Data sharing and transparency: the impact on evidence synthesis

---

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy by

Sarah Nevitt

August 2017

Abstract

Thesis title: Data sharing and transparency: the impact on evidence synthesis

Introduction and aims

The use of individual participant data (IPD) for evidence synthesis is widely regarded as the 'gold standard' approach to analysis, in particular for clinical outcomes of a 'time to event' (TTE) nature. However, the undertaking of IPD syntheses can be methodologically complex, time consuming and open to sources of bias if not conducted rigorously.

The aim of this thesis is to investigate and document the practical aspects and challenges of conducting IPD syntheses. Such challenges are of particular relevance in the current research environment with changing functionalities of and attitudes towards data sharing.

Methods

This thesis reports two novel systematic reviews regarding the reporting of aggregate TTE outcomes and analyses in epilepsy monotherapy trials published to 2012 and regarding data retrieval rates and characteristics associated with a high proportion of data retrieved for all published IPD meta-analyses (IPD-MAs) from 1987 to 2015. This thesis also presents the results of an IPD-network meta-analysis (NMA) of antiepileptic drug therapy including detailed documentation of the statistical methodology of the IPD-NMA, IPD requesting and preparation processes, and methods for incorporating summary statistics with IPD for NMA.

Results

The first systematic review of reporting in epilepsy monotherapy trials showed concerning reporting inadequacies relating to definitions, analysis and reporting of TTE outcomes in these trials, suggesting that an IPD approach synthesis is the only feasible option for this topic. The systematic review of 760 published IPD-MAs using systematic methods to identify eligible studies showed that only 25% of these IPD-MAs have had access to all IPD and that IPD-MAs that included only randomised trials, had an authorship policy, included fewer eligible participants and were conducted outside of the Cochrane Database of Systematic Reviews were associated with a high or complete IPD retrieval rate.

IPD was provided for a total of 12,391 out of a total of 17,961 eligible participants (69% of total data) from 36 out of the 77 eligible trials (47%) for the IPD-NMA of antiepileptic drug therapy. This reflects a decline in the IPD retrieval rate from requests made by the Cochrane Epilepsy Group from 1995-2005 to requests made in 2012-2015.

A range of methodological approaches to modelling the relationship between treatment-effect and epilepsy type within the NMA and for including AD with IPD in NMA show that incorporation of the small amount of additional AD available with the IPD in NMA had a negligible impact on results. However, the methodological approach to the relationship between treatment and epilepsy type did impact on numerical results and conclusions.

Conclusions

The work of this thesis has provided a detailed insight into the conduct of an IPD-NMA in epilepsy and highlighted many inadequacies of the conduct and reporting of AD and IPD syntheses across a wide range of clinical disciplines. The work of this thesis was undertaken during a time of great change within the research community regarding how clinical trial data is shared for secondary research and has identified some of the early benefits and importantly challenges and restrictions of new methods of sharing clinical trial data. Unless emerging limitations are addressed with urgency, new methods of data sharing, intended to improve access to clinical trial data, may become a hindrance to IPD synthesis.

Acknowledgements

I would firstly like to thank my supervisors Professors Catrin Tudur Smith and Anthony Marson for their excellent guidance and continuous support in all of my work and studies for the last six years.

I am also very grateful to Professor Rumona Dickson and the Liverpool Reviews and Implementations Group for providing the funding from the UK NIHR HTA Programme for both this PhD and for many training courses I have attended. My thanks also go to the National Institute of Health Research for funding my time working with the Cochrane Epilepsy Group.

The work within this thesis would not have been possible without the contribution of many people; thanks to Maria Sudell, Jennifer Weston, Laura Sutton, Becky Davie, Sally Reynolds and Lisa Williams for providing methodological support and co-review during the systematic reviews and meta-analyses presented in this thesis and to the editorial team of the Cochrane Epilepsy Group, particularly to Graham Chan for designing and performing all systematic searches outlined in this review.

My sincere thanks must also go to my many colleagues and great friends across all of the research groups I have worked with; the Department of Biostatistics at the University of Liverpool, the Cochrane Epilepsy Group, the Cochrane Cystic Fibrosis Disorders and Genetic Disorders Group and the Liverpool Reviews and Implementations Group. The provision of large quantities of caffeine, sugar and alcohol as required over the years has been much appreciated.

