods of cost-effectiveness analysis in healthcare: OUP Oxford; 2010.
 110. Beecham J, Knapp M. Costing psychiatric interventions: Gaskell In *Measuring Mental Health Needs*. Edited by Thornicroft G. London; 2001.
 111. McDaid D, Byford S, Sefton T. Because it's worth it: a practical guide to conducting economic evaluations in the social welfare field: Joseph Rowntree Foundation; 2003.
 112. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health technology assessment*. 1998;2(14):i-iv, 1-74.
 113. Fayers P, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes: John Wiley & Sons; 2013.
 114. Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical care*. 1992;473-83.
 115. Brooks R, Group E. EuroQol: the current state of play. *Health policy*. 1996;37(1):53-72.
 116. WHO. WHOQOL-BREF: introduction, administration, scoring and generic version of the assessment: field trial version, December 1996. 1996.
 117. Dolan P, Gudex C, Kind P, Williams A. A social tariff for EuroQol: results from a UK general population survey: Centre for Health Economics University of York, UK; 1995.
 118. Michalak EE, Yatham LN, Lam RW. Quality of life in bipolar disorder: a review of the literature. *Health and Quality of Life Outcomes*. 2005;3(1):1.
 119. IsHak WW, Brown K, Aye SS, Kahloon M, Mobaraki S, Hanna R. Health-related quality of life in bipolar disorder. *Bipolar disorders*. 2012;14(1):6-18.
 120. Gray AM, Clarke PM, Wolstenholme JL, Wordsworth S. Applied Methods of Cost-effectiveness Analysis in Healthcare. Gray AM, Briggs A, editors. Oxford: Oxford University Press; 2011. 328 p.
 121. NICE. Guide to the Methods of Technology Appraisal. London: National Institute of Health and Clinical Excellence; 2013.
 122. Glick HA, Doshi JA, Sonnad SS, Polsky D. Economic Evaluation in Clinical Trials. Gray AM, Briggs A, editors. Oxford: Oxford University Press
2007. 256 p.
123. Mooney CZ, Duval RD, Duval R. Bootstrapping: A nonparametric approach to statistical inference: Sage; 1993.
 124. Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. *Health economics*. 2005;14(4):339-47.
 125. Dakin H, Wordsworth S. Cost-Minimisation Analysis Versus Cost-Effectiveness Analysis, Revisited. *Health economics*. 2013;22(1):22-34.
 126. McIntosh E. Applied methods of cost-benefit analysis in health care: Oxford University Press; 2010.
 127. Briggs A. Economic evaluation and clinical trials: size matters. *BMJ*. 2000;321(7273):1362-3.
 128. NICE. Guide to the methods of technology appraisal. National Institute for Health and Clinical Excellence; 2008.
 129. Briggs A, Sculpher M, Claxton K. Decision Modelling for Health Economic Evaluation. Gray AM, Briggs A, editors. Oxford: Oxford University Press; 2006. 256 p.
 130. Petrou S, Gray A. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. *BMJ*. 2011;342.
 131. Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models: a suggested framework and example of application. *PharmacoEconomics*. 2000;17(5):461-77.

132. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good Practice Guidelines for Decision-Analytic Modelling in Health Technology Assessment: A Review and Consolidation of Quality Assessment. *Pharmacoeconomics*. 2006;24(4):355-71.
133. Luyten J, Naci H, Knapp M. Economic evaluation of mental health interventions: an introduction to cost-utility analysis. *Evidence-based mental health*. 2016;19(2):49-53.
134. van der Loos ML, Mulder PG, Hartong EG, Blom MB, Vergouwen AC, de Keyser HJ, et al. Efficacy and safety of lamotrigine as add-on treatment to lithium in bipolar depression: a multicenter, double-blind, placebo-controlled trial. *The Journal of clinical psychiatry*. 2009;70(2):223-31.
135. Vieta E, Cruz N, Garcia-Campayo J, de Arce R, Manuel Crespo J, Valles V, et al. A double-blind, randomized, placebo-controlled prophylaxis trial of oxcarbazepine as adjunctive treatment to lithium in the long-term treatment of bipolar I and II disorder. *Int J Neuropsychopharmacol*. 2008;11(4):445-52.
136. Vieta E, Suppes T, Eggens I, Persson I, Paulsson B, Brecher M. Efficacy and safety of quetiapine in combination with lithium or divalproex for maintenance of patients with bipolar I disorder (international trial 126). *J Affect Disord*. 2008;109(3):251-63.
137. Keck PE, Jr., Calabrese JR, McIntyre RS, McQuade RD, Carson WH, Eudicone JM, et al. Aripiprazole monotherapy for maintenance therapy in bipolar I disorder: a 100-week, double-blind study versus placebo. *The Journal of clinical psychiatry*. 2007;68(10):1480-91.
138. Löthgren M. Economic evidence in affective disorders: a review. *The European Journal of Health Economics*. 2004;5(0):s12-s20.
139. Soares-Weiser K, Vergel YB, Beynon S, Dunn G, Barbieri M, Duffy S, et al. A systematic review and economic model of the clinical effectiveness and cost-effectiveness of interventions for preventing relapse in people with bipolar disorder. *Health technology assessment*. 2007;11(39):iii-119.
140. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLoS Med*. 2009;6(7):e1000100.
141. Shemilt I, Mugford M, Byford S, Drummond M, Eisenstein E, Knapp M, et al. Incorporating economics evidence. In: Higgins JP, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.0.1 ed: The Cochrane Collaboration; 2008.
142. CRD. Systematic reviews of economic evaluations. 2009. In: *Systematic Reviews [Internet]*. York: University of York. 3rd. CRD's guidance for undertaking reviews in health care; [199-218].
143. Rothwell PM. Commentary: External validity of results of randomized trials: disentangling a complex concept. *International journal of epidemiology*. 2010;39(1):94-6.
144. Revicki DA, Paramore LC, Sommerville KW, Swann AC, Zajecka JM, Depakote Comparator Study G. Divalproex sodium versus olanzapine in the treatment of acute mania in bipolar disorder: health-related quality of life and medical cost outcomes. *Journal of Clinical Psychiatry*. 2003;64(3):288-94.
145. Lam DH, McCrone P, Wright K, Kerr N. Cost-effectiveness of relapse-prevention cognitive therapy for bipolar disorder: 30-month study. *British Journal of Psychiatry*. 2005;186:500-6.
146. Revicki DA, Hirschfeld RMA, Ahearn EP, Weisler RH, Palmer C, Keck PE, Jr. Effectiveness and medical costs of divalproex versus lithium in the treatment of bipolar disorder: results of a naturalistic clinical trial. *Journal of affective disorders*. 2005;86(2-3):183-93.
147. Bauer MS, McBride L, Williford WO, Glick H, Kinosian B, Altshuler L, et al. Collaborative care for bipolar disorder: Part II. Impact on clinical outcome, function, and costs. *Psychiatric Services*. 2006;57(7):937-45.
148. Simon GE, Ludman EJ, Bauer MS, Unutzer J, Operskalski B. Long-term effectiveness and cost of a systematic care program for bipolar disorder. *Archives of General Psychiatry*. 2006;63(5):500-8.

149. Stevanovic D. Quality of Life Enjoyment and Satisfaction Questionnaire—short form for quality of life assessments in clinical practice: A psychometric study. *Journal of psychiatric and mental health nursing*. 2011;18(8):744-50.
150. Ekman M, Lindgren P, Miltenburger C, Meier G, Locklear JC, Chatterton ML. Cost effectiveness of quetiapine in patients with acute bipolar depression and in maintenance treatment after an acute depressive episode. *Pharmacoeconomics*. 2012;30(6):513-30.
151. Kasteng F, Eriksson J, Sennfalt K, Lindgren P. Metabolic effects and cost-effectiveness of aripiprazole versus olanzapine in schizophrenia and bipolar disorder. *Acta psychiatrica Scandinavica*. 2011;124(3):214-25.
152. Woodward TC, Tafesse E, Quon P, Lazarus A. Cost effectiveness of adjunctive quetiapine fumarate extended-release tablets with mood stabilizers in the maintenance treatment of bipolar I disorder (Provisional abstract). *Pharmacoeconomics* [Internet]. 2010; (9):[751-64 pp.]. Available from: <http://www.mrw.interscience.wiley.com/cochrane/cleed/articles/NHSEED-22010001667/frame.html>.
153. Woodward TC, Tafesse E, Quon P, Kim J, Lazarus A. Cost-effectiveness of quetiapine with lithium or divalproex for maintenance treatment of bipolar I disorder. *Journal of Medical Economics*. 2009;12(4):259-68.
154. Fajutrao L, Paulsson B, Liu S, Locklear J. Cost-effectiveness of quetiapine plus mood stabilizers compared with mood stabilizers alone in the maintenance therapy of bipolar I disorder: results of a Markov model analysis (Structured abstract). *Clinical Therapeutics* [Internet]. 2009; (Part 1):[1456-68 pp.]. Available from: <http://www.mrw.interscience.wiley.com/cochrane/cleed/articles/NHSEED-22009103261/frame.html>.
155. McKendrick J, Cerri KH, Lloyd A, D'Ausilio A, Dando S, Chinn C. Cost effectiveness of olanzapine in prevention of affective episodes in bipolar disorder in the United Kingdom (Structured abstract). *Journal of Psychopharmacology* [Internet]. 2007; (6):[588-96 pp.]. Available from: <http://www.mrw.interscience.wiley.com/cochrane/cleed/articles/NHSEED-22007002244/frame.html>.
156. Klok RM, Al Hadithy AFY, van Schayk NPJT, Antonisse AJJ, Caro JJ, Brouwers JRB, et al. Pharmacoeconomics of quetiapine for the management of acute mania in bipolar I disorder. *Expert Review of Pharmacoeconomics and Outcomes Research*. 2007;7(5):459-67.
157. NICE. Bipolar disorder: The management of bipolar disorder in adults, children and adolescents, in primary and secondary care. Leicester: National Institute for Health and Clinical Excellence; 2006.
158. Calvert NW, Burch SP, Fu AZ, Reeves P, Thompson TR. The cost-effectiveness of lamotrigine in the maintenance treatment of adults with bipolar I disorder (Structured abstract). *Journal of Managed Care Pharmacy* [Internet]. 2006; (4):[322-30 pp.]. Available from: <http://www.mrw.interscience.wiley.com/cochrane/cleed/articles/NHSEED-22006006514/frame.html>.
159. Keck PE, Jr., Nabulsi AA, Taylor JL, Henke CJ, et al. A pharmacoeconomic model of divalproex vs. lithium in the acute and prophylactic treatment of bipolar I disorder. *Journal of Clinical Psychiatry*. 1996;57(5):213-22.
160. Chisholm D, Saxena S. Cost effectiveness of strategies to combat neuropsychiatric conditions in sub-Saharan Africa and South East Asia: mathematical modelling study. *BMJ*. 2012;344:e609.
161. Chisholm D, van Ommeren M, Ayuso-Mateos J-L, Saxena S. Cost-effectiveness of clinical interventions for reducing the global burden of bipolar disorder. *British Journal of Psychiatry*. 2005;187:559-67.
162. Menzin J, Sussman M, Tafesse E, Duczakowski C, Neumann P, Friedman M. A model of the economic impact of a bipolar disorder screening program in primary care (Provisional abstract). *Journal of Clinical Psychiatry* [Internet]. 2009; (9):[1230-6 pp.]. Available from: <http://www.mrw.interscience.wiley.com/cochrane/cleed/articles/NHSEED-22009103895/frame.html>.

163. Edejer T-T, Baltussen R, Adam T, Hutubessy RCW, Acharya A, Evans DB, et al. Making choices in health: WHO guide to cost-effectiveness analysis. Geneva: WHO 2003.
164. Cooper N, Coyle D, Abrams K, Mugford M, Sutton A. Use of evidence in decision models: an appraisal of health technology assessments in the UK since 1997. *Journal of health services research & policy*. 2005;10(4):245-50.
165. Magalhaes P, Dean O, Bush A, Copolov D, Weisinger D, Malhi G, et al. Systemic illness moderates the impact of n-acetyl cysteine in bipolar disorder. *Progress in Neuro Psychopharmacology & Biological Psychiatry* Dec. 2011(Pagination):No Pagination Specified.
166. Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. *BMJ*. 2011;342:d1548.
167. Deakin B, Ferrier N, Holt RI, Millar H, Nutt DJ, Reynolds G, et al. The physical health challenges in patients with severe mental illness: cardiovascular and metabolic risks. *Journal of Psychopharmacology*. 2010;24 8(Suppl 1):1-8.
168. Johnson MD, Binford SH, Ansted CJ, Kennedy RS. Meeting highlights on mood disorders: The 3rd annual chair summit. *Health Outcomes Research in Medicine*. 2011;2(3):e141-e55.
169. Caro JJ, Möller J, Getsios D. Discrete event simulation: the preferred technique for health economic evaluations? *Value in Health*. 2010;13(8):1056-60.
170. Mavranouzouli I, Lokkerbol J. A Systematic Review and Critical Appraisal of Economic Evaluations of Pharmacological Interventions for People with Bipolar Disorder. *PharmacoEconomics*. 2016:1-26.
171. Mohiuddin S. A Systematic and Critical Review of Model-Based Economic Evaluations of Pharmacotherapeutics in Patients with Bipolar Disorder. *Applied health economics and health policy*. 2014;12(4):359-72.
172. Joslyn C, Hawes DJ, Hunt C, Mitchell PB. Is age of onset associated with severity, prognosis, and clinical features in bipolar disorder? A meta-analytic review. *Bipolar disorders*. 2016.
173. Fineberg NA, Haddad PM, Carpenter L, Gannon B, Sharpe R, Young AH, et al. The size, burden and cost of disorders of the brain in the UK. *Journal of Psychopharmacology*. 2013;27(9):761-70.
174. Gustavsson A, Svensson M, Jacobi F, Allgulander C, Alonso J, Beghi E, et al. Cost of disorders of the brain in Europe 2010. *European Neuropsychopharmacology*. 2011;21(10):718-79.
175. Dilsaver SC. An estimate of the minimum economic burden of bipolar I and II disorders in the United States: 2009. *Journal of affective disorders*. 2011;129(1–3):79-83.
176. Andlin-Sobocki P, Wittchen HU. Cost of affective disorders in Europe. *European Journal of Neurology*. 2005;12:34-8.
177. Das Gupta R, Guest JF. Annual cost of bipolar disorder to UK society. *The British Journal of Psychiatry*. 2002;180(3):227-33.
178. Young AH, Rigney U, Shaw S, Emmas C, Thompson JM. Annual cost of managing bipolar disorder to the UK healthcare system. *Journal of affective disorders*. 2011;133(3):450-6.
179. Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman self-rating mania scale. *Biological psychiatry*. 1997;42(10):948-55.
180. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biological psychiatry*. 2003;54(5):573-83.
181. DOH. National Schedule of Reference Costs. In: Health Do, editor. 2010-11
182. Committee JF. British national formulary: Pharmaceutical Press; 2013.
183. Curtis L. Unit costs of health and social care 2012: Personal Social Services Research Unit; 2012.
184. CMH. Mental health at work: developing the business case. London: Sainsbury Centre for Mental Health; 2007.

185. NICE. Managing long-term sickness absence and incapacity for work. London: National Institute for Health and Clinical Excellence; 2009.
186. Bovill D. Patterns of Pay: Results from the Annual Survey of Hours and Earnings, 1997 to 2012. London: Office for National Statistics; 2013.
187. Faria R, Weatherly H, van den Berg B. A review of approaches to measure and monetarily value informal care. Unit costs of health and social care 2012 [Internet]. 2012.
188. Spineli LM, Pandis N, Salanti G. Reporting and handling missing outcome data in mental health: a systematic review of Cochrane systematic reviews and meta-analyses. *Research synthesis methods*. 2015;6(2):175-87.
189. MacNeil Vroomen J, Eekhout I, Dijkgraaf MG, van Hout H, de Rooij SE, Heymans MW, et al. Multiple imputation strategies for zero-inflated cost data in economic evaluations: which method works best? *The European Journal of Health Economics*. 2016;17(8):939-50.
190. Royston P, White IR. Multiple imputation by chained equations (MICE): implementation in Stata. *Journal of Statistical Software*. 2011;45(4):1-20.
191. Twisk JW. Applied longitudinal data analysis for epidemiology: Cambridge University Press; 2013.
192. Raikou M, McGuire A. Estimating costs for economic evaluations. In: Jones A, editor. *The Elgar Companion to Health Economics*. 2 ed. Cheltenham, UK: Edward Elgar Publishing Limited; 2012. p. 463-72.
193. Szmaragd C, Clarke P, Steele F. Subject specific and population average models for binary longitudinal data: a tutorial. *Longitudinal and life course studies*. 2013;42(2).
194. Cui J, Qian G. Selection of working correlation structure and best model in GEE analyses of longitudinal data. *Communications in Statistics—Simulation and Computation®*. 2007;36(5):987-96.
195. Pan W. Akaike's information criterion in generalized estimating equations. *Biometrics*. 2001;57(1):120-5.
196. Williams R. Using the margins command to estimate and interpret adjusted predictions and marginal effects. *Stata Journal*. 12(2):308.
197. Dell'Aglio Jr. JC, Basso LA, Argimon IldL, Arteché A. Systematic review of the prevalence of bipolar disorder and bipolar spectrum disorders in population-based studies. *Trends in Psychiatry and Psychotherapy*. 2013;35:99-105.
198. Clemente AS, Diniz BS, Nicolato R, Kapczinski FP, Soares JC, Fírmio JO, et al. Bipolar disorder prevalence: a systematic review and meta-analysis of the literature. *Revista Brasileira de Psiquiatria*. 2015;37:155-61.
199. Merikangas KR, Jin R, He J, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Archives of General Psychiatry*. 2011;68(3):241-51.
200. Ferrari AJ, Baxter AJ, Whiteford HA. A systematic review of the global distribution and availability of prevalence data for bipolar disorder. *Journal of affective disorders*. 2011;134(1–3):1-13.
201. Pini S, de Queiroz V, Pagnin D, Pezawas L, Angst J, Cassano GB, et al. Prevalence and burden of bipolar disorders in European countries. *European Neuropsychopharmacology*. 2005;15(4):425-34.
202. Angst J. The emerging epidemiology of hypomania and bipolar II disorder. *Journal of affective disorders*. 1998;50(2–3):143-51.
203. Szádóczky E, Papp Z, Vitrai J, Ríhmer Z, Füredi J. The prevalence of major depressive and bipolar disorders in Hungary: Results from a national epidemiologic survey. *Journal of affective disorders*. 1998;50(2–3):153-62.
204. UN. United Nations Statistics Division. Composition of macro geographical (continental regions), geographical sub-regions, and selected economic and other groupings 2013 [Available from: <http://unstats.un.org/unsd/methods/m49/m49regin.htm>].
205. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled clinical trials*. 1986;7(3):177-88.
206. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-
analyses. *Bmj*. 2003;327(7414):557-60.

207. ONS. Mid-2002 to Mid-2010 Population Estimates revised following the 2011 Census. In: Statistics OfN, editor. Newport: Office for National Statistics; 2012.
208. HM-Treasury. Public Expenditure Statistical Analysis. London2012.
209. Modini M, Tan L, Brinchmann B, Wang M-J, Killackey E, Glozier N, et al. Supported employment for people with severe mental illness: systematic review and meta-analysis of the international evidence. *The British Journal of Psychiatry*. 2016.
210. Drake RE, Bond GR, Becker DR. Individual placement and support: an evidence-based approach to supported employment: Oxford University Press; 2012.
211. Bebbington P, Ramana R. The epidemiology of bipolar affective disorder. *Social Psychiatry and Psychiatric Epidemiology*. 1995;30(6):279-92.
212. Marwaha S, Durrani A, Singh S. Employment outcomes in people with bipolar disorder: a systematic review. *Acta psychiatrica Scandinavica*. 2013;128(3):179-93.
213. Lam DH, Watkins ER, Hayward P, Bright J, Wright K, Kerr N, et al. A randomized controlled study of cognitive therapy for relapse prevention for bipolar affective disorder: outcome of the first year. *Archives of General Psychiatry*. 2003;60(2):145-52.
214. Tafalla M, Salvador-Carulla L, Saiz-Ruiz J, Diez T, Cordero L. Pattern of healthcare resource utilization and direct costs associated with manic episodes in Spain. *BMC psychiatry*. 2010;10(1):31.
215. Ekman M, Granström O, Omérov S, Jacob J, Landén M. The societal cost of bipolar disorder in Sweden. *Social Psychiatry and Psychiatric Epidemiology*. 2013;48(10):1601-10.
216. Runge C, Grunze H. Annual costs of bipolar disorders in Germany. *Der Nervenarzt*. 2004;75(9):896-903.
217. Hakkaart-van Roijen L, Hoeijenbos M, Regeer E, Ten Have M, Nolen W, Veraart C, et al. The societal costs and quality of life of patients suffering from bipolar disorder in the Netherlands. *Acta psychiatrica Scandinavica*. 2004;110(5):383-92.
218. Perlis RH, Ostacher MJ, Patel JK, Marangell LB, Zhang H, Wisniewski SR, et al. Predictors of recurrence in bipolar disorder: primary outcomes from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *FOCUS: The Journal of Lifelong Learning in Psychiatry*. 2006;4(4):553-61.
219. Begley CE, Annegers JF, Swann AC, Lewis C, Coan S, Schnapp WB, et al. The lifetime cost of bipolar disorder in the US. *PharmacoEconomics*. 2001;19(5):483-95.
220. Atkinson M, Zibin S, Chuang H. Characterizing quality of life among patients with chronic mental illness: a critical examination of the self-report methodology. *American journal of psychiatry*. 1997;154(1):99-105.
221. Dickey B, Wagenaar H, Stewart A. Using health status measures with the seriously mentally ill in health services research. *Medical care*. 1996;34(2):112-6.
222. Fitzpatrick R, Gibbons E, Mackintosh A. An overview of patient-reported outcome measures for people with anxiety and depression
University of Oxford, Group P-rOM; 2009.
223. Hayes J, Prah P, Nazareth I, King M, Walters K, Petersen I, et al. Prescribing trends in bipolar disorder: cohort study in the United Kingdom THIN primary care database 1995–2009. *PLoS One*. 2011;6(12):e28725.
224. Hayes JF, Pitman A, Marston L, et al. Self-harm, unintentional injury, and suicide in bipolar disorder during maintenance mood stabilizer treatment: A uk population-based electronic health records study. *JAMA Psychiatry*. 2016;73(6):630-7.
225. Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychological methods*. 2002;7(2):147.
226. HM-Treasury. Spending Review London2010
227. Brazier J. Measuring and valuing mental health for use in economic evaluation. *Journal of health services research & policy*. 2008;13(suppl 3):70-5.
228. Black N. Patient reported outcome measures could help transform healthcare. *BMJ (Clinical research ed)*. 2013;346:f167.

229. Gilbody SM, House AO, Sheldon T. Outcomes measurement in psychiatry: A critical review of outcomes measurement in psychiatric research and practice: NHS Centre for Reviews and Dissemination; 2003.
230. McCrone P. Mental health economics: current methodological issues. *Epidemiology and Psychiatric Sciences*. 2011;20(03):239-43.
231. Payakachat N, Ali MM, Tilford JM. Can The EQ-5D Detect Meaningful Change? A Systematic Review. *Pharmacoeconomics*. 2015;33(11):1137-54.
232. Connell J, O'Cathain A, Brazier J. Measuring quality of life in mental health: Are we asking the right questions? *Social Science & Medicine*. 2014;120:12-20.
233. Mulhern B, Mukuria C, Barkham M, Knapp M, Byford S, Soeteman D, et al. Using generic preference-based measures in mental health: psychometric validity of the EQ-5D and SF-6D. *The British Journal of Psychiatry*. 2014;bjp. bp. 112.122283.
234. Trauer T. Outcome measurement in mental health: theory and practice: Cambridge University Press; 2010.
235. Longworth L, Rowen D. Mapping to obtain EQ-5D utility values for use in NICE health technology assessments. *Value in Health*. 2013;16(1):202-10.
236. Bilderbeck AC, Atkinson LZ, McMahon HC, Voysey M, Simon J, Price J, et al. Psychoeducation and online mood tracking for patients with bipolar disorder: A randomised controlled trial. *Journal of affective disorders*. 2016;205:245-51.
237. Brazier J, Longworth L. NICE DSU technical support document 8: an introduction to the measurement and valuation of health for NICE submissions. NICE Decision Support Unit, London. 2011.
238. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use: Oxford University Press, USA; 2014.
239. Feng Y, Devlin N, Herdman M. Assessing the health of the general population in England: how do the three-and five-level versions of EQ-5D compare? *Health and quality of life outcomes*. 2015;13(1):1.
240. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment*. 1994;6(4):284.
241. McDowell I. Measuring health: a guide to rating scales and questionnaires: Oxford university press; 2006.
242. Cohen J. Methods in psychology. A power primer *Psychological Bulletin*. 1992;112(1):155-9.
243. Cameron IM, Crawford JR, Cardy AH, du Toit SW, Lawton K, Hay S, et al. Psychometric properties of the Quick Inventory of Depressive Symptomatology (QIDS-SR) in UK primary care. *Journal of psychiatric research*. 2013;47(5):592-8.
244. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health and Quality of Life Outcomes*. 2006;4(1):1-5.
245. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of clinical epidemiology*. 2008;61(2):102-9.
246. Eisen SV, Ranganathan G, Seal P, Spiro III A. Measuring clinically meaningful change following mental health treatment. *The journal of behavioral health services & research*. 2007;34(3):272-89.
247. Sonntag M, Konnopka A, Leichsenring F, Salzer S, Beutel ME, Herpertz S, et al. Reliability, validity and responsiveness of the EQ-5D in assessing and valuing health status in patients with social phobia. *Health and quality of life outcomes*. 2013;11(1):1.
248. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Medical care*. 2003;41(5):582-92.
249. Cohen J. Statistical power analysis for the behavioral sciences. Vol. 2. Lawrence Earlbaum Associates, Hillsdale, NJ. 1988.

250. Petrou S, Rivero-Arias O, Dakin H, Longworth L, Oppe M, Froud R, et al. The MAPS reporting statement for studies mapping onto generic preference-based outcome measures: explanation and elaboration. *PharmacoEconomics*. 2015;33(10):993-1011.
251. Brown TA. *Confirmatory factor analysis for applied research*: Guilford Publications; 2015.
252. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. 2008.
253. Reise SP, Morizot J, Hays RD. The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*. 2007;16(1):19-31.
254. Rodriguez A, Reise SP, Haviland MG. Applying Bifactor Statistical Indices in the Evaluation of Psychological Measures. *Journal of personality assessment*. 2015:1-15.
255. Muthén LK, Muthén BO. *MPLUS (7). Computer software*) Los Angeles: Muthén & Muthén <http://statmodel.com>. 2012.
256. StataCorp. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP2015.
257. Szende A, Janssen B, Cabases J. *Self-reported population health: an international perspective based on EQ-5D*: Springer; 2014.
258. Peasgood T, Brazier J, Papaioannou D. A systematic review of the validity and responsiveness of EQ-5D and SF-6D for depression and anxiety. *HEDS Discussion paper 12/15*. 2012.
259. Copay AG, Subach BR, Glassman SD, Polly DW, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *The Spine Journal*. 2007;7(5):541-6.
260. Feng Y, Parkin D, Devlin NJ. Assessing the performance of the EQ-VAS in the NHS PROMs programme. *Quality of Life Research*. 2014;23(3):977-89.
261. Whynes D, McCahon R, Ravenscroft A, Hodgkinson V, Evley R, Hardman J. Responsiveness of the EQ-5D health-related quality-of-life instrument in assessing low back pain. *Value in Health*. 2013;16(1):124-32.
262. Tordrup D, Mossman J, Kanavos P. Responsiveness of the EQ-5D to clinical change: is the patient experience adequately represented? *International journal of technology assessment in health care*. 2014;30(01):10-9.
263. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PMM. On assessing responsiveness of health-related quality of life instruments: Guidelines for instrument evaluation. *Quality of Life Research*. 2003;12(4):349-62.
264. Subero MM, Berk L, Dodd S, Kulkarni J, De Castella A, Fitzgerald PB, et al. To a broader concept of remission: Rating the health-related quality of life in bipolar disorder. *Journal of affective disorders*. 2013.
265. Hayhurst H, Palmer S, Abbott R, Johnson T, Scott J. Measuring health-related quality of life in bipolar disorder: relationship of the EuroQol (EQ-5D) to condition-specific measures. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*. 2006;15(7):1271-80.
266. Saarni SI, Viertiö S, Perälä J, Koskinen S, Lönnqvist J, Suvisaari J. Quality of life of people with schizophrenia, bipolar disorder and other psychotic disorders. *The British Journal of Psychiatry*. 2010;197(5):386-94.
267. Allison PD, editor *Handling missing data by maximum likelihood*. SAS global forum; 2012.
268. Mulhern B, O’Gorman H, Rotherham N, Brazier J. Comparing the measurement equivalence of EQ-5D-5L across different modes of administration. *Health and quality of life outcomes*. 2015;13(1):1.
269. Keding A, Böhnke JR, Croudace TJ, Richmond SJ, MacPherson H. Validity of single item responses to short message service texts to monitor depression: an mHealth sub-study of the UK ACUDep trial. *BMC medical research methodology*. 2015;15(1):1.
270. Goodwin G, Haddad P, Ferrier I, Aronson J, Barnes T, Cipriani A, et al. Evidence-based guidelines for treating bipolar disorder: Revised third edition recommendations from the

- British Association for Psychopharmacology. *Journal of Psychopharmacology*. 2016;30(6):495-553.
271. Vieta E, Pacchiarotti I, Valentí M, Berk M, Scott J, Colom F. A critical update on psychological interventions for bipolar disorders. *Current psychiatry reports*. 2009;11(6):494-502.
 272. Bourne C, Bilderbeck A, Drennan R, Atkinson L, Price J, Geddes JR, et al. Verbal learning impairment in euthymic bipolar disorder: BDI v BDII. *Journal of affective disorders*. 2015;182:95-100.
 273. Bower P, Gilbody S, Richards D, Fletcher J, Sutton A. Collaborative care for depression in primary care. *The British Journal of Psychiatry*. 2006;189(6):484-93.
 274. Miklowitz DJ, Goodwin GM, Bauer MS, Geddes JR. Common and specific elements of psychosocial treatments for bipolar disorder: a survey of clinicians participating in randomized trials. *Journal of psychiatric practice*. 2008;14(2):77.
 275. Parker GB, Graham RK, Fletcher K, Futeran SM, Friend P. The impact of being newly diagnosed with a bipolar disorder and the short-term outcome of disorder-specific management. *Bipolar disorders*. 2014;16(2):172-9.
 276. Reilly TJ, MacGillivray SA, Reid IC, Cameron IM. Psychometric properties of the 16-item Quick Inventory of Depressive Symptomatology: A systematic review and meta-analysis. *Journal of psychiatric research*. 2015;60:132-40.
 277. Lako IM, Wigman JT, Klaassen RM, Slooff CJ, Taxis K, Bartels-Velthuis AA. Psychometric properties of the self-report version of the Quick Inventory of Depressive Symptoms (QIDS-SR16) questionnaire in patients with schizophrenia. *BMC Psychiatry*. 2014;14(1):247.
 278. Miller CJ, Johnson SL, Eisner L. Assessment tools for adult bipolar disorder. *Clinical Psychology: Science and Practice*. 2009;16(2):188-201.
 279. Sen A. *The idea of justice*: Harvard University Press; 2011.
 280. Lorgelly PK. Choice of Outcome Measure in an Economic Evaluation: A Potential Role for the Capability Approach. *PharmacoEconomics*. 2015;33(8):849-55.
 281. Simon J, Anand P, Gray A, Rugkåsa J, Yeeles K, Burns T. Operationalising the capability approach for outcome measurement in mental health research. *Social Science & Medicine*. 2013;98:187-96.
 282. DOH. Department of Health website 2013 [Available from: <https://www.gov.uk/government/publications/nhs-reference-costs-2012-to-2013>].
 283. Curtis L. Unit Costs of Health and Social Care Kent: Personal Social Services Research Unit; 2013. Available from: <http://www.pssru.ac.uk>
 284. BNF. British National Formulary (BNF) 67 ed2014.
 285. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health economics*. 2005;14(5):487-96.
 286. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. *Clinical trials*. 2012;9(4):396-407.
 287. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *Bmj*. 1999;319(7211):670-4.
 288. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ: British Medical Journal*. 2009;338.
 289. Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *PharmacoEconomics*. 2014;32(12):1157-70.
 290. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in medicine*. 2011;30(4):377-99.
 291. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health economics*. 2004;13(5):461-75.