Last but not least, I am eternally grateful to my family for always encouraging me and challenging me to be the best that I can be. My final and biggest thanks go to my husband Will, not only for proof-reading this whole thesis, but also for being by my side along every twist and turn of this journey and still wanting to marry me in the middle of it all. I am honoured to submit this work in our family name.

Statement of contribution

The Cochrane Individual Participant Data Meta-Analyses (IPD-MAs) and Network Meta-Analysis (IPD-NMA) reported throughout this thesis were performed by a research team based at the University of Liverpool. The contributions of the research team were as follows:

Sarah J Nevitt (née Nolan, SJN): lead statistician and author, coordination of IPD requests.

Anthony G Marson (AGM): co-ordinating Editor of the Cochrane Epilepsy Group, conception of IPD-MAs and IPD-NMA, second reviewer for the selection of eligible studies, support to IPD requests and to clinical interpretation of results and supervision of this thesis.

Catrin Tudur Smith (CTS): senior statistician and methodological reviewer, original lead author of many of the IPD-MAs and original IPD-NMA, support to IPD requests, to statistical analysis of the current Cochrane IPD-MAs and IPD-NMA and primary supervision of this thesis.

Maria Sudell (MS): statistical support in preparing individual participant data for analysis.

Jennifer Weston (née Pulman, JW): methodological support and second reviewer for risk of bias assessment.

In addition, all members of the research team provided support in writing the Cochrane IPD-MAs and IPD-NMAs and read and approved final manuscripts for publication.

Laura Sutton (LS) was a second reviewer for the systematic review presented in Chapter 3 of this thesis and Becky Davie (BD), Sally Reynolds (SR) and Lisa Williams (LW) were second reviewers for the systematic review presented in Chapter 4 of this thesis. All were based at the University of Liverpool at the time the work was conducted. These reviewers were involved in selecting eligible studies for the reviews and performing double data extraction.

Graham Chan (Information Specialist, Cochrane Epilepsy Group) designed and performed all systematic searches described in this thesis as described in Chapters 3, 4 and 5.

Table of Contents

Abstract ..... **iii**
Acknowledgements ..... **iv**
Statement of contribution ..... **v**
List of tables ..... **x**
List of figures ..... **xii**
Abbreviations ..... **xv**
Publications and presentations of work in this thesis ..... **xvii**
Chapter 1: Introduction **1**
1.1 Evidence synthesis 1
1.1.1 Systematic reviews and meta - analysis 1
1.1.2 Indirect comparisons and network meta - analysis 2
1.1.3 Aggregate data and individual participant data synthesis 4
1.1.4 Time-to-event data 6
1.2 Individual participant data meta-analyses and network meta-analyses in Epilepsy 7
1.2.1 Clinical Setting 7
1.2.2 Rationale for individual participant data approach 8
1.3 Individual participant data sharing: recent initiatives 9
1.4 Thesis objective and structure 12
Chapter 2: Meta-analysis of time-to-event data: a methodology review **13**
2.1 Introduction 13
2.2 Analysis of TTE data in a single trial 14
2.2.1 Direct calculation of the Hazard Ratio 14
2.2.2 Use and interpretation of the Hazard Ratio 15
2.3 Meta-analysis of TTE data 16
2.3.1 Literature review methods 16
2.3.2 Methods for aggregate data 17
2.3.3 Methods for individual participant data (IPD) 24
2.3.4 Other meta-analytic methods 31
2.4 Discussion 35
Chapter 3: Aggregate time-to-event (TTE) data: a systematic review of reporting of outcomes and statistical analyses in epilepsy monotherapy trials **37**
3.1 Introduction 37
3.1.1 Summary of previous reviews of aggregate TTE data 37
3.1.2 TTE outcomes in published epilepsy monotherapy trials 43
3.2 Methods 47
3.2.1 Systematic Search 47
3.2.2 Eligibility Criteria 47