292. Bruffaerts C, Verardi V, Vermandele C. A generalized boxplot for skewed and heavy-tailed distributions. *Statistics & probability letters*. 2014;95:110-7.
293. StataCorp. *Stata Statistical Software: Release 14*. College Station, Texas: StataCorp LP; 2015.
294. Pari AA, Simon J, Wolstenholme J, Geddes J, Goodwin G, editors. *Economic burden of bipolar disorder in the United Kingdom: a patient-level exploratory analysis*. BIPOLAR DISORDERS; 2015: WILEY-BLACKWELL 111 RIVER ST, HOBOKEN 07030-5774, NJ USA.
295. Goodwin G, Psychopharmacology CGotBAf. Evidence-based guidelines for treating bipolar disorder: revised second edition—recommendations from the British Association for Psychopharmacology. *Journal of Psychopharmacology*. 2009;23(4):346-88.
296. Kendall T, Morriss R, Mayo-Wilson E, Marcus E. Assessment and management of bipolar disorder: summary of updated NICE guidance. *BMJ*. 2014;349:g5673.
297. Canady VA. FDA approves first generic Abilify to treat mental illnesses. *Mental Health Weekly*. 2015;25(18):6-7.
298. Ramsey SD, Willke RJ, Glick H, Reed SD, Augustovski F, Jonsson B, et al. Cost-effectiveness analysis alongside clinical trials II—an ISPOR Good Research Practices Task Force report. *Value in Health*. 2015;18(2):161-72.
299. Bond K, Anderson IM. Psychoeducation for relapse prevention in bipolar disorder: a systematic review of efficacy in randomized controlled trials. *Bipolar disorders*. 2015;17(4):349-62.
300. Hamer RM, Simpson PM. Last observation carried forward versus mixed models in the analysis of psychiatric clinical trials. *American Journal of Psychiatry*. 2009.
301. Freedland KE, Mohr DC, Davidson KW, Schwartz JE. Usual and unusual care: existing practice control groups in randomized controlled trials of behavioral interventions. *Psychosomatic medicine*. 2011;73(4):323.
302. Morriss R, Lobban F, Riste L, Davies L, Holland F, Long R, et al. Clinical effectiveness and acceptability of structured group psychoeducation versus optimised unstructured peer support for patients with remitted bipolar disorder (PARADES): a pragmatic, multicentre, observer-blind, randomised controlled superiority trial. *The Lancet Psychiatry*. 2016;3(11):1029-38.
303. Kallestad H, Wullum E, Scott J, Stiles TC, Morken G. The long-term outcomes of an effectiveness trial of group versus individual psychoeducation for bipolar disorders. *Journal of affective disorders*. 2016;202:32-8.
304. Thornicroft G, Patel V. *Global Mental Health Trials*: OUP Oxford; 2014.
305. Parker AG, Hetrick SE, Jorm AF, Yung AR, McGorry PD, Mackinnon A, et al. The effectiveness of simple psychological and exercise interventions for high prevalence mental health problems in young people: a factorial randomised controlled trial. *Trials*. 2011;12(1):1.
306. Jensen PS, Weersing R, Hoagwood KE, Goldman E. What is the evidence for evidence-based treatments? A hard look at our soft underbelly. *Mental health services research*. 2005;7(1):53-74.
307. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. *PLoS med*. 2013;10(1):e1001362.
308. Muralidharan A, Miklowitz DJ, Craighead WE, Nathan P, Gorman J. *Psychosocial Treatments for Bipolar Disorder. A Guide to Treatments That Work*. 2015:307.
309. Miklowitz DJ, Scott J. Psychosocial treatments for bipolar disorder: cost-effectiveness, mediating mechanisms, and future directions. *Bipolar Disorders*. 2009;11 Suppl 2:110-22.
310. Pari A, Ahmed A, Simon J, Wolstenholme J, Geddes JR, Goodwin GM. Economic evaluations in bipolar disorder: a systematic review and critical appraisal. *Bipolar disorders*. 2014;16(6):557-82.
311. Marshall A, Billingham LJ, Bryan S. Can we afford to ignore missing data in cost-effectiveness analyses? *The European Journal of Health Economics*. 2009;10(1):1-3.
312. Mihaylova B, Briggs A, O'Hagan A, Thompson SG. Review of statistical methods for analysing healthcare resources and costs. *Health economics*. 2011;20(8):897-916.

313. Hollingworth W, McKell-Redwood D, Hampson L, Metcalfe C. Cost–utility analysis conducted alongside randomized controlled trials: Are economic end points considered in sample size calculations and does it matter? *Clinical Trials*. 2012;1740774512465358.
314. McCrone P, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the Price: The Cost of Mental Health Care in England to 20262008.
315. Vöhringer PA, Perlis RH. Discriminating Between Bipolar Disorder and Major Depressive Disorder. *Psychiatric Clinics of North America*. 2016;39(1):1-10.
316. Saunders KEA, Bilderbeck AC, Price J, Goodwin GM. Distinguishing bipolar disorder from borderline personality disorder: A study of current clinical practice. *European Psychiatry*. 30(8):965-74.
317. Saunders K, Geddes JR. The management of bipolar disorder. *British journal of hospital medicine (London, England: 2005)*. 2016;77(3):175.
318. Cheniaux E, Filgueiras A, da Silva RdA, Silveira LAS, Nunes ALS, Landeira-Fernandez J. Increased energy/activity, not mood changes, is the core feature of mania. *Journal of affective disorders*. 2014;152:256-61.
319. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of life research*. 2011;20(10):1727-36.
320. Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. *Quality of Life Research*. 2012;21(1):167-76.
321. Tacconelli E. Systematic reviews: CRD's guidance for undertaking reviews in health care. *The Lancet Infectious Diseases*. 2010;10(4):226.
322. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. 2004.
323. Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H. Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. *British Journal of Psychiatry*. 2004;184:526-33.
324. Clark RE, Teague GB, Ricketts SK, Bush PW, Xie H, McGuire TG, et al. Cost-effectiveness of assertive community treatment versus standard case management for persons with co-occurring severe mental illness and substance use disorders. *Health Services Research*. 1998;33(5 Pt 1):1285-308.
325. Caro JJ, Huybrechts KF, Xenakis JG, O'Brien JA, Rajagopalan K, Lee K. Budgetary impact of treating acute bipolar mania in hospitalized patients with quetiapine: an economic analysis of clinical trials (Structured abstract). *Current Medical Research and Opinion* [Internet]. 2006; (11):[2233-42 pp.]. Available from: <http://www.mrw.interscience.wiley.com/cochrane/cleed/articles/NHSEED-22006002421/frame.html>.
326. Conney J, Kaston B. Pharmacoeconomic and health outcome comparison of lithium and divalproex in a VA geriatric nursing home population: influence of drug-related morbidity on total cost of treatment. *American Journal of Managed Care*. 1999;5(2):197-204.
327. Dean BB, Gerner D, Gerner RH. A systematic review evaluating health-related quality of life, work impairment, and healthcare costs and utilization in bipolar disorder. *Current Medical Research and Opinion*. 2004;20(2):139-54.
328. DSI. A health technology assessment to investigate the consequences of implementing specialised outpatient clinics for treatment of patients with severe affective disorders Copenhagen: Danish Institute for Health Services Research; 2005.
329. Fleurence RL, Chatterton ML, Dixon JM, Rajagopalan K. Economic outcomes associated with atypical antipsychotics in bipolar disorder: a systematic review. *Primary Care Companion to the Journal of Clinical Psychiatry*. 2007;9(6):419-28.
330. Goldberg JF, Altshuler LL, Frye M, Keck PE, Jr. Pharmacoeconomics of lithium and divalproex sodium in the treatment of bipolar disorder. 1999:141-67.
331. Greenhalgh J, Knight C, Hind D, Beverley C, Walters S. Clinical and cost-effectiveness of electroconvulsive therapy for depressive illness, schizophrenia, catatonia and mania: systematic reviews and economic modelling studies (Structured abstract). *Health technology*

assessment [Internet]. 2005:[1 p.]. Available from: <http://www.mrw.interscience.wiley.com/cochrane/clhta/articles/HTA-32005000183/frame.html>.

332. Hakkaart-van Roijen L, Hoeijenbos MB, Regeer EJ, Ten Have M, Nolen WA, Veraart CPWM, et al. The societal costs and quality of life of patients suffering from bipolar disorder in the Netherlands. *Acta psychiatrica Scandinavica*. 2004;110(5):383-92.
333. Hong J, Reed C, Novick D, Haro JM, Aguado J. Clinical and economic consequences of medication non-adherence in the treatment of patients with a manic/mixed episode of bipolar disorder: results from the European Mania in Bipolar Longitudinal Evaluation of Medication (EMBLEM) study. *Psychiatry Research*. 2011;190(1):110-4.
334. Keith S. Use of long-acting risperidone in psychiatric disorders: focus on efficacy, safety and cost-effectiveness. *Expert Review of Neurotherapeutics*. 2009;9(1):9-31.
335. Li SC, Aggarwal SK. Estimation of resource utilisation difference between lithium and valproate treatment groups from the VALID study. *Journal of Medical Economics*. 2011;14(3):350-6.
336. Namjoshi MA, Rajamannar G, Jacobs T, Sanger TM, Risser R, Tohen MF, et al. Economic, clinical, and quality-of-life outcomes associated with olanzapine treatment in mania. Results from a randomized controlled trial. *Journal of affective disorders*. 2002;69(1-3):109-18.
337. Myhr G, Payne K. Cost-effectiveness of cognitive-behavioural therapy for mental disorders: Implications for public health care funding policy in Canada. *Canadian Journal of Psychiatry*. 2006;51(10):662-70.
338. Reifman A, Wyatt RJ. Lithium: a brake in the rising cost of mental illness. *Archives of General Psychiatry*. 1980;37(4):385-8.
339. Scott J, Colom F, Popova E, Benabarre A, Cruz N, Valenti M, et al. Long-term mental health resource utilization and cost of care following group psychoeducation or unstructured group support for bipolar disorders: a cost-benefit analysis. *Journal of Clinical Psychiatry*. 2009;70(3):378-86.
340. Scott J, Leboyer M. Consequences of delayed diagnosis of bipolar disorders. *Encephale*. 2011;37 Suppl 3:S173-5.
341. Shaya FT, Wang J, M, Lee S, Levine A, Varghese R. Clinical and economic evaluation of olanzapine-fluoxetine HCl combination in the treatment of bipolar depression: a managed care approach. *Journal of Medical Economics*. 2007;10(1):67-77.
342. Simons WR, Krishnan AA. The economic value of lamotrigine as a mood stabilizer: a U.S. managed care perspective. *Managed Care Interface*. 2004;17(12):44-9.
343. Steffens DC, Krishnan KR. Decision model for the acute treatment of mania. *Depression and anxiety*. 1996;4(6):289-93.
344. Virtanen M, Kawachi I, Oksanen T, Salo P, Tuisku K, Pulkki-Raback L, et al. Socio-economic differences in long-term psychiatric work disability: prospective cohort study of onset, recovery and recurrence. *Occupational & Environmental Medicine*. 2011;68(11):791-8.
345. Zhang Y. Cost-saving effects of olanzapine as long-term treatment for bipolar disorder. *The Journal of Mental Health Policy & Economics*. 2008;11(3):135-46.

Appendix 1: Search strategies used to identify references which summarise the economic evaluations of management strategies in Bipolar disorder in Chapter 4

Search strategy Ovid, Embase March 2012

Search History (67 searches)

#	Searches	Results
67	remove duplicates from 66	3193
66	64 not 65	3227
65	limit 64 to animals	25
64	61 not 63	3252
63	limit 61 to (conference abstract or "conference review")	519
62	limit 61 to animals	36
61	limit 60 to yr="1980 -Current"	3771
60	59 not 57	3811
59	58 not 56	3826
58	5 and 52	4053
57	"animal studies"/	1496063
56	or/53-55	1618881
55	note.pt.	467455
54	editorial.pt.	392963
53	letter.pt.	758463
52	or/9,51	1431539
51	or/10-50	575140
50	(sf20 or sf 20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty).tw.	269
49	(sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).tw.	24

48	(sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).tw.	2629
47	(sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).tw.	1267
46	(sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).tw.	16405
45	(resource adj (allocat\$ or use\$ or utilit\$)).tw.	8984
44	(hye or hyes).tw.	61
43	tto.tw.	616
42	willingness to pay.tw.	2242
41	disutili\$.tw.	232
40	rosser.tw.	79
39	simulation model\$.tw.	5670
38	markov.tw.	10614
37	monte carlo.tw.	21385
36	decision tree\$.sh.	4247
35	decision analy\$.tw.	5209
34	qaly\$.tw.	5552
33	(valu\$ adj2 quality).tw.	1313
32	(qtwist or q twist).tw.	137
31	time trade.tw.	745
30	standard gamble.tw.	649
29	visual analog\$.tw.	32348
28	(euroqol or euro qol).tw.	2118
27	(weight\$ adj3 preference\$).tw.	362
26	utility value\$.tw.	736
25	((disability or quality) adj adjusted).tw.	7330
24	((life adj2 year\$) or health year equivalent\$).tw.	29383
23	(health adj utilit\$).tw.	1064

22	hui\$1.tw.	1145
21	(quality adj3 well\$).tw.	4769
20	qwb.tw.	163
19	(eq5d or (eq adj 5d)).tw.	2995
18	(qald\$ or qale\$ or qtime\$).tw.	117
17	(well being or wellbeing).tw.	43199
16	(health adj2 stat\$).tw.	55035
15	((adjusted adj2 life) or qaly\$).tw.	8667
14	(quality adj2 life).tw.	156730
13	(daly or qol or hql or hqol or hrqol or hr ql or hrql).tw.	31025
12	wellbeing/	24241
11	statistical model/	80052
10	exp "quality of life"/ or "value of life"/ or quality adjusted life year/	285604
9	or/6-8	1030942
8	(budget\$ or cost\$ or econom\$ or expenditure\$ or financ\$ or fiscal\$ or funding or pharmaco-economic\$ or price or prices or pricing or (value adj3 money) or (burden adj3 (disease\$ or illness\$))).tw.	599813
7	resource allocation.sh. or "deductibles and coinsurance"/	59955
6	(budget\$ or cost\$ or economic\$ or expenditure\$ or fee\$1 or fees\$ or financ\$ or health resource\$ or money or pharmaco-economic\$ or socioeconomic\$).hw.	713732
5	or/1-4	50413
4	(hypomania or mania or manic psychosis).sh. or (hypomania\$ or mania\$ or manic\$).ti,ab.	21120
3	(bipolar depression or bipolar disorder or bipolar i disorder or bipolar ii disorder or cyclothymia or rapid cycling bipolar disorder).sh.	29398
2	((cyclothymi\$ or rapid or ultradian) adj5 cycl\$) or RCBD).tw.	4905
1	((bipolar or bi?polar or bi polar) adj5 (disorder\$ or depress\$)).tw.	22869

Search strategy HTA and NHS EED

Search: BD	
Interface: Cochrane	Databases: HTA NHS EED
Filters: N/A	Date Range: Inception to date [10.03.2012]
Hits: HTA #16; NHS EED #112	
Notes: simplified strategy used for the purposes of search	

Current Search History

ID	Search	Hits
#1	<u>MeSH descriptor Bipolar Disorder explode all trees</u>	1366
#2	<u>(bipolar or bi?polar or "bi polar") near/5 (disorder* or depress*):ti or (bipolar or bi?polar or "bi polar") near/5 (disorder* or depress*):ab</u>	1613
#3	<u>((cyclothymi* or rapid or ultradian) near/5 cycl*) or rcbd:ti or ((cyclothymi* or rapid or ultradian) near/5 cycl*) or rcbd:ab</u>	236
#4	<u>(hypomania* or mania* or manic*):ti or (hypomania* or mania* or manic*):ab</u>	1257
#5	<u>(#1 OR #2 OR #3 OR #4)</u>	2779
#6	<u>(#1 OR #2 OR #3 OR #4)</u>	23
#7	<u>(#1 OR #2 OR #3 OR #4)</u>	22

Search strategy, OVID Medline indexed from 1980 to Mar 2012

Search History (66 searches)

#	Searches	Results
66	remove duplicates from 65	2073
65	63 not 64	2099
64	limit 63 to animals	69
63	limit 62 to yr="1980 -Current"	2168
62	61 not 60	2234
61	5 and 52	2286
60	or/53-59	1321280
59	interview.pt.	22445
58	lectures.pt.	5413
57	comment.pt.	491320
56	clinical conference.pt.	6169
55	news.pt.	143598
54	editorial.pt.	301207
53	letter.pt.	750880
52	or/9,51	1060010
51	or/10-50	374700
50	(sf20 or sf 20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty).tw.	312
49	(sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).tw.	19
48	(sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).tw.	1976
47	(sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).tw.	1182
46	(sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).tw.	12724
45	(resource adj (allocat\$ or use\$ or utilit\$)).tw.	7446

44	(hye or hyes).tw.	51
43	tto.tw.	478
42	willingness to pay.tw.	1697
41	disutili\$.tw.	170
40	rosser.tw.	69
39	simulation model\$.tw.	4772
38	markov.tw.	9459
37	monte carlo.tw.	23631
36	decision tree\$.sh.	7849
35	decision analy\$.tw.	4319
34	qaly\$.tw.	3822
33	(valu\$ adj2 quality).tw.	1060
32	(qtwist or q twist).tw.	81
31	time trade.tw.	626
30	standard gamble.tw.	589
29	visual analog\$.tw.	25722
28	(euroqol or euro qol).tw.	1605
27	(weight\$ adj3 preference\$).tw.	296
26	utility value\$.tw.	525
25	((disability or quality) adj adjusted).tw.	5900
24	((life adj2 year\$) or health year equivalent\$).tw.	24695
23	(health adj utilit\$).tw.	837
22	hui\$1.tw.	922
21	(quality adj3 well\$).tw.	3716
20	qwb.tw.	150
19	(eq5d or (eq adj 5d)).tw.	1971
18	(qald\$ or qale\$ or qtime\$).tw.	89
17	(well being or wellbeing).tw.	35806

16	(health adj2 stat\$).tw.	48203
15	((adjusted adj2 life) or qaly\$).tw.	6576
14	(quality adj2 life).tw.	117149
13	(daly or qol or hql or hqol or hrqol or hr ql or hrql).tw.	21848
12	wellbeing/	0
11	statistical model/	55493
10	exp "quality of life"/ or "value of life"/ or quality adjusted life year/	105008
9	or/6-8	749900
8	(budget\$ or cost\$ or econom\$ or expenditure\$ or financ\$ or fiscal\$ or funding or pharmaco-economic\$ or price or prices or pricing or (value adj3 money) or (burden adj3 (disease\$ or illness\$))).tw.	504817
7	resource allocation.sh. or "deductibles and coinsurance"/	7971
6	(budget\$ or cost\$ or economic\$ or expenditure\$ or fee\$1 or fees\$ or financ\$ or health resource\$ or money or pharmaco-economic\$ or socio-economic\$).hw.	392796
5	or/1-4	39427
4	(hypomania or mania or manic psychosis).sh. or (hypomania\$ or mania\$ or manic\$).ti,ab.	13779
3	(bipolar depression or bipolar disorder or bipolar i disorder or bipolar ii disorder or cyclothymia or rapid cycling bipolar disorder).sh.	27860
2	((cyclothymi\$ or rapid or ultradian) adj5 cycl\$) or RCBD).tw.	4316
1	((bipolar or bi?polar or bi polar) adj5 (disorder\$ or depress\$)).tw.	17210

Search strategy PubMed Medline

Search	Query	Items found
#47	Search (#45) NOT #38	150
#46	Search (#44) NOT #38	2126
#45	Search (((((((((((((((((((((((((((((((statistical) AND model))) OR (visual analogue scale[Text Word])) OR (standard gamble[Text Word])) OR (decision analysis[Text Word])) OR (decision tree[MeSH Terms])) OR ((eq5d[Text Word]) OR eq-5d[Text Word])) OR (((hui[Text Word]) OR "health utility index"[Text Word]) OR "utility value"[Text Word])) OR ((euroqol[Text Word]) OR "euro qol"[Text Word])) OR (((((((sf20[Text Word]) OR shortform 20[Text Word]) OR short form 20[Text Word]) OR sf twenty[Text Word]) OR sftwenty[Text Word]) OR shortform twenty[Text Word]) OR short form twenty[Text Word]) OR sf 20[Text Word])) OR (((((((sf16[Text Word]) OR shortform 16[Text Word]) OR short form 16[Text Word]) OR sf sixteen[Text Word]) OR sfsixteen[Text Word]) OR shortform sixteen[Text Word]) OR short form sixteen[Text Word]) OR sf 16[Text Word])) OR (((((((sf12[Text Word]) OR shortform 12[Text Word]) OR short form 12[Text Word]) OR sf twelve[Text Word]) OR sftwelve[Text Word]) OR shortform twelve[Text Word]) OR short form twelve[Text Word]) OR sf 12[Text Word])) OR (((((((sf6[Text Word]) OR sf 6[Text Word]) OR short form 6[Text Word]) OR shortform 6[Text Word]) OR sf six[Text Word]) OR sfsix[Text Word]) OR shortform six[Text Word]) OR short form six[Text Word])) OR (((((((sf36[Text Word]) OR sf-36[Text Word]) OR Shortform 36[Text Word]) OR short form thirty six[Text Word]) OR sf thirtysix[Text Word]) OR sf thirty six[Text Word]) OR shortform thirtysix[Text Word]) OR shortform thirty six[Text Word]) OR short form thirtysix[Text Word]) OR sf 36[Text Word]) OR short form 36[Text Word])) OR (willingness to pay[Text Word])) OR (disutility)) OR (markov[Text Word])) OR (monte carlo[Text Word])) OR ("time trade off")) OR (((Quality-adjusted life expectancy) OR disability adjusted life year) OR health adjusted life expectancy) OR quality of life) OR health related quality of life)) OR (qale) OR (((((daly) OR qol) OR hql) OR hqol) OR hrqol) OR hrql)) OR (wellbeing)) OR (((quality of life) OR value of life) OR quality adjusted life year))) OR (((((((pricing) OR fiscal) OR value for money)) OR (((burden of illness) OR burden of disease) OR burden)) OR ((resource allocation) OR deductibles and co-insurance)) OR (((((((economics) OR fees) OR finance) OR money) OR pharmacoeconomic) OR socioeconomic) OR expenditure) OR cost) OR budget)))) AND (((((((("bipolar disorder")) OR (Bipolar Disorder)) OR (cyclothymia)) OR (mania)) OR (bipolar depression)) AND ("1980/01/01"[PDAT] : "2012/04/01"[PDAT]) NOT "animals"[MeSH Terms:noexp] NOT MEDLINE[sb])	151
#44	Search (((((((((((((((((((((((((((((((statistical) AND model))) OR (visual analogue scale[Text Word])) OR (standard gamble[Text Word])) OR (decision analysis[Text Word])) OR (decision tree[MeSH Terms])) OR ((eq5d[Text Word]) OR eq-5d[Text Word])) OR (((hui[Text Word]) OR "health utility index"[Text Word]) OR "utility value"[Text Word])) OR ((euroqol[Text Word]) OR "euro qol"[Text Word])) OR (((((((sf20[Text Word]) OR	2176

health adjusted life expectancy) OR quality of life) OR health related quality of life)) OR (qale)) OR ((((((daly) OR qol) OR hql) OR hqol) OR hrqol) OR hrql)) OR (wellbeing)) OR (((quality of life) OR value of life) OR quality adjusted life year))) OR (((((((pricing) OR fiscal) OR value for money)) OR (((burden of illness) OR burden of disease) OR burden)) OR ((resource allocation) OR deductibles and co-insurance)) OR (((((((((economics) OR fees) OR finance) OR money) OR pharmacoeconomic) OR socioeconomic) OR expenditure) OR cost) OR budget)))) AND (((((((("bipolar disorder") OR (Bipolar Disorder)) OR (cyclothymia)) OR (mania)) OR (bipolar depression)) AND (("1980/01/01"[PDat] : "2012/04/01"[PDat])) Limits: Animals

[#42](#) Search ((((((((((((((((((((((((((((((((((statistical) AND model))) OR (visual analogue scale[Text Word])) OR (standard gamble[Text Word])) OR (decision analysis[Text Word])) OR (decision tree[MeSH Terms])) OR ((eq5d[Text Word] OR eq-5d[Text Word])) OR (((hui[Text Word]) OR "health utility index"[Text Word]) OR "utility value"[Text Word])) OR ((euroqol[Text Word]) OR "euro qol"[Text Word])) OR (((((((sf20[Text Word]) OR shortform 20[Text Word]) OR short form 20[Text Word]) OR sf twenty[Text Word]) OR sftwenty[Text Word]) OR shortform twenty[Text Word]) OR short form twenty[Text Word]) OR sf 20[Text Word])) OR (((((((sf16[Text Word]) OR shortform 16[Text Word]) OR short form 16[Text Word]) OR sf sixteen[Text Word]) OR sfsixteen[Text Word]) OR shortform sixteen[Text Word]) OR short form sixteen[Text Word]) OR sf 16[Text Word])) OR (((((((sf12[Text Word]) OR shortform 12[Text Word]) OR short form 12[Text Word]) OR sf twelve[Text Word]) OR sftwelve[Text Word]) OR shortform twelve[Text Word]) OR short form twelve[Text Word]) OR sf 12[Text Word])) OR (((((((sf6[Text Word]) OR sf 6[Text Word]) OR short form 6[Text Word]) OR shortform 6[Text Word]) OR sf six[Text Word]) OR sfsix[Text Word]) OR shortform six[Text Word]) OR short form six[Text Word])) OR (((((((sf36[Text Word]) OR sf-36[Text Word]) OR Shortform 36[Text Word]) OR short form thirty six[Text Word]) OR sf thirtysix[Text Word]) OR sf thirty six[Text Word]) OR shortform thirtysix[Text Word]) OR shortform thirty six[Text Word]) OR short form thirtysix[Text Word]) OR sf 36[Text Word]) OR short form 36[Text Word])) OR (willingness to pay[Text Word])) OR (disutility)) OR (markov[Text Word])) OR (monte carlo[Text Word])) OR ("time trade off")) OR (((Quality-adjusted life expectancy) OR disability adjusted life year) OR health adjusted life expectancy) OR quality of life) OR health related quality of life)) OR (qale)) OR (((((daly) OR qol) OR hql) OR hqol) OR hrqol) OR hrql)) OR (wellbeing)) OR (((quality of life) OR value of life) OR quality adjusted life year))) OR (((((((pricing) OR fiscal) OR value for money)) OR (((burden of illness) OR burden of disease) OR burden)) OR ((resource allocation) OR deductibles and co-insurance)) OR (((((((economics) OR fees) OR finance) OR money) OR pharmacoeconomic) OR socioeconomic) OR expenditure) OR cost) OR budget)))) AND (((((((("bipolar disorder") OR (Bipolar Disorder)) OR (cyclothymia)) OR (mania)) OR (bipolar depression)) AND (("1980/01/01"[PDat] : "2012/04/01"[PDat])))

#41 Search ((((((((((((((((((((((statistical) AND model)) OR (visual analogue scale[Text Word])) OR (standard gamble[Text Word])) OR (decision analysis[Text Word])) OR (decision tree[MeSH Terms])) OR ((eq5d[Text Word]) OR eq-5d[Text Word])) OR (((hui[Text Word]) OR "health utility index"[Text Word]) OR "utility value"[Text Word])) OR ((euroqol[Text Word]) OR "euro qol"[Text Word])) OR (((((((sf20[Text Word]) OR shortform 20[Text Word]) OR short form 20[Text Word]) OR sf twenty[Text Word]) OR sftwenty[Text Word]) OR shortform twenty[Text Word]) OR short form twenty[Text Word]) OR sf 20[Text Word])) OR (((((((sf16[Text Word]) OR shortform 16[Text Word])

OR short form 16[Text Word]) OR sf sixteen[Text Word]) OR sfsixteen[Text Word]) OR shortform sixteen[Text Word]) OR short form sixteen[Text Word]) OR sf 16[Text Word])) OR (((((((sf12[Text Word]) OR shortform 12[Text Word]) OR short form 12[Text Word]) OR sf twelve[Text Word]) OR sftwelve[Text Word]) OR shortform twelve[Text Word]) OR short form twelve[Text Word]) OR sf 12[Text Word])) OR (((((((sf6[Text Word]) OR sf 6[Text Word]) OR short form 6[Text Word]) OR shortform 6[Text Word]) OR sf six[Text Word]) OR sfsix[Text Word]) OR shortform six[Text Word]) OR short form six[Text Word])) OR (((((((sf36[Text Word]) OR sf-36[Text Word]) OR Shortform 36[Text Word]) OR short form thirty six[Text Word]) OR sf thirtysix[Text Word]) OR sf thirty six[Text Word]) OR shortform thirtysix[Text Word]) OR shortform thirty six[Text Word]) OR short form thirtysix[Text Word]) OR sf 36[Text Word]) OR short form 36[Text Word])) OR (willingness to pay[Text Word]) OR (disutility)) OR (markov[Text Word])) OR (monte carlo[Text Word])) OR ("time trade off")) OR (((((Quality-adjusted life expectancy) OR disability adjusted life year) OR health adjusted life expectancy) OR quality of life) OR health related quality of life)) OR (qale)) OR (((((daly) OR qol) OR hql) OR hqol) OR hrqol) OR hrql)) OR (wellbeing)) OR (((quality of life) OR value of life) OR quality adjusted life year))) OR (((((((pricing) OR fiscal) OR value for money)) OR (((burden of illness) OR burden of disease) OR burden)) OR ((resource allocation) OR deductibles and co-insurance)) OR (((((((economics) OR fees) OR finance) OR money) OR pharmacoeconomic) OR socioeconomic) OR expenditure) OR cost) OR budget)))) AND (((("bipolar disorder")) OR (Bipolar Disorder)) OR (cyclothymia)) OR (mania)) OR (bipolar depression))

#39 Search (((((((((((((((((((((((statistical) AND model)) OR (visual analogue scale[Text Word])) OR (standard gamble[Text Word])) OR (decision analysis[Text Word])) OR (decision tree[MeSH Terms])) OR ((eq5d[Text Word]) OR eq-5d[Text Word])) OR (((hui[Text Word]) OR "health utility index"[Text Word]) OR "utility value"[Text Word])) OR ((euroqol[Text Word]) OR "euro qol"[Text Word])) OR (((((((sf20[Text Word]) OR shortform 20[Text Word]) OR short form 20[Text Word]) OR sf twenty[Text Word]) OR sftwenty[Text Word]) OR shortform twenty[Text Word]) OR short form twenty[Text Word]) OR sf 20[Text Word])) OR (((((((sf16[Text Word]) OR shortform 16[Text Word]) OR short form 16[Text Word]) OR sf sixteen[Text Word]) OR sfsixteen[Text Word]) OR shortform sixteen[Text Word]) OR short form sixteen[Text Word]) OR sf 16[Text Word])) OR (((((((sf12[Text Word]) OR shortform 12[Text Word]) OR short form 12[Text Word]) OR sf twelve[Text Word]) OR sftwelve[Text Word]) OR shortform twelve[Text Word]) OR short form twelve[Text Word]) OR sf 12[Text Word])) OR (((((((sf6[Text Word]) OR sf 6[Text Word]) OR short form 6[Text Word]) OR shortform 6[Text Word]) OR sf six[Text Word]) OR sfsix[Text Word]) OR shortform six[Text Word]) OR short form six[Text Word])) OR (((((((sf36[Text Word]) OR sf-36[Text Word]) OR Shortform 36[Text Word]) OR short form thirty six[Text Word]) OR sf thirtysix[Text Word]) OR sf thirty six[Text Word]) OR shortform thirtysix[Text Word]) OR shortform thirty six[Text Word]) OR short form thirtysix[Text Word]) OR sf 36[Text Word]) OR short form 36[Text Word])) OR (willingness to pay[Text Word]) OR (disutility)) OR (markov[Text Word])) OR (monte carlo[Text Word])) OR ("time trade off")) OR (((((Quality-adjusted life expectancy) OR disability adjusted life year) OR health adjusted life expectancy) OR quality of life) OR health related quality of life)) OR (qale)) OR (((((daly) OR qol) OR hql) OR hqol) OR hrqol) OR hrql)) OR (wellbeing)) OR (((quality of life) OR value of life) OR quality adjusted life year))) OR (((((((pricing) OR fiscal) OR value for money)) OR (((burden of illness) OR burden of disease) OR burden))

[1138501](#)

	OR ((resource allocation) OR deductibles and co-insurance)) OR (((((((((economics) OR fees) OR finance) OR money) OR pharmacoeconomic) OR socioeconomic) OR expenditure) OR cost) OR budget))	
#38	Search (((note[Publication Type]) OR letter[Publication Type]) OR editorial[Publication Type]) OR news[Publication Type]) OR "conference abstract"[Publication Type]	1196493
#37	Search (((((((((((((((((((((statistical) AND model)) OR (visual analogue scale[Text Word])) OR (standard gamble[Text Word])) OR (decision analysis[Text Word])) OR (decision tree[MeSH Terms])) OR ((eq5d[Text Word]) OR eq-5d[Text Word])) OR (((hui[Text Word]) OR "health utility index"[Text Word]) OR "utility value"[Text Word])) OR ((euroqol[Text Word]) OR "euro qol"[Text Word])) OR (((((((sf20[Text Word]) OR shortform 20[Text Word]) OR short form 20[Text Word]) OR sf twenty[Text Word]) OR sftwenty[Text Word]) OR shortform twenty[Text Word]) OR short form twenty[Text Word]) OR sf 20[Text Word])) OR (((((((sf16[Text Word]) OR shortform 16[Text Word]) OR short form 16[Text Word]) OR sf sixteen[Text Word]) OR sfsixteen[Text Word]) OR shortform sixteen[Text Word]) OR short form sixteen[Text Word]) OR sf 16[Text Word])) OR (((((((sf12[Text Word]) OR shortform 12[Text Word]) OR short form 12[Text Word]) OR sf twelve[Text Word]) OR sftwelve[Text Word]) OR shortform twelve[Text Word]) OR short form twelve[Text Word]) OR sf 12[Text Word])) OR (((((((sf6[Text Word]) OR sf 6[Text Word]) OR short form 6[Text Word]) OR shortform 6[Text Word]) OR sf six[Text Word]) OR sfsix[Text Word]) OR shortform six[Text Word]) OR short form six[Text Word])) OR (((((((sf36[Text Word]) OR sf-36[Text Word]) OR Shortform 36[Text Word]) OR short form thirty six[Text Word]) OR sf thirtysix[Text Word]) OR sf thirty six[Text Word]) OR shortform thirtysix[Text Word]) OR shortform thirty six[Text Word]) OR short form thirtysix[Text Word]) OR sf 36[Text Word]) OR short form 36[Text Word])) OR (willingness to pay[Text Word])) OR (disutility)) OR (markov[Text Word])) OR (monte carlo[Text Word])) OR ("time trade off")) OR (((((Quality-adjusted life expectancy) OR disability adjusted life year) OR health adjusted life expectancy) OR quality of life) OR health related quality of life)) OR (qale)) OR (((((daly) OR qol) OR hql) OR hqol) OR hrqol) OR (wellbeing)) OR (((quality of life) OR value of life) OR quality adjusted life year)	330035
#36	Search (statistical) AND model	60470
#35	Search visual analogue scale[Text Word]	10189
#34	Search standard gamble[Text Word]	595
#33	Search decision analysis[Text Word]	3069
#32	Search decision tree[MeSH Terms]	7820
#31	Search (eq5d[Text Word]) OR eq-5d[Text Word]	2038
#30	Search ((hui[Text Word]) OR "health utility index"[Text Word]) OR "utility value"[Text Word]	787
#29	Search (euroqol[Text Word]) OR "euro qol"[Text Word]	1632