3.2.3 Screening of Studies 48
3.2.4 Data Extraction 48
3.2.5 Data analysis and presentation of results 49
3.3 Results 49
3.3.1 Results of the search 49
3.3.2 Outcomes reported in Epilepsy Monotherapy studies 50
3.3.3 Characteristics of Epilepsy Monotherapy studies reporting at least one time-to-event outcome 52
3.3.4 TTE outcomes reported in Epilepsy Monotherapy studies 57
3.4 Discussion 63
3.4.1 Summary of key results and implications 63
3.4.2 Strengths and limitations 65
3.4.3 Comparison to previous work 66
3.4.4 Concluding remarks 68
Chapter 4: IPD retrieval: A systematic review of individual participant data meta-analyses (IPD-MA) 69
4.1 Introduction 69
4.2 Methods 70
4.2.1 Systematic search methods 70
4.2.2 Eligibility criteria 71
4.2.3 Data extraction 71
4.2.4 Statistical analysis and presentation of results 72
4.2.5 Additional and sensitivity analyses 74
4.3 Results 75
4.3.1 Results of the search 75
4.3.2 Characteristics of IPD-MAs 75
4.3.3 IPD retrieval rate in systematic IPD-MA 79
4.3.4 Unavailability of IPD and the impact on analysis 86
4.4 Discussion 88
4.4.1 Summary of main results 88
4.4.2 Strengths and weaknesses 89
4.4.3 Relation to other studies and implications 90
4.4.4 Concluding remarks 91
Chapter 5: Individual participant data requests: 20 years' experience of the Cochrane Epilepsy Group 92
5.1 Introduction 92
5.2 Selection of studies for IPD-NMA 93
5.2.1. Inclusion criteria 93
5.2.2 Study selection 94
5.3 IPD request methodology and preparation of IPD for network meta-analysis ... 105

5.3.1 IPD requests prior to 2012 105
5.3.2 IPD requests from 2012 to 2015 106
5.3.3 Preparation of IPD for analysis 107
5.4 Results of IPD Requests 110
5.4.1 IPD requests prior to 2012 110
5.4.2 IPD requests from 2012-2015 111
5.4.3 Total IPD available for Cochrane IPD-NMA 115
5.5 Discussion 122
5.5.1 Reflection of experiences Group and implications 122
5.5.2 Relation to the wider context of clinical trial data sharing 124
5.5.3 Concluding remarks 127
Chapter 6: Approaches to network meta-analysis of antiepileptic drug monotherapy .. 128
6.1 Introduction 128
6.2 Statistical methodology 128
6.2.2 First-stage: IPD-NMA with separate models by epilepsy type 130
6.2.3 First stage: IPD-NMA with treatment-by-epilepsy type interaction 131
6.2.4 First stage: IPD-NMA via the 'meta-analysis of interactions' approach 133
6.2.5 Second stage: IPD-NMA methods 134
6.3 Results 137
6.3.1. IPD-NMA results according to the approach for modelling epilepsy type 137
6.3.2 Additional consideration of treatment-covariate interaction 142
6.3 Discussion 144
6.3.1 Summary of main results and implications 144
6.3.2 Strengths and weaknesses 145
6.3.3 Concluding remarks 147
Chapter 7: A Cochrane IPD-NMA of antiepileptic drug monotherapy: additional considerations 148
7.1 Additional methodological considerations 148
7.1.1 Direct (pairwise) evidence 148
7.1.2 Investigation of consistency 149
7.1.3 Additional analyses and sensitivity analyses 150
7.2 Results 152
7.2.1 Primary outcome: Time-to-withdrawal of allocated treatment 155
7.2.2 Secondary outcome: Time-to-first seizure after randomisation 162
7.2.3 Additional analyses and sensitivity analyses 169
7.3 Discussion 173
7.3.1 Summary of main results 173
7.3.2 Strengths and weaknesses 174
7.3.3 Clinical implications and relation to other studies 175
7.3.4 Concluding remarks 177