#28	Search ((((((sf20[Text Word]) OR shortform 20[Text Word]) OR short form 20[Text Word]) OR sf twenty[Text Word]) OR sftwenty[Text Word]) OR shortform twenty[Text Word]) OR short form twenty[Text Word]) OR sf 20[Text Word]	306
#27	Search ((((((sf16[Text Word]) OR shortform 16[Text Word]) OR short form 16[Text Word]) OR sf sixteen[Text Word]) OR sfsixteen[Text Word]) OR shortform sixteen[Text Word]) OR short form sixteen[Text Word]) OR sf 16[Text Word]	15
#26	Search ((((((sf12[Text Word]) OR shortform 12[Text Word]) OR short form 12[Text Word]) OR sf twelve[Text Word]) OR sftwelve[Text Word]) OR shortform twelve[Text Word]) OR short form twelve[Text Word]) OR sf 12[Text Word]	2027
#25	Search ((((((sf6[Text Word]) OR sf 6[Text Word]) OR short form 6[Text Word]) OR shortform 6[Text Word]) OR sf six[Text Word]) OR sfsix[Text Word]) OR shortform six[Text Word]) OR short form six[Text Word]	1244
#24	Search ((((((((((sf36[Text Word]) OR sf-36[Text Word]) OR Shortform 36[Text Word]) OR short form thirty six[Text Word]) OR sf thirtysix[Text Word]) OR sf thirty six[Text Word]) OR shortform thirtysix[Text Word]) OR shortform thirty six[Text Word]) OR short form thirtysix[Text Word]) OR sf 36[Text Word]) OR short form 36[Text Word]	13034
#21	Search willingness to pay[Text Word]	1748
#20	Search disutility	145
#19	Search markov[Text Word]	12117
#18	Search monte carlo[Text Word]	30494
#17	Search "time trade off"	640
#16	Search (((Quality-adjusted life expectancy) OR disability adjusted life year) OR health adjusted life expectancy) OR quality of life) OR health related quality of life	176825
#15	Search qale	68
#14	Search ((((((daly) OR qol) OR hql) OR hqol) OR hrqol) OR hrql	29904
#13	Search wellbeing	4284
#12	Search ((((((pricing) OR fiscal) OR value for money)) OR (((burden of illness) OR burden of disease) OR burden)) OR ((resource allocation) OR deductibles and co-insurance)) OR (((((((economics) OR fees) OR finance) OR money) OR pharmacoeconomic) OR socioeconomic) OR expenditure) OR cost) OR budget)	857421
#10	Search ((pricing) OR fiscal) OR value for money	171972
#9	Search ((burden of illness) OR burden of disease) OR burden	93647
#7	Search (resource allocation) OR deductibles and co-insurance	17719
#6	Search (((((((economics) OR fees) OR finance) OR money) OR pharmacoeconomic) OR socioeconomic) OR expenditure) OR cost) OR budget	792196

#5	Search (((("bipolar disorder")) OR (Bipolar Disorder)) OR (cyclothymia)) OR (mania) OR (bipolar depression)	34742
#4	Search bipolar depression	33199
#3	Search mania	34200
#2	Search Bipolar Disorder	32698

Search strategy OVID, Psychinfo

Search History (61 searches)

#	Searches	Results
61	remove duplicates from 60	1466
60	58 not 59	1468
59	limit 58 to animal	16
58	limit 57 to yr="1980 -Current"	1484
57	53 not 56	1573
56	54 or 55	34053
55	Editorial.dt.	22516
54	Letter.dt.	11537
53	5 and 52	1620
52	or/9,51	275164
51	or/10-50	102611
50	(sf20 or sf 20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty).tw.	48
49	(sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).tw.	3
48	(sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).tw.	559
47	(sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).tw.	12

46	(sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).tw.	2951
45	(resource adj (allocat\$ or use\$ or utilit\$)).tw.	2609
44	(hye or hyes).tw.	22
43	tto.tw.	88
42	willingness to pay.tw.	750
41	disutili\$.tw.	101
40	rosser.tw.	58
39	simulation model\$.tw.	946
38	markov.tw.	1908
37	monte carlo.tw.	2571
36	decision tree\$.sh.	0
35	decision analy\$.tw.	769
34	qaly\$.tw.	474
33	(valu\$ adj2 quality).tw.	382
32	(qtwist or q twist).tw.	12
31	time trade.tw.	140
30	standard gamble.tw.	151
29	visual analog\$.tw.	3980
28	(euroqol or euro qol).tw.	362
27	(weight\$ adj3 preference\$).tw.	218
26	utility value\$.tw.	157
25	((disability or quality) adj adjusted).tw.	737
24	((life adj2 year\$) or health year equivalent\$).tw.	4759
23	(health adj utilit\$).tw.	245
22	hui\$1.tw.	384
21	(quality adj3 well\$).tw.	1432
20	qwb.tw.	78

19	(eq5d or (eq adj 5d)).tw.	534
18	(qald\$ or qale\$ or qtime\$).tw.	6
17	(well being or wellbeing).tw.	40755
16	(health adj2 stat\$).tw.	15769
15	((adjusted adj2 life) or qaly\$).tw.	838
14	(quality adj2 life).tw.	32531
13	(daly or qol or hql or hqol or hrqol or hr ql or hrql).tw.	6987
12	wellbeing/	0
11	statistical model/	0
10	exp "quality of life"/ or "value of life"/ or quality adjusted life year/	21918
9	or/6-8	188362
8	(budget\$ or cost\$ or econom\$ or expenditure\$ or financ\$ or fiscal\$ or funding or pharmaco-economic\$ or price or prices or pricing or (value adj3 money) or (burden adj3 (disease\$ or illness\$))).tw.	166532
7	resource allocation.sh. or "deductibles and coinsurance"/	1832
6	(budget\$ or cost\$ or economic\$ or expenditure\$ or fee\$1 or fees\$ or financ\$ or health resource\$ or money or pharmaco-economic\$ or socioeconomic\$).hw.	57367
5	or/1-4	30511
4	(hypomania or mania or manic psychosis).sh. or (hypomania\$ or mania\$ or manic\$).ti,ab.	15944
3	(bipolar depression or bipolar disorder or bipolar i disorder or bipolar ii disorder or cyclothymia or rapid cycling bipolar disorder).sh.	16470
2	((cyclothymi\$ or rapid or ultradian) adj5 cycl\$) or RCBD).tw.	2182
1	((bipolar or bi?polar or bi polar) adj5 (disorder\$ or depress\$)).tw.	18989

Search strategy EMBSCO CINAHL

# ▼	Searches
61	You will need to Edit search and then limit to publication date
60	S59 NOT S57
59	S58 NOT S56
58	S5 AND S52
57	(MH "Animal Studies")
53	(ZT "commentary") or (ZT "editorial") or (ZT "letter") or (ZT "conference abstract")
52	S9 OR S51
51	S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49 OR S50
50	TI (sf20 or sf 20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty) OR AB (sf20 or sf 20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty)
49	TI (sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen) OR AB (sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen)
48	TI (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve) OR AB (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve)
47	TI (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six) OR AB (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six)
46	TI (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six) OR AB (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six)
45	TI (resource allocate* OR resource use* OR resource utility*) OR AB (resource allocate* OR resource use* OR resource utility*)
44	TI (hye or hyes) OR AB (hye or hyes)
43	TI (tto) OR AB (tto)
42	TI (willingness to pay) OR AB (willingness to pay)
41	TI (disutili*) OR AB (disutili*)

40	TI (rosser) OR AB (rosser)
39	TI (simulation model*) OR AB (simulation model*)
38	TI (markov) OR AB (markov)
37	TI (monte carlo) OR AB (monte carlo)
36	TI (decision tree*) OR AB (decision tree*)
35	TI (decision analy*) OR AB (decision analy*)
34	TI (qaly*) OR AB (qaly*)
33	TI (valu* n2 quality) OR AB (valu* n2 quality)
32	TI (qtwist or q twist) OR AB (qtwist or q twist)
31	TI (time trade) OR AB (time trade)
30	TI (standard gamble) OR AB (standard gamble*)
29	TI (visual analog*) OR AB (visual analog*)
28	TI (euroqol or euro qol) OR AB (euroqol or euro qol)
27	TI (weight* n3 preference*) OR AB (weight* n3 preference*)
26	TI (utility value*) OR AB (utility value*)
25	TI (disability adjusted OR quality adjusted) OR AB (disability adjusted OR quality adjusted)
24	TI ((life n2 year*) or health year equivalent*) OR AB ((life n2 year*) or health year equivalent*)
23	TI (health utilit*) OR AB (health utility*)
22	TI (hui*) OR AB (hui*)
21	TI (quality n3 well*) OR AB (quality n3 well*)
20	TI (qwb) OR AB (qwb)
19	TI (eq5d or eq 5d) OR AB (eq5d or eq 5d)
18	TI (qald* or qale* or qtime*) OR AB (qald* or qale* or qtime*)
17	TI (well being or wellbeing) OR AB (well being or wellbeing)
16	TI (health n2 stat*) OR AB (health n2 stat*)
15	TI ((adjusted n2 life) or qaly*) OR AB ((adjusted n2 life) or qaly*)
14	TI (quality n2 life) OR AB (quality n2 life)
13	TI (daly or qol or hql or hqol or hrqol or hr ql or hrql) OR AB (daly or qol or hql or hqol or hrqol or hr ql or hrql)
12	(MH "Psychological Well-Being")
11	(MH "Models, Statistical")
10	(MH "Quality of Life") OR (MH "Quality of Working Life") OR (MH "Quality-Adjusted Life Years")
9	S6 OR S7 OR S8

8	TI (budget* or cost* or econom* or expenditure* or financ* or fiscal* or funding or pharmaco-economic* or price or prices or pricing or (value n3 money) or (burden n3 (disease* or illness*))) OR AB (budget* or cost* or econom* or expenditure* or financ* or fiscal* or funding or pharmaco-economic* or price or prices or pricing or (value n3 money) or (burden n3 (disease* or illness*)))
7	(MH "Resource Allocation+")
6	TI (budget* or cost* or economic* or expenditure* or fee*1 or fees* or financ* or health resource* or money or pharmaco-economic* or socio-economic*) OR AB (budget* or cost* or economic* or expenditure* or fee*1 or fees* or financ* or health resource* or money or pharmaco-economic* or socio-economic*)
5	S1 OR S2 OR S3 OR S4
4	TI (hypomania or mania or manic psychosis) OR AB (hypomania or mania or manic psychosis) OR MW (hypomania or mania or manic psychosis)
3	MW (bipolar depression or bipolar disorder or bipolar i disorder or bipolar ii disorder or cyclothymia or rapid cycling bipolar disorder)
2	TI (((cyclothymi* or rapid or ultradian) n5 cycl*) or RCBD) OR AB (((cyclothymi* or rapid or ultradian) n5 cycl*) or RCBD)
1	TI ((bipolar or bi?polar or bi polar) n5 (disorder* or depress*)) OR AB ((bipolar or bi?polar or bi polar) n5 (disorder* or depress*))

Appendix 2: Data extraction template for collecting data on studies included in Chapter 4

DATA EXTRACTION FORM

Reviewer:

Date of Review:

Authors:

Publication Date:

Title:

Country:

Language:

Objective:

Economic study design:

CEA ☐

CBA ☐

CUA ☐

CMA ☐

CCA ☐

Source of data for effect size measure(s):

Meta-analysis

☐ RCT

☐

Quasi experimental study

☐

Controlled study with concurrent controls

☐

Controlled study with historical controls

☐

Mirror image (before-after) study

☐

Observational studies

☐

Retrospective databases

☐

Primary outcome measure(s) (please list):

Interventions compared (please describe):

Treatment:

Comparator:

Setting (please describe):

Patient population characteristics (please describe):

Patient sample:

Perspective of analysis:

Societal ☐ Patient and patient family ☐
Health care system ☐ Health care provider ☐
Individual clinician ☐ Insurer ☐
Other: _____

Time frame of analysis: _____

Cost data:

Primary ☐ Secondary ☐

If secondary please specify: _____

Costs included:

Direct medical	Direct non-medical	Productivity losses
direct treatment ☐	social care ☐	income forgone due to illness ☐
in-patient ☐	social benefits ☐	income forgone due to death ☐
out-patient ☐	travel costs ☐	income forgone by caregiver ☐
day care ☐	caregiver out-of-pocket ☐	
community health care ☐	criminal justice ☐	
medication ☐	training of staff ☐	
side effect costs ☐		
or		
staff ☐		
medication ☐		
labs/diagnostic ☐		
overhead ☐		
capital equipment ☐		
real estate ☐		

Others: _____

Currency: _____ **Year of costing:** _____

Was discounting used?

Yes, for benefits and costs ☐ Yes, but only for costs ☐ No ☐

Rate: _____

QoL estimates derived from:

Patients ☐	Experts ☐
General population ☐	Care givers ☐

Using:

Generic measures ☐ Direct elicitation ☐

Generic measures:

Rosser Index	☐	Health Utilities Index (HUI)	☐
EQ-5D	☐	Quality of Well Being (QWB)	☐
15D	☐	SF-36	☐
SF-12	☐	SF-6	☐

Converted into utilities?

Yes	☐	No	☐
-----	--------------------------	----	--------------------------

Using:

Direct elicitation:

Standard Gamble	☐	VAS	☐
Time trade-off	☐	Person trade-off	☐

Utility values combined with survival to form QALYs?

Yes	☐	No	☐
-----	--------------------------	----	--------------------------

Result(s):

Validity checklist Used:**CHEC-LIST:****score (Yes/NA/All):**

...../...../.....

Risk of bias:

⇒ High	⇒ Low
--------	-------

Comments, limitations of the study:

Appendix 3: Centre for review and dissemination (CRD) checklist for assessing economic evaluations(104, 321)*

Study Design:

- 1 The research question is stated.
- 2 The economic importance of the research question is stated.
- 3 The viewpoint(s) of the analysis is (are) clearly stated and justified.
- 4 The rationale for choosing alternative programmes or interventions compared is stated.
- 5 The alternatives being compared are clearly described.
- 6 The form of the economic evaluation used is stated.
- 7 The choice of form of economic evaluation is justified in relation to the questions addressed.

Data Collection:

- 8 The source(s) of effectiveness estimates used is(are) stated.
- 9 Details of the design and results of effectiveness study are given (if based on a single study).
- 10 Details of the methods of synthesis or meta-analysis of estimates are given (if based on a
- 11 The primary outcome measures(s) for the economic evaluation are clearly stated.
- 12 Methods to value benefits are stated.
- 13 Details of the subjects from whom valuations were obtained were given.
- 14 Productivity changes (if included) are reported separately.
- 15 The relevance of productivity changes to the study question is discussed.
- 16 Quantities of resource use are reported separately from their unit costs.
- 17 Methods for the estimation of quantities and unit costs are described.
- 18 Currency and price data are recorded.
- 19 Details of currency of price adjustments for inflation or currency conversion are given.
- 20 Details of any model used are given.
- 21 The choice of model used and the key parameters on which it is based are justified.

Analysis and Interpretation of Results:

- 22 Time horizon of costs and benefits is stated.
- 23 The discount rate(s) is (are) justified.
- 24 The choice of discount rate(s) is (are) justified.
- 25 An explanation is given if costs and benefits are not discounted.
- 26 Details of statistical tests and confidence intervals are given for stochastic data.
- 27 The approach to sensitivity analysis is given.
- 28 The choice of variables for sensitivity analysis is justified.
- 29 The ranges over which the variables are varied are justified.
- 30 Relevant alternatives are compared.
- 31 Incremental analysis is reported.
- 32 Major outcomes are presented in a disaggregated as well as aggregated form.
- 33 The answer to the study question is given.
- 34 Conclusions follow from the data reported.
- 35 Conclusions are accompanied by the appropriate caveats.
- 36 Were generalisability issues addressed.

*based on Drummond's check-list for assessing economic evaluations

Appendix 4: Review of guidelines for good practice in decision-analytic modelling in health technology assessment; Philips Checklist (322)

Structure (S)							
S1	Rationale for structure	S4	Structural assumptions	S7	Time horizon		
S2	Statement of scope/ The	S5	Strategies/ comparators	S8	Disease states / pathways		
S3	Statement of decision	S6	Model type	S9	Cycle length		
S1	Is there a clear statement of the decision problem?						
S1	Is the objective of the evaluation and model specified and consistent with the stated decision?						
S1	Is the primary decision maker specified?						
S2	Is the perspective of the model stated clearly?						
S2	Are the model inputs consistent with the stated perspective?						
S2	Has the scope of the model been stated and justified?						
S2	Are the outcomes of the model consistent with the perspective, scope and overall objective of the						
S3	Has the evidence regarding the model structure been described?						
S3	Is the structure of the model consistent with a coherent theory of the health condition under						
S3	Have any competing theories regarding model structure been considered?						
S3	Are the sources of data used to develop the model specified?						
S3	Are the causal relationships described by the model structure justified appropriately?						
S4	Are the structural assumptions transparent and justified?						
S4	Are the structural assumptions reasonable given overall objective, perspective and scope of the						
S5	Is there a clear definition of the options under evaluation?						
S5	Have all feasible and practical options been evaluated?						
S5	Is there justification for the exclusion of feasible options?						
S6	Is the chosen model type appropriate given the decision problem and specified causal relationships						
S7	Is the time horizon of the model sufficient to reflect all important differences between options?						
S7	Is the time horizon of the model, and the duration of treatment and treatment effect described and						
S7	Has a lifetime horizon been used? If not, has a shorter time horizon been justified?						
S8	Do disease states (state transition model) or pathways (decision tree model) reflect underlying						
S9	Is the cycle length defined and justified in terms of the natural history of disease?						
Consistency (C)							
C1 Internal consistency			C2 External consistency				
C	Is there evidence that the mathematical logic of the model has been tested thoroughly before use?						
C	Are the conclusions valid given the data presented?						
C	Are any counterintuitive results from the model explained and justified?						
C	If the model has been calibrated against independent data, have any differences been explained and						
C	Have the results of the model been compared with those of previous models and any differences in						
Data (D)							
D1	Data identification	D2B	Treatment effects	D4	Assessment of	D4C	Heterogeneity
D2	Pre-model data	D2C	Quality-of-life	D4A	Methodological	D4D	Parameter
D2A	Baseline data	D3	Data incorporation	D4B	Structural		
D1a	Are the data identification methods transparent and appropriate given the objectives of the model?						
D1b	Where choices have been made between data sources, are these justified appropriately?						
D1c	Has particular attention been paid to identifying data for the important parameters in the model?						
D1d	Has the process of selecting key parameters been justified and systematic methods used to identify						
D1e	Has the quality of the data been assessed appropriately?						
D1f	Where expert opinion has been used, are the described and justified?						
D2a	Are the pre-model data analysis methodology based on justifiable statistical and epidemiological						
D2A	Is the choice of baseline data described and justified?						
D2A	Are transition probabilities calculated appropriately?						
D2A	Has a half cycle correction been applied to both cost and outcome?						
D2A	If not, has this omission been justified?						
D2B	If relative treatment effects have been derived from trial data, have they been synthesised using						
D2B	Have the methods and assumptions used to extrapolate short-term results to final outcomes been						
D2B	Have alternative assumptions been explored through sensitivity analysis?						