Chapter 8: Combining individual participant data with aggregate data in network meta-analysis 179
8.1 Introduction 179
8.1.1 Individual patient level compared to aggregate level approaches to meta-analysis 179
8.1.2 Combining IPD and AD in meta-analysis 180
8.1.3 Combining IPD and AD in network meta-analysis 183
8.1.4 Objective 185
8.2 Methods 185
8.2.1 Extraction of aggregate data from epilepsy studies 185
8.2.2 Methods for combining IPD and AD in NMA 189
8.3 Results 191
8.3.1. Individuals with partial seizures 191
8.3.2 Individuals with generalised seizures 192
8.3.3 All individuals (regardless of epilepsy type) 193
8.4 Discussion 194
8.4.1 Summary of results and clinical implications 194
8.4.2 Limitations of the methodological approach and future work 196
8.4.3 Concluding remarks 199
Chapter 9: Discussion and Conclusions 200
9.1 Summary of main findings of the thesis 200
9.2. Implications for practice and research 203
9.2.1 Clinical implications 203
9.2.2 Methodological implications: IPD retrieval rate 205
9.2.3 Methodological implications: conduct and reporting of IPD syntheses 206
9.3 Limitations and future work 209
9.4 Concluding remarks 211
References 212
Appendices 228
Appendix 1: Search strategies and reference list for Chapter 2 228
Appendix 2: Search strategies used in Chapter 3 235
Appendix 3: Data extraction form used in Chapter 3 236
Appendix 4: Screenshot of Microsoft Access Database used in Chapter 3 247
Appendix 5: References of 108 epilepsy monotherapy trials included in Chapter 3 248
Appendix 6: Additional results from Chapter 3 254
Appendix 7: Search strategies used in Chapter 4 268
Appendix 8: Data extraction form used in Chapter 4 273
Appendix 9: Additional results from Chapter 4 275
Appendix 10: References of studies included in Cochrane Epilepsy IPD-NMA 280
Appendix 11: Template IPD request letter and data request form 284

Appendix 12: Procedure for preparing a raw IPD dataset into an analysis ready dataset 288
Appendix 13: Example data structure for NMA in Chapter 6 301
Appendix 14: Additional analysis and sensitivity analyses of IPD-NMA performed to account for inconsistencies in IPD (Chapter 7) 303
Appendix 15: Additional results of the IPD-NMA: time to 12 month and time to 6 month remission 312
Appendix 16: Investigation of inconsistency in Chapter 7 314
Appendix 17: Additional results from Chapter 7 333
Appendix 18: Published work from this thesis 339

List of tables

Table 1: Characteristics of previous reviews on the reporting of aggregate TTE data and current systematic review of epilepsy monotherapy trials 38
Table 2: Summary of the reporting of TTE outcomes and statistical analysis; previous work and current systematic review of epilepsy 39
Table 3: Study designs and participant characteristics of 54 epilepsy monotherapy studies 53
Table 4: Definition of events and censoring in 98 TTE outcomes reported in 54 epilepsy monotherapy studies 59
Table 5: Characteristics of 1280 individual participant data reviews 77
Table 6: Characteristics of all systematic IPD-MA according to proportion of IPD provided 81
Table 7: Multivariable logistic regression model results: Characteristics associated with retrieving 100% of IPD or receiving more than 80% of IPD in 503 IPD-MAs 83
Table 8: Reasons reported for unavailability of IPD in 571 systematic IPD-MA without 100% of IPD retrieved 87
Table 9: Approach to accounting for missing IPD in 571 systematic IPD-MA without 100% of IPD retrieved 87
Table 10: Characteristics of trials included in Cochrane review and IPD-NMA 97
Table 11: Methodological quality of studies included in the Cochrane Epilepsy Review and IPD-NMA 103
Table 12: Outcome of individual participant data requests conducted between approx. 1995 to 2005 and 2012 to 2015. 113
Table 13: Number of participants randomised to each drug for trials with IPD provided or not provided 116
Table 14: Reasons why access could not be provided to de-identified study data following enquiries to CSDR and YODA 125
Table 15: Fixed-effects IPD-NMA results for time-to-withdrawal of allocated treatment according to approach of modelling epilepsy type (individuals with partial seizures) 137