D2B	Have assumptions regarding the continuing effect of treatment once treatment is complete been
D2B	Have alternative assumptions been explored through sensitivity analysis?
D2C	Are the utilities incorporated into the model appropriate?
D2C	Is the source for the utility weights referenced?
D2C	Are the methods of derivation for the utility weights justified?
D3a	Have all data incorporated into the model been described and referenced in sufficient detail?
D3b	Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices
D3c	Is the process of data incorporation transparent?
D3d	If data have been incorporated as distributions, has the choice of distribution for each parameter
D3e	If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?
D4a	Have the four principal types of uncertainty been addressed?
D4b	If not, has the omission of particular forms of uncertainty been justified?
D4A	Have methodological uncertainties been addressed by running alternative versions of the model
D4B	Is there evidence that structural uncertainties have been addressed via sensitivity analysis?
D4C	Has heterogeneity been dealt with by running the model separately for different sub-groups?
D4D	Are the methods of assessment of parameter uncertainty appropriate?
D4D	Has probabilistic sensitivity analysis been done, if not has this been justified?
D4D	If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly

Appendix 5: list of excluded studies

Study	Reason for exclusion
Andrews et al. 2004 (323)	Not an economic evaluation
Clark et al. 1998 (324)	Results of CEA for BD not presented separately
Caro et al. 2006 (325)	Budgetary impact analysis (not an economic evaluation)
Conney et al.1999 (326)	Costing study
Dean et al. 2004 (327)	Review
DSI et al. 2005 (328)	Could not retrieve
Fleurence et al. 2007 (329)	Review
Goldberg et al. 1999 (330)	Review
Greenhalgh et al. 2005 (331)	Doesn't present results for BD
Roijen et al. 2004 (332)	Not an economic evaluation
Hong et al. 2011 (333)	Doesn't compare any interventions
Kieth et al. 2009 (334)	Review
Li et al. 2011 (335)	Costing study
Löthgren et al. 2004 (138)	Review
Namjoshi et al.2002 (336)	Before-and-after costing study
Myhr et al. 2006 (337)	Doesn't present results for BD
Reifman et al. 1980 (338)	Not an economic evaluation
Scott et al. 2009 (339)	Costing study
Scott et al. 2011 (340)	Review
Shaya et al. 2007 (341)	Budgetary impact analysis (not an economic evaluation)
Simons et al. 2004 (342)	Costing study
Steffens et al.1996 (343)	Not an economic evaluation
Virtanen et al. 2011 (344)	Not an economic evaluation
Zhang et al. 2008 (345)	Costing study

Appendix 6: Assessing the quality of Trial-based economic evaluations in Bipolar disorder using centre of review and dissemination's checklist

Study	Lam, et al 2005	Revidki, et al, 2005	Simon GE, 2006	Bauer et al, 2006	Namjoshi, et al, 2002	Revidki, et al, 2002			
Study design									
1									
2									
3									
4									
5									
6									
7									
Data collection									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									
18									
19									

20									
21									
Analysis & interpretation of results									
22									
23									
24									
25									
26									
27									
28									
29									
30									
31									
32									
33									
34									
35									
36									

Appendix 7: Assessing the quality of Modelling-based economic evaluations of management strategies in Bipolar disorder using Philips's checklist

Study	Keck, 1997	Birdle, 2004	Chisholm, 2005	Calvert, 2006	NICE, 2006	Klok, 2007	McKendrick, 2007	Soares 2007	Fajutrao, 2009	Menzin, 2009	Woodward, 2009	Woodward, 2010	Kasteng, 2011	Chisholm, 2012	Eckman, 2012
Structure															
S1a	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
S1b	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
S1c	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+
S2a	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+
S2b	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
S2c	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
S2d	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+
S3a	-	-	NA	-	-	+	-	+	-	+	+	+	+	NA	+
S3b	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
S3c	-	-	NA	+	+	NA	+	+	+	NA	+	+	+	NA	+
S4a	-	-	-	+	+	-	+	+	+	+	+	+	+	-	+
S4b	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
S5a	+	+	-	+	+	+	+	+	+	+	+	+	+	-	+
S5b	+	+	-	-	-	+	-	-	-	+	-	+	-	-	+
S5c	NA	NA	+	-	+	NA	+	+	+	NA	-	NA	-	+	+
S6a	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
S7a	-	-	+	-	+	+	-	+	-	+	-	-	+	+	+
S7b	+	+	-	+	+	+	+	+	+	NA	+	+	+	-	+
S8a	-	-	NA	-	-	NA	-	+	-	+	+	+	+	NA	+
S9a	-	-	-	+	+	NA	+	+	+	NA	+	+	+	-	NA

Study	Keck, 1998	Birdle, 2004	Chisholm, 2005	Calvert, 2006	NICE, 2006	Klok, 2007	McKendrick, 2007	Soares 2007	Fajutrao, 2009	Menzin, 2009	Woodward, 2009	Woodward, 2010	Kasteng, 2011	Chisholm, 2012	Eckaman, 2012
Data															
D1a	+	+	+	+	+	-	+	-	+	+	+	+	+	+	+
D1b	-	+	-	-	+	-	-	+	-	-	-	-	-	-	+
D1c	-	+	+	+	+	-	+	+	+	-	-	-	-	+	+
D1d	-	-	+	+	+	-	+	+	+	-	-	+	-	+	+
D1e	-	-	-	+	+	-	+	+	+	-	-	-	-	-	+
D1f	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D1g	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Aa	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Ab	-	+	+	+	-	+	+	+	+	+	+	+	+	+	+
D2Ac	+	+	+	+	-	+	+	-	+	+	+	+	-	+	+
D2Ad	+	+	+	+	-	+	+	-	+	+	+	+	-	+	+
D2Ae	+	+	+	+	-	+	+	-	+	+	+	+	-	+	+
D2Ba	-	+	-	-	+	-	-	+	-	-	-	-	-	-	+
D2Bb	+	+	-	+	+	+	-	+	+	-	+	+	+	-	+
D2Bc	+	+	+	+	-	-	-	+	-	+	-	-	+	+	+
D2Bd	+	-	-	+	+	+	-	+	-	+	-	-	+	-	+
D2Be	+	-	-	+	-	+	+	+	-	+	-	-	+	-	+
D2Ca	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Cb	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Cc	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Cd	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+
D2Da	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Db	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Dc	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+

Study	Keck, 1998	Birdle, 2004	Chisholm, 2005	Calvert, 2006	NICE, 2006	Klok, 2007	McKendrick, 2007	Soares 2007	Fajutrao, 2009	Menzin, 2009	Woodward, 2009	Woodward, 2010	Kasteng, 2011	Chisholm, 2012	Eckaman, 2012
Data															
D1a	+	+	+	+	+	-	+	-	+	+	+	+	+	+	+
D1b	-	+	-	-	+	-	-	+	-	-	-	-	-	-	+
D1c	-	+	+	+	+	-	+	+	+	-	-	-	-	+	+
D1d	-	-	+	+	+	-	+	+	+	-	-	+	-	+	+
D1e	-	-	-	+	+	-	+	+	+	-	-	-	-	-	+
D1f	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D1g	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Aa	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Ab	-	+	+	+	-	+	+	+	+	+	+	+	+	+	+
D2Ac	+	+	+	+	-	+	+	-	+	+	+	+	-	+	+
D2Ad	+	+	+	+	-	+	+	-	+	+	+	+	-	+	+
D2Ae	+	+	+	+	-	+	+	-	+	+	+	+	-	+	+
D2Ba	-	+	-	-	+	-	-	+	-	-	-	-	-	-	+
D2Bb	+	+	-	+	+	+	-	+	+	-	+	+	+	-	+
D2Bc	+	+	+	+	-	-	-	+	-	+	-	-	+	+	+
D2Bd	+	-	-	+	+	+	-	+	-	+	-	-	+	-	+
D2Be	+	-	-	+	-	+	+	+	-	+	-	-	+	-	+
D2Ca	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Cb	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Cc	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Cd	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+
D2Da	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Db	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D2Dc	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+

Study	Keck, 1998	Birdle, 2004	Chisholm, 2005	Calvert, 2006	NICE, 2006	Klok, 2007	McKendrick, 2007	Soares 2007	Fajutrao, 2009	Menzin, 2009	Woodward, 2009	Woodward, 2010	Kasteng, 2011	Chisholm, 2012	Eckaman, 2012
D3a	+	+	-	+	+	-	+	-	+	+	+	+	+	-	+
D3b	+	-	-	+	+	-	+	+	+	-	+	+	+	-	+
D3c	+	+	-	+	N/A	-	N/A	-	+	-	-	-	+	-	+
D3d	N/A	-	N/A	N/A	N/A	N/A	N/A	-	-	N/A	-	+	+	+	+
D3e	N/A	+	N/A	N/A	N/A	N/A	N/A	+	-	N/A	-	+	+	+	+
D4.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D4.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D4a	-	-	+	-	+	-	-	+	-	+	-	-	+	-	+
D4b	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D4c	-	-	+	+	+	-	-	+	-	-	-	-	+	+	+
D4Da	-	-	-	-	-	-	-	+	+	+	-	+	+	+	+
D4Db	-	-	+	+	+	-	+	+	+	-	+	+	+	+	+
Consistency															
C1a	-	-	+	-	+	+	-	+	-	-	-	-	-	+	+
C2a	-	+	-	+	+	+	+	+	+	+	-	+	+	-	+
C2b	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C2c	N/A	N/A	N/A	+	N/A	N/A	+	+	+	N/A	+	+	N/A	N/A	+
C2d	N/A	N/A	N/A	+	N/A	N/A	+	+	+	N/A	+	+	N/A	N/A	+

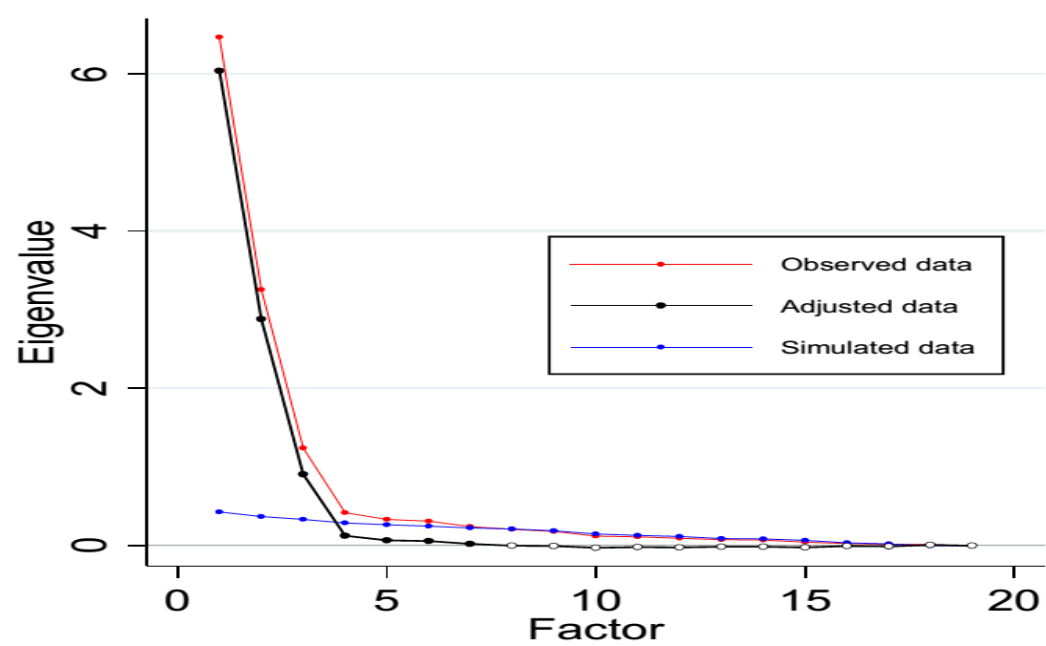
Supplementary table 1: Model-based economic evaluations marked using explicit quality criteria (Philip's Checklist)

+	Meets quality criteria
-	Does not meet quality criteria
N/A	Not applicable

Appendix 8: Correlation matrix between the ASRM (mania), EQ-5D-3L, and VAS

Items	EQ-5D-3L	Mobility	Self-care	Usual activities	Pain/discomfort	Depression/anxiety	VAS
ALTMAN score	0.092*	-0.004	-0.003	-0.079*	0.010	-0.157*	0.126*
Elevated mood	0.158*	-0.016	-0.075*	-0.122*	-0.032	-0.216*	0.195*
Increased self-esteem	0.083*	-0.027	-0.009	-0.080*	0.010	-0.129*	0.143*
Decreased need of sleep	-0.074*	0.099*	0.070*	0.089*	0.105*	0.005	-0.061*
Pressured speech	0.006	0.030	0.078*	-0.007	0.077*	-0.102*	0.066*
Psychomotor agitation	0.114*	-0.019	-0.048	-0.105*	-0.028	-0.136*	0.133*
* statistically significant at P value<0.05							

Appendix 9: Horn's parallel analysis



Appendix 10: Characteristics of studies exploring appropriateness of EQ-5D-3L in Bipolar disorder. Partly adapted and updated from Brazier et al 2014.

Study	Study context	Measures used	Reliability	Known group validity	Convergent validity	Responsiveness	Overall
Vojta et al, 2001	USA, n=86	SF-12, ISS, EQ-5D-3L	NR	EQ-5D-3L scores showed rank order with euthymic= manic/hypomanic> mixed =depressed states SF-12 mental subscale discriminated across all the mood states in expected direction; no difference in SF-12 physical subscale scores	Poor Correlations between EQ-5D-3L and symptom and functional measures	NR	Manic states have lower HRQoL than euthymia; depression primary determinant of HRQoL
Hakkaart-van Roijen, 2004	Netherlands, n=40	SF-36, EQ-5D-3L	NR	Mean EQ-5D-3L score was not significantly different from that in the general population (0.82 Vs 0.88, P=0.084)	NR	NR	No significant differences observed in EQ-5D-3L
Hayhurst et al, 2006	UK, n=221/253	ASRM, BDI, HRSD-17, LIFE, SAS, EQ-5D-3L	NR	Significant differences observed in EQ-5D scores across depression states	moderate significant correlation (-0.41- -0.45) with scores on all depression measures	NR	EQ-5D can be used to assess the HRQoL of depressed BD patients
Arbuckle et al, 2009	USA, n=194/225	BFSQ CGI-BP, EQ-5D-3L, SDS, SF-36, MADRS, YMRS	NR	NR	moderate significant correlations (-0.51) between EQ-5D-3L and SDS	NR	The SDS has strong validity, reliability, and ability
Saarni et al, 2010	Finland, n=20	BDI, 15 D, EQ-5D-3L, SAPS, SANS	NR	No differences in EQ-5D-3L scores in relation to disease severity after controlling for age, gender, income, education and BDI score	significant correlations (-0.71) between EQ-5D-3L and depressive symptoms		The EQ-5D-3L is problematic as an outcome measure in BD and psychotic disorders

Subero et al, 2013	Australia, n=176/240	CGI-BP , YMRS , HAMD21, EQ-5D-3L , SF-36	NR	Differences in EQ-5D scores noted between patients and general population mean score (0.77 Vs 0.88)	NR	NR	cut-off value of ≥ 0.88 on the EQ-5D score are appropriate to define full remission in BD
15D:15 dimensions ; BDI :Beck Depression Inventory; BFSQ :Bipolar Functional Status Questionnaire; CGI-BP: Clinical Global Impressions Bipolar Version Severity of Illness scale ; HRSD-17/21:Hamilton Rating Scale for Depression 17/21 items; ISS: Internal State Scale; LIFE: LIFE-II ratings of Depression and Mania; MADRS: Montgomery-Åsberg Depression Rating Scale; MSSS: Major Symptoms of Schizophrenia Scale; SDS: Sheehan Disability Scale; SF-12: Short form-12; SF-36: Short form-36; YMRS:Young Mania Rating Scale							

Appendix 11: Sample Questionnaire used in the OXTEXT-6 Trial-based economic evaluation in Bipolar disorder

OXTEXT ID: OXT Date Sent by OXTEXT:



OXTEXT-6: FIMM Follow-up Questionnaire (12 Months)

What date did you fill out this questionnaire?

SECTION 0: Integrated Mood Management (IMM)

0.1 Using the IMM Manual

As part of Integrated Mood Management, you were recommended to read the Manual and various handouts, and do various exercises.

We wrote to you on and asked you how much time you had spent on these tasks.

We now need to know how much time you have spent on these tasks between then and now.

Since please estimate how much time in total you have spent reading the IMM manual, engaging in any IMM homework activities or using your Mood Management Plan.

 hours

Please note that this figure should be the total amount of time, in hours, that you have spent on these activities in the last 9 months (40 weeks).

For example:

- if you did about one hour per month, the figure you enter would be 9;
- if you did about one hour per week, the figure would be 40;
- if you did none, the figure would be 0.

Please be honest; we don't mind whether you used the manual and plan a lot or a little.

1. HEALTH ECONOMICS QUESTIONNAIRE

This questionnaire asks about your contacts with health and social services (Section 1), changes to your medication (Section 2), accommodation (Section 3), distance to care (Section 4), employment (Section 5) and informal care (Section 6).

during the specified time period:

If you do not know the exact answer, please give your best estimate.

If an answer is 0 (zero), please indicate this instead of leaving the relevant box blank.

Dates of interest: From: To:

SECTION 1: Contact with health and social services during the specified time period

1.1 Contact with health and social care staff

Contact with	Number of contacts		
	By phone*	At health or social services premises	At home or in the community (e.g. parks, cafes)
Psychiatrist			
Psychologist			
GP			
Community Mental Health Nurse (CPN)			
Community Nurse (General Health)			
Practice Nurse (at the GP clinic)			
Drug/alcohol service worker			
Social worker			
Community Support worker (unqualified)			
Home help/home care worker			
Housing worker			
Voluntary/Charity worker			
Emergency services (ambulance)			

* Please include only non-administrative phone contacts (i.e. contacts directly related to your health status at the time).

1.2. Contact with alternative therapist during the specified time period

How often have you visited an alternative practitioner such as an acupuncturist or aromatherapist?
(If the therapy was paid for by you, the patient, consider it a private service.)

Number of visits	
NHS services	Private services

1.3 Hospital visits/admissions during the specified time period

1.3.1 Mental health outpatient contacts

Were any of your contacts with the Psychiatrists/Psychologists listed in question 1.1 with staff from the specialist mood disorder clinic at the Warneford Hospital, Oxford? If so, please give the number of contacts here

Psychiatrist (from the specialist mood disorder clinic at the Warneford Hospital)

Psychologist (from the specialist mood disorder clinic at the Warneford Hospital)

Number of contacts	
By phone*	Face-to-face visits

*Please include only non-administrative phone contacts (i.e. contacts directly related to your health status at the time).

1.3.2 Non-mental health outpatient visits

a) How many times have you visited a hospital as an outpatient for reasons not related to your psychiatric illness?

Number of visits

b) If the answer to question a) is other than 0 (zero), please give details in the following table (if necessary, please continue in Section 7).

Type of outpatient clinic (e.g. cardiology, orthopaedics, diabetes, general surgery)	Number of visits

1.3.3 Day-patient/day hospital admissions

a) How many times have you been admitted to hospital as a day-patient (i.e. no overnight stay) or visited a day hospital?

Number of visits

b) If the answer to question a) is other than 0 (zero), please give details in the following table (if necessary, please continue in Section 7).

Type of department (e.g. psychiatry, radiology, ophthalmology, surgery)	Number of visits

1.3.4 Accident and Emergency (A&E) Department visits

Number of visits

How many times have you visited an Accident and Emergency (A&E) Department?

1.3.5 Mental health hospital admissions

Number of
admissions

a) How many times have you been admitted to a psychiatric ward overnight?

b) If the answer to question a) is other than 0 (zero), please give details in the following table (if necessary, please continue in Section 7).

Admission	Total number of nights in hospital	How many of the nights were in Psychiatric Intensive Care Unit (PICU)?
1		
2		
3		

1.3.6 Non-mental health hospital admissions

Number of
admissions

a) How many times have you been admitted to hospital overnight to any other (non-psychiatric) ward?

b) If the answer to question a) is other than 0 (zero), please give details in the following table (if necessary, please continue in Section 7).

Admission	Type of department	Total number of nights in hospital
1		
2		
3		

1.4 Attendance at other groups/programmes during the specified time period

	Number of visits
a) How many times have you attended a day centre (i.e. groups/programmes not run by health care staff)?	
b) How many times have you attended a drop-in centre (including street agencies)?	
c) How many times have you attended a self-help/support group related to your psychiatric illness?	

1.5 Other services used during the specified time period

a) Have you made use of any other health services or social services not mentioned above (e.g. private psychologist, dietician, nutritionist, occupational therapist, etc.)? Yes ☐ No ☐

b) If the answer to question a) is Yes, please give details in the following table (if necessary, please continue in Section 7). If the answer to question a) is No, please leave the following table blank.

Type of Service	Number of times

SECTION 2: Changes to psychiatric medication during the specified time period

a) Have you made any changes to the **medicines** that you usually take **for mental health problems** during the specified time period? Yes ☐ No ☐

If the answer to question a) is Yes, please give details of the change(s) in the following table (if necessary, please continue in Section 7). If the answer to question a) is No, please leave the following tables blank.

2.1 Changes to oral psychiatric medication taken daily

Drug name	Date of change/start (dd/mm/yyyy)	Total dose/day	Unit (e.g mg)	Reason for change/start (if known)
e.g, Lamotrigine	10/11/1999	400	mg	
	 / /			
	 / /			
	 / /			

2.2 Changes to oral psychiatric medication taken regularly, but not daily

Drug name	Date of change/start (dd/mm/yyyy)	Frequency (occasions per month)	Dose taken at once	Unit (e.g. mg)	Reason for change/start (if known)
<i>e.g. Diazepam</i>	<i>11/08/1998</i>	<i>5</i>	<i>2</i>	<i>mg</i>	
	 / /				
	 / /				

2.3 Changes to non-oral psychiatric medication taken regularly (such as depot injection)

Drug name	Date of change/start (dd/mm/yy)	Frequency (occasions per month)	Dose given at once	Unit (e.g. mg, ml)	Reason for change/start (if known)
<i>e.g. Clopixol</i>	<i>07/02/2010</i>	<i>1</i>	<i>2</i>	<i>ml</i>	
	 / /				
	 / /				

SECTION 3: Accommodation

3.1 Where do you normally live?

- i) In own home (owned) ☐
- ii) In own home (rented) ☐
- iii) In a relative's/friend's home ☐
- iv) In a nursing home ☐
- v) In a residential home ☐
- vi) In sheltered accommodation ☐
- vii) Other ☐ Please specify

SECTION 4: Distance to care

4.1 What is the estimated distance between your accommodation and the outpatient psychiatric clinic you usually attend for your illness?

 miles

SECTION 5: Employment and time off work due to your illness during the specified time period

5.1 Are you currently in paid employment, self-employed, or in full-time education?

- i) Yes, in full-time employment / full-time self-employed ☐
- ii) Yes, in part-time employment / part-time self-employed ☐
- iii) Yes, in full-time education ☐
- iv) No, not in paid employment, self-employed, or in full-time education ☐

(Please specify e.g. retired, on benefits, unemployed:)

If the answer to question 5.1 is i), ii) or iii), please complete question 5.2.

5.2 During the specified time period, how many days have you been unable to work/study because you have been unwell?

 days

SECTION 6: Informal care during the specified time period

6.1 On average, how many hours per week have friends and family members spent helping you with your illness during the specified time period? (e.g. personal care, support with daily living skills)

 hours

SECTION 7: Additional information**

***Please indicate also the number of the question the information belongs to.*

2. HEALTH RELATED QUALITY OF LIFE



Health Questionnaire

By placing a tick in one box in each group below, please indicate which statements best describe your own health state today.

Mobility

- I have no problems in walking about ☐
- I have some problems in walking about ☐
- I am confined to bed ☐

Self-Care

- I have no problems with self-care ☐
- I have some problems washing or dressing myself ☐
- I am unable to wash or dress myself ☐

Usual Activities *(e.g. work, study, housework, family or leisure activities)*

- I have no problems with performing my usual activities ☐
- I have some problems with performing my usual activities ☐
- I am unable to perform my usual activities ☐

Pain/Discomfort

- I have no pain or discomfort ☐
- I have moderate pain or discomfort ☐
- I have extreme pain or discomfort ☐

Anxiety/Depression

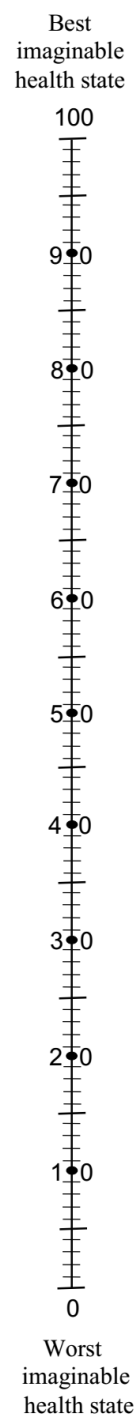
- I am not anxious or depressed ☐
- I am moderately anxious or depressed ☐
- I am extremely anxious or depressed ☐

UK (English) © 1990 EuroQol Group EQ-5D™ is a trade mark of the EuroQol Group

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and the worst state you can imagine is marked 0.

We would like you to indicate on this scale how good or bad your own health is today, in your opinion. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state is today.

**Your own
health state
today**



3. GENERAL QUALITY OF LIFE

ABOUT YOUR OVERALL QUALITY OF LIFE

Please indicate which statements best describe your overall quality of life at the moment by placing a tick () in **ONE** box for each of the five groups below.

1. Feeling settled and secure

- I am able to feel settled and secure in **all** areas of my life
- I am able to feel settled and secure in **many** areas of my life
- I am able to feel settled and secure in **a few** areas of my life
- I am **unable** to feel settled and secure in **any** areas of my life

☐	4
☐	3
☐	2
☐	1

2. Love, friendship and support

- I can have **a lot** of love, friendship and support
- I can have **quite a lot** of love, friendship and support
- I can have **a little** love, friendship and support
- I **cannot** have **any** love, friendship and support

☐	4
☐	3
☐	2
☐	1

3. Being independent

- I am able to be **completely** independent
- I am able to be independent in **many** things
- I am able to be independent in **a few** things
- I am **unable** to be at all independent

☐	4
☐	3
☐	2
☐	1

4. Achievement and progress

- I can achieve and progress in **all** aspects of my life
- I can achieve and progress in **many** aspects of my life
- I can achieve and progress in **a few** aspects of my life
- I **cannot** achieve and progress in **any** aspects of my life

☐	4
☐	3
☐	2
☐	1

5. Enjoyment and pleasure

- I can have **a lot** of enjoyment and pleasure
- I can have **quite a lot** of enjoyment and pleasure
- I can have **a little** enjoyment and pleasure
- I **cannot** have **any** enjoyment and pleasure

☐	4
☐	3
☐	2
☐	1

Please ensure you have only ticked ONE box for each of the five groups.

4. GENERAL QUALITY OF LIFE (OxCAP-MH)

This questionnaire asks about your overall quality of life.

1	Does your health in any way limit your daily activities, compared to most people of your age? <i>[Please tick one]</i>	☐ Yes ☐ No
2	Are you able to meet socially with friends or relatives? <i>[Please tick one]</i>	☐ Yes ☐ No
3	In the past 4 weeks, how often have you lost sleep over worry? <i>[Please tick one]</i>	☐ Always ☐ Most of the time ☐ Some of the time ☐ Hardly ever ☐ Never
4	In the past 4 weeks, how often have you been able to enjoy your recreational activities? <i>[Please tick one]</i>	☐ Always ☐ Most of the time ☐ Some of the time ☐ Hardly ever ☐ Never
5	How suitable or unsuitable is your accommodation for your current needs? <i>[Please tick one]</i>	☐ Very suitable ☐ Fairly suitable ☐ Neither suitable nor unsuitable ☐ Fairly unsuitable ☐ Very unsuitable
6	Please indicate how safe you feel walking alone in the area near your home: <i>[Please tick one]</i>	☐ Very safe ☐ Fairly safe ☐ Neither safe nor unsafe ☐ Fairly unsafe ☐ Very unsafe
7	Please indicate how likely you believe it to be that you will be assaulted in the future (including sexual and domestic assault): <i>[Please tick one]</i>	☐ Very likely ☐ Fairly likely ☐ Neither likely nor unlikely ☐ Fairly unlikely ☐ Very unlikely
8a	How likely do you think it is that you will experience discrimination? <i>[Please tick one]</i>	☐ Very likely ☐ Fairly likely ☐ Neither likely nor unlikely ☐ Fairly unlikely ☐ Very unlikely

8b On what grounds do you think it is likely that you will be discriminated against? <i>(Please complete only if your answer to question 8A is 'Very likely' or 'Fairly likely'.)</i> <i>[Please tick up to three]</i>	☐ Race/ethnicity ☐ Gender ☐ Religion ☐ Sexual orientation ☐ Age ☐ Health or disability (including mental health)				
9 Please indicate how strongly you agree or disagree with the following statements: <i>[Please tick one box for each statement]</i>					
	Strongly Agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
9a I am able to influence decisions affecting my local area.					
9b I am free to express my views, including political and religious views.					
9c I am able to appreciate and value plants, animals and the world of nature.					
9d I respect, value and appreciate people around me.					
9e I find it easy to enjoy the love, care and support of my family and friends.					
9f I am free to decide for myself how to live my life.					
9g I am free to use my imagination and to express myself creatively (e.g. through art, literature, music, etc.).					
9h I have access to interesting forms of activity (or employment).					
10 The next question asks you to think about your life expectancy: The average life expectancy in the UK is 77 years for men and 81 years for women. Some people think they are going to live longer than the average person whilst other people believe they are going to live shorter than the average person.					
10a Given your family history, dietary habits, lifestyle and health status, do you expect to live? <i>[Please tick one]</i>	☐ Longer than average ☐ Same as average ☐ Shorter than average				

10b Until what age do you believe you are going to live? <i>(Please complete only if your answer to question 10A is 'Longer than average' or 'Shorter than average'.)</i>	years
--	---

5. SOCIODEMOGRAPHIC INFORMATION

1 What is your weight? 	st lbs or kg
2 Which religion, religious denomination or religious body do you belong to? <i>[Please tick one]</i>	☐ Christian ☐ Hindu ☐ Jewish ☐ Muslim ☐ Sikh ☐ Other ☐ None ☐ Prefer not to answer
3 What is your marital status? <i>[Please tick one]</i>	☐ Married/civil partnership ☐ Living with partner ☐ Separated ☐ Divorced ☐ Widowed ☐ Never married ☐ Prefer not to answer
4 How many dependent children do you have – that is children dependent on your income? <i>[Please tick one]</i>	☐ None ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ More than 4 ☐ Prefer not to answer
5 What is the highest educational qualification you have? <i>[Please tick one]</i>	☐ Postgraduate degree ☐ First degree ☐ Higher education below degree ☐ Highers/A Levels or equivalent ☐ Standard grades 1-3/GCSEs or equivalent ☐ Standard grades 4-7/CSEs or Equivalent ☐ Foreign or other ☐ No qualification ☐ Prefer not to answer

6. BIPOLAR MOOD MANAGEMENT QUESTIONNAIRE

(2011 05 24)

This questionnaire asks about your knowledge and understanding of self-management for bipolar disorder.

- Please read each statement carefully.
- Then, place a tick in the box corresponding to the answer which best describes how much you agree with that statement.
- Do not spend too long on each question – your first impressions are probably best.
- Please give a totally honest answer.
- Please give an answer for every part (a-d) to every question (1-10) as in the example below:

Example:

	Agree	Neither agree nor disagree	Disagree
4 a It's important for people with bipolar disorder to keep an eye out for the early symptoms which show that their mood is changing	☒	☐	☐
b I know which early symptoms show that my mood is changing	☐	☒	☐
c I know how to keep a close eye out for the early symptoms which show that my mood is changing	☐	☐	☒
d I keep a close eye out for the early symptoms which show that my mood is changing	☐	☐	☒

	Agree	Neither agree nor disagree	Disagree
1 a People with bipolar disorder need to actively manage their mood stability	☐	☐	☐
b I need to actively manage <i>my</i> mood stability	☐	☐	☐
c I know how to actively manage my mood stability	☐	☐	☐
d I usually do actively manage my mood stability	☐	☐	☐
2 a I understand how daily routine affects mood in people with bipolar disorder	☐	☐	☐
b I understand how my daily routine affects <i>my</i> mood	☐	☐	☐
c I know how to manage my daily routine to help to manage my mood	☐	☐	☐
d I manage my daily routine to help to manage my mood	☐	☐	☐

		Agree	Neither agree nor disagree	Disagree
3	a Having a personal Mood Management Plan is important for people with bipolar disorder	☐	☐	☐
	b Having a personal Mood Management Plan is important for <i>me</i>	☐	☐	☐
	c I have a personal Mood Management Plan	☐	☐	☐
	d I use a personal Mood Management Plan	☐	☐	☐
4	a It's important for people with bipolar disorder to keep an eye out for the early symptoms which show that their mood is changing	☐	☐	☐
	b I know which early symptoms show that my mood is changing	☐	☐	☐
	c I know how to keep a close eye out for the early symptoms which show that my mood is changing	☐	☐	☐
	d I keep a close eye out for the early symptoms which show that my mood is changing	☐	☐	☐
5	a I understand how medicines can help manage mood instability in people with bipolar disorder	☐	☐	☐
	b I understand how medicines can help me to manage <i>my</i> mood instability	☐	☐	☐
	c I know which mood medicines I should take and when I should take them	☐	☐	☐
	d I usually take my mood medicines as prescribed	☐	☐	☐
6	a I know which factors increase the risk of unstable mood in people with bipolar disorder	☐	☐	☐
	b I know which factors in my life increase <i>my</i> risk of unstable mood	☐	☐	☐
	c I know how to manage these risk factors in my life	☐	☐	☐
	d I actively manage these risk factors	☐	☐	☐
7	a Using alcohol / substances can worsen mood stability in people with bipolar disorder	☐	☐	☐
	b My use of alcohol / substances can worsen <i>my</i> mood stability	☐	☐	☐
	c I know how to cut down or stop my use of alcohol / substances when necessary	☐	☐	☐
	d I actively cut down or stop my use of alcohol / substances when necessary	☐	☐	☐

		Agree	Neither agree nor disagree	Disagree
8	a I understand how sleep routine affects mood in people with bipolar disorder	☐	☐	☐
	b I understand how my sleep routine affects <i>my</i> mood	☐	☐	☐
	c I know how to alter my sleep routine to help to manage my mood	☐	☐	☐
	d I alter my sleep routine to help to manage my mood	☐	☐	☐
9	a Regularly monitoring mood can help people with bipolar disorder to manage their mood	☐	☐	☐
	b Regularly monitoring my mood can help <i>me</i> to manage my mood	☐	☐	☐
	c I know how to monitor my mood	☐	☐	☐
	d I am regularly monitoring my mood	☐	☐	☐
10	a I know which factors reduce the risk of unstable mood in people with bipolar disorder	☐	☐	☐
	b I know which factors in my life reduce <i>my</i> risk of unstable mood	☐	☐	☐
	c I know how to strengthen these protective factors in my life	☐	☐	☐
	d I actively strengthen these protective factors	☐	☐	☐

That is the end of the questionnaire.