Table 16: Fixed-effects IPD-NMA results for time-to-withdrawal of allocated treatment according to approach of modelling epilepsy type (individuals with generalised seizures) 139
Table 17: One stage meta-analysis results for time-to-withdrawal of allocated treatment according to approach of modelling treatment-by-epilepsy type interaction 143
Table 18: Reasons for withdrawal from allocated treatment..... 153
Table 19: Pairwise and network meta-analysis results - Time-to-withdrawal of allocated treatment for individuals with partial seizures..... 156
Table 20: Pairwise and network meta-analysis results - Time-to-withdrawal of allocated treatment for individuals with generalised seizures 158
Table 21: Pairwise and network meta-analysis results - Time-to-first seizure for individuals with partial seizures..... 163
Table 22: Pairwise and network meta-analysis results - Time-to-first seizure for individuals with partial seizures..... 165
Table 23: Aggregate data available for time-to-withdrawal of allocated treatment 187
Table 24: Aggregate data available for time-to-first seizure 188
Table 25: Network meta-analysis results also incorporating aggregate data: Time-to-withdrawal of allocated treatment for individuals with partial seizures..... 192
Table 26: Network meta-analysis results also incorporating aggregate data. Time-to-first seizure for individuals with partial seizures..... 192
Table 27: Network meta-analysis results also incorporating aggregate data. Time-to-withdrawal of allocated treatment for individuals with generalised seizures. 193
Table 28: Network meta-analysis results also incorporating aggregate data. Time-to-first seizure for individuals with generalised seizures. 193
Table 29: Network meta-analysis results also incorporating aggregate data. Time-to-withdrawal of allocated treatment for all individuals (regardless of epilepsy type)..... 194
Table 30: Network meta-analysis results also incorporating aggregate data. Time-to-first seizure for all individuals (regardless of epilepsy type)..... 194
Table 31: Outcomes reported in 54 epilepsy monotherapy studies without a time-to-event outcome defined..... 254
Table 32: Outcomes reported in 54 epilepsy monotherapy studies with at least one time-to-event outcome clearly or possibly defined..... 260
Table 33 Characteristics of 85 reviews in which IPD was extracted from included studies 275
Table 34: Results of univariate logistic regression analysis 277
Table 35: Multivariable logistic regression: further examination of authorship policy..... 277
Table 36: Multivariable logistic regression: inclusion of type of study 277
Table 37: Multivariable logistic regression results: excluding studies with no information regarding funding..... 278
Table 38: Multivariable logistic regression results: a. less than 80% of IPD was retrieved 278
Table 39: Multivariable logistic regression results: b. 80% or more IPD was retrieved 278
Table 40: Multivariable logistic regression results: C. 100% of IPD was retrieved 279

Table 41: Multivariable logistic regression results: proportion of IPD known compared to proportion of IPD unknown 279
Table 42: Results of fractional logistic regression 279
Table 43 Multivariate logistic regression by proportion of study data retrieved 279
Table 44: Example structure of a dataset in 'pairs' format in Stata version 14. 301
Table 45: Example structure of a dataset in 'augmented' format in Stata version 14. 302

List of figures

Figure 1: Direct and indirect evidence from the network of interventions A, B and C 3
Figure 2: Study Flow Diagram of identification of eligible epilepsy monotherapy studies 51
Figure 3: Study Flow Diagram of identification of eligible IPD-MA 76
Figure 4: Number of distinct systematic IPD-MA published to August 2015 and proportion of IPD provided 79
Figure 5: Characteristics of systematic IPD-MA and proportion of IPD provided. 80
Figure 6: Timeline of IPD requests for Cochrane Epilepsy IPD-NMA 92
Figure 7: Study Flow Diagram: Selection of studies for Cochrane Epilepsy Individual Participant Data Network Meta-Analysis 95
Figure 8: Duration and outcome of data requests for 39 randomised controlled trials of antiepileptic drugs 112
Figure 9: Network plots of pairwise comparisons for all trials, for trials providing IPD and for trials not providing IPD for a network meta-analysis of 10 antiepileptic drugs 119
Figure 10: Characteristics of studies providing or not providing IPD 120
Figure 11: Comparison-adjusted network funnel plot of trial-specific ln HRs for time-to-withdrawal of allocated treatment 122
Figure 12: Network plots of pairwise comparisons for all individuals, participants with partial seizures and participants generalised seizures. 129
Figure 13: All AEDs compared to carbamazepine (CBZ) for time to treatment withdrawal, individuals with partial seizures 159
Figure 14: All AEDs compared to lamotrigine (LTG) for time to treatment withdrawal, individuals with partial seizures 160
Figure 15: All AEDs compared to sodium valproate (VPS) for time to treatment withdrawal, individuals with generalised seizures 160
Figure 16: All AEDs compared to carbamazepine (CBZ) for time-to-first seizure, individuals with partial seizures 166
Figure 17: All AEDs compared to lamotrigine (LTG) for time-to-first seizure, individuals with partial seizures 167
Figure 18: All AEDs compared to sodium valproate (VPS) for time-to-first seizure, individuals with generalised seizures 167