Please now:

- a) Run through the questionnaire and check that you have answered all the questions.***
- b) Return the questionnaire to us in the stamped addressed envelope provided.***

~ THANK YOU ~

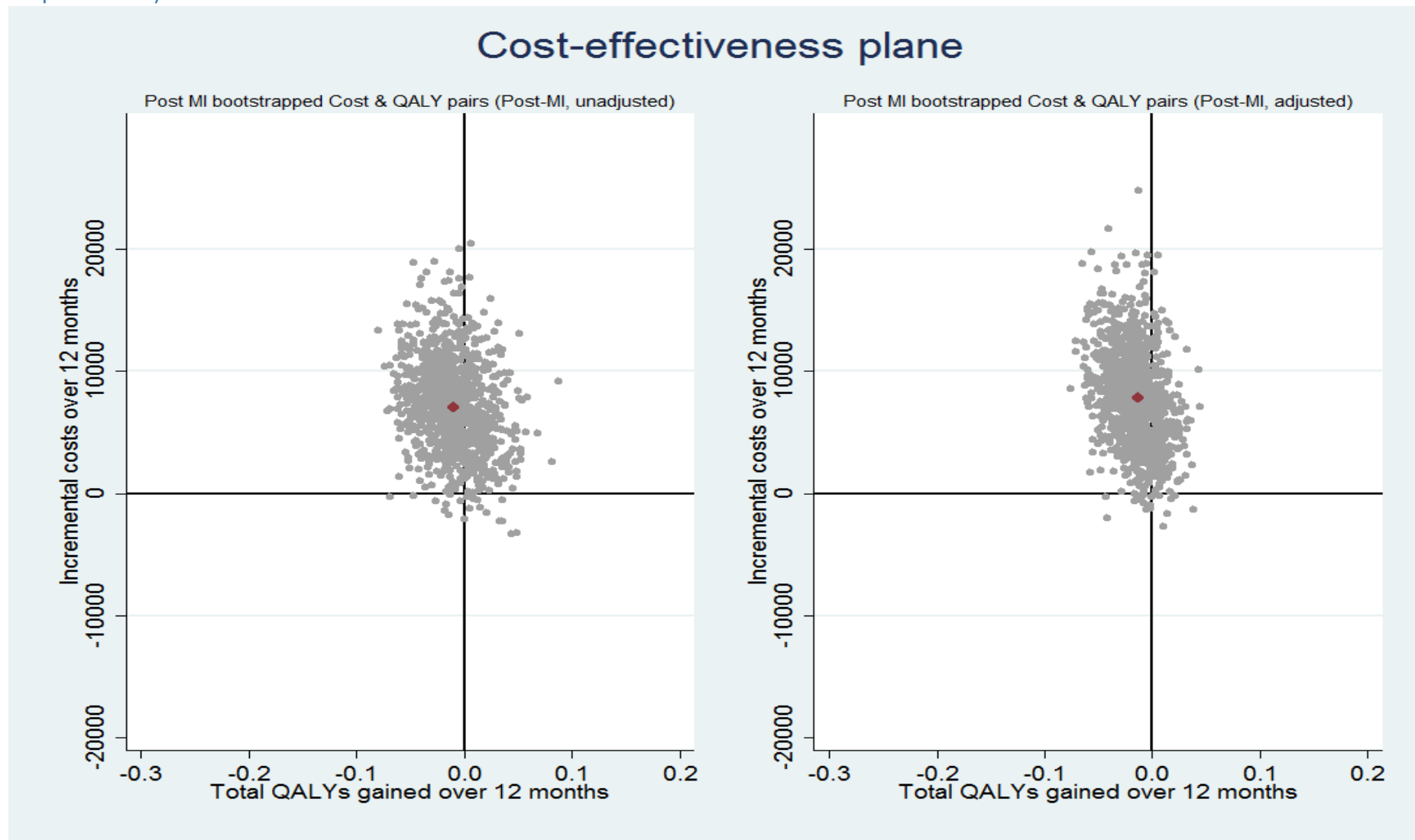
Appendix 12: Services included and unit costs (£) for 2012-2013

Resource use	Unit costs	Source of estimate	Details of estimate
Non mental health in-patient services			
Non-mental health inpatient admissions (per day)	£154-£1536	Department of Health;[25] , various categories	Weighted average costs weighted on full consultant episodes, 2012/13
Non-mental health day care procedures (per day)	£67-£2031	Department of Health;[25]	Mean cost per visit for various specialities 2012/13
Non-mental health out-patient & emergency services			
Outpatient (per visit)	£30-£445	Department of Health;[25]	Mean cost per visit for various specialities 2012/13
Laboratory investigations & screening (per test)	£4-£91	Department of Health;[25]	Mean cost per test for various procedures 2012/13
Diagnostic imaging (per test)	£37-£283	Department of Health;[25]	Mean cost per test for various procedures 2012/13
A & E (per visit)	£115	Department of Health;[25] sheet: AE	Weighted average costs weighted on full consultant episodes
Ambulance (per call-out)	£96	Department of Health;[25] Index sheet: AMB	Weighted average costs
Mental health out-patient			
Hospital Psychiatrist (per hour)	£70	Curtis & Netten (2013), Section 15.7	£140 Cost per contact hour; Consult cost assumed for 30-minute consultation time.
Clinical psychologist (per minute)	£59.00	Curtis & Netten (2013), Section 9.5	Cost per consultation lasting one hour with qualifications

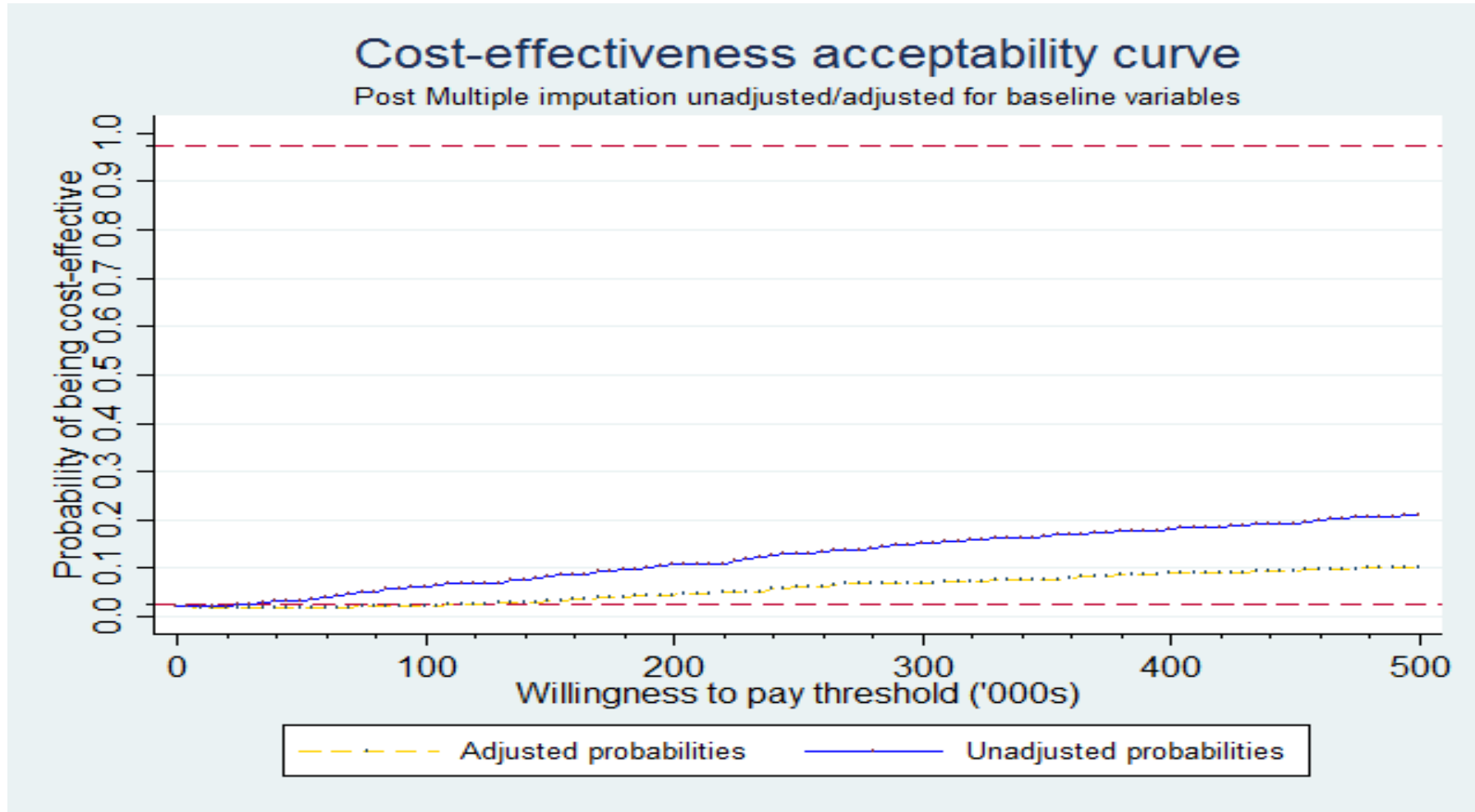
Mental health in-patient			
Psychiatry (per night)	£345	Department of Health;[24]	Weighted average costs weighted on full consultant episodes, and inflated to 2013 prices
Psychiatric intensive care unit (PICU) (per night)	£690	Code SCU1, Department of Health;[24]	inflated to 2013 prices
Mental health day-care	£130	Department of Health;[25]	Weighted average costs weighted on full consultant episodes, 2012/13
Community-based services			
Community psychiatry nurse (per hour)	£39.00	Curtis & Netten (2013), Section 10.2	Cost per consultation lasting one hour with qualifications
Community support worker (per visit)	£12.50	Curtis & Netten (2013), Section 10.5	(£30 per hour of visit) Per visit lasting 25 minutes; duration assumed to be similar to that for a nurse specialist
Drug and alcohol services (per contact)	£60.50	Department of Health;[25] Sheet MH	Estimated £121 per activity based on weighted average figures; Consult cost assumed for 30-minute consultation time.
GP (clinic)	£45.00	Curtis & Netten (2013), Section 10.8b	Per contact lasting 11.7 minutes (£230 per hour of patient contact)
GP (home)	£114	Curtis & Netten (2013), Section 10.8b	Per visit lasting 23.4 minutes
GP (on phone)	£27	Curtis & Netten (2013), Section 10.8b	Per consultation lasting 7.1 minutes

GP practise nurse (clinic)	£13.43	Curtis & Netten (2013), Section 10.6	Average consultation lasting 15.5 minutes (£52 per hour of face-to-face contact)
GP practise nurse (home)	£21.67	Curtis & Netten (2013), Section 10.6/10.7	Per visit lasting 25 minutes; duration assumed to be similar to that for a nurse specialist
GP practise nurse (on phone)	£5.20	Curtis & Netten (2013), Section 10.6/10.7	Per consultation lasting 6 minutes; duration assumed to be similar to that for a nurse specialist
Community nurse/district nurses (per visit)	£29.17	Curtis & Netten (2013), Section 10.1	(£70 per hour of visit) Per visit lasting 25 minutes; duration assumed to be similar to that for a nurse specialist
Home care worker (per hour)	£17	Curtis & Netten (2013), Section 11.6	Based on mean hourly cost of all home care including local authority funded and independent provision
Housing worker (per session)	£31	Curtis & Netten (2013), Section 8.10.1	Based on unit cost of housing support service
Social worker (per visit)	£57	Curtis & Netten (2013), Section 11.2	Based on mean hourly cost with qualifications
Voluntary worker (per minute)	£15.6	assuming national average salary, ONS (2012)	
Other services	£26-£241	Curtis & Netten (2013) & Department of Health;[25]	Mean cost various private and other visits 2012/13
Medicines	Various	BNF 2012-13	
Sick leave (per day)	£93	NICE & Sainsbury MH centre 2010-2011	Inflated to 2013 using Consumer price index
Informal care (per hour)	£15.6	assuming national average salary, ONS (2012)	

Appendix 13: Scatter plot of incremental costs and QALYs on the cost-effectiveness plane (after multiple imputation)



Appendix 14: Cost effectiveness acceptability curve (after multiple imputation), FIMM Vs MIMM



Appendix 15. Results of main and sensitivity analysis (in GBP)

Scenario/Analyses	Mean differences (95% CI, P value)	ICER, Cost (£) per QALY	Probability of FIMM being Cost-effective at a threshold of £20,000
Base Case:			
Societal costs	2963 (-6138 - 12064; 0.510)		
QALY (12-month)	-0.002 (-0.08 - 0.073; 0.956)	-1481500	27%
Sensitivity Analyses:			
(1) Base Case CEA, adjusted for baseline covariates			
Societal costs	3248 (-6047 - 12544; 0.488)		
QALY (12-month)	-0.013 (-0.07 - 0.04; 0.659)	-225714	25%
(2) CEA, with multiple imputation of missing data			
Societal costs	7507 (-666 - 15680; 0.071)		
QALY (12-month)	-0.008 (-0.06 - 0.05; 0.783)	-938375	2%
(3) CEA, with multiple imputation of missing data*			
Societal costs	8311 (127 - 16495; 0.047)		
QALY (12-month)	-0.012 (-0.06 - 0.03; 0.596)	-692583	2%
(4) CEA after removal of outliers			
Societal costs	1712 (-4467 - 7891; 0.553)		
QALY (12-month)	0.014 (-0.063 - 0.090; 0.724)	122286	32%
(5) CEA after removal of outliers*			
Societal costs	1773 (-4510 - 8056; 0.575)		
QALY (12-month)	-0.002 (-0.059 - 0.054; 0.938)	-886500	29%

(6) CEA after multiple imputation & removal of outliers

Societal costs	3075 (-1629 - 7780; 0.197)		
QALY (12-month)	0.002 (-0.054 - 0.058; 0.943)	1537500	8%

(7) CEA after multiple imputation & removal of outliers*

Societal costs	3420 (-1320 - 8160; 0.155)		
QALY (12-month)	-0.003 (-0.049 - 0.043; 0.905)	-1140000	5%

(8) CEA with Total NHS and social services

Total NHS and social services costs	1021 (-1768 - 3809; 0.467)		
QALY (12-month)	-0.002 (-0.08 - 0.073; 0.956)	-486109	25%

(9) CEA with Total NHS and social services

Total NHS and social services costs	1118 (-1582 - 3817; 0.411)		
QALY (12-month)	-0.013 (-0.07 - 0.04; 0.659)	-88690	17%

(10) CEA after multiple imputation with Total NHS and social services

Total NHS and social services costs	4230 (-1547 - 10007; 0.150)		
QALY (12-month)	-0.007 (-0.06 - 0.05; 0.783)	-612999	2%

(11) CEA after multiple imputation with Total NHS and social services*

Total NHS and social services costs	4423 (-1332 - 10178; 0.131)		
QALY (12-month)	-0.012 (-0.06 - 0.03; 0.596)	-368571	2%

(12) CEA with Total NHS and social services without outliers

Total NHS and social services costs	2206 (462 - 3951; 0.014)		
QALY (12-month)	-0.006 (-0.08 - 0.07; 0.873)	-361668	3%

(13) CEA with Total NHS and social services without outliers*

Total NHS and social services costs	2220 (465 - 3974; 0.014)		
QALY (12-month)	-0.013 (-0.07 - 0.04; 0.644)	-166892	1%

(14) CEA after multiple imputation & removal of outliers

Total NHS and social services costs	2144 (195 - 4093; 0.032)		
QALY (12-month)	-0.010 (-0.06- 0.05; 0.713)	-210168	0%

(15) CEA after multiple imputation & removal of outliers*

Total NHS and social services costs	2223 (272 - 4173; 0.026)	-185229	0%
QALY (12-month)	-0.012 (-0.06 - 0.03; 0.596)		