Figure 19: Time-to-withdrawal of allocated treatment and time-to-first seizure adjusted for age (partial seizures) 304
Figure 20: Time-to-withdrawal of allocated treatment and time-to-first seizure adjusted for age (generalised seizures) 304
Figure 21: Time-to-withdrawal of allocated treatment using parametric AFT model (both seizure types) 305
Figure 22: Time-to-first seizure using parametric AFT model (both seizure types) 305
Figure 23: Time-to-withdrawal of allocated treatment and time-to-first seizure, Stephen 2007 excluded (partial seizures) 306
Figure 24: Time-to-withdrawal of allocated treatment and time-to-first seizure, Stephen 2007 excluded (generalised seizures) 306
Figure 25: Time-to-withdrawal of allocated treatment, Reunanen 1996 excluded (both seizure types) 307
Figure 26: Time-to-withdrawal of allocated treatment, Placencia 1996 excluded (both seizure types) 307
Figure 27: Time-to-first seizure, Banu 2007 excluded (both seizure types) 308
Figure 28: Time-to-first seizure, Nieto-Barrera 2001 excluded (both seizure types) 308
Figure 29: Time-to-withdrawal of allocated treatment and time-to-first seizure, seizure type reclassification 1 (see Notes below, results for individuals with partial seizures) 309
Figure 30: Time-to-withdrawal of allocated treatment and time-to-first seizure, seizure type reclassification 1 (see Notes below, results for individuals with generalised seizures) 309
Figure 31: Time-to-withdrawal of allocated treatment and time-to-first seizure, seizure type reclassification 2 (see Notes below, results for individuals with partial seizures) 310
Figure 32: Time-to-withdrawal of allocated treatment and time-to-first seizure, seizure type reclassification 2 (see Notes below, results for individuals with generalised seizures) 310
Figure 33: Time-to-withdrawal of allocated treatment, seizure type reclassification 1 with random-effects (see Notes below, both seizure types) 311
Figure 34: Time to 12 month remission: All IPD-NMA results by epilepsy type 312
Figure 35: Time to 6 month remission: All IPD-NMA results by epilepsy type 313
Figure 36: Time-to-withdrawal of allocated treatment and time-to-first seizure (carbamazepine (CBZ) reference) for individuals with generalised seizures 314
Figure 37: Time-to-withdrawal of allocated treatment and time-to-first seizure (gabapentin (GBP) reference) for individuals with generalised seizures 315
Figure 38: Time-to-withdrawal of allocated treatment and time-to-first seizure (levetiracetam (LEV) reference) for individuals with generalised seizures 316
Figure 39: Time-to-withdrawal of allocated treatment and time-to-first seizure (lamotrigine (LTG) reference) for individuals with generalised seizures 317
Figure 40: Time-to-withdrawal of allocated treatment and time-to-first seizure (oxcarbazepine (OXC) reference) for individuals with generalised seizure 318

Figure 41: Time-to-withdrawal of allocated treatment and time-to-first seizure (phenobarbitone (PHB) reference) for individuals with generalised seizures 319
Figure 42: Time-to-withdrawal of allocated treatment and time-to-first seizure (phenytoin (PHT) reference) for individuals with generalised seizures 320
Figure 43: Time-to-withdrawal of allocated treatment and time-to-first seizure (topiramate (TPM) reference) for individuals with generalised seizures 321
Figure 44: Time-to-withdrawal of allocated treatment and time-to-first seizure (sodium valproate (VPS) reference) for individuals with generalised seizures 322
Figure 45: Time-to-withdrawal of allocated treatment and time-to-first seizure (carbamazepine (CBZ) reference) for individuals with partial seizures 323
Figure 46: Time-to-withdrawal of allocated treatment and time-to-first seizure (gabapentin (GBP) reference) for individuals with partial seizures 324
Figure 47: Time-to-withdrawal of allocated treatment and time-to-first seizure (levetiracetam (LEV) reference) for individuals with partial seizures 325
Figure 48: Time-to-withdrawal of allocated treatment and time-to-first seizure (lamotrigine (LTG) reference) for individuals with partial seizures 326
Figure 49: Time-to-withdrawal of allocated treatment and time-to-first seizure (oxcarbazepine (OXC) reference) for individuals with partial seizures 327
Figure 50: Time-to-withdrawal of allocated treatment and time-to-first seizure (phenobarbitone (PHB) reference) for individuals with partial seizures 328
Figure 51: Time-to-withdrawal of allocated treatment and time-to-first seizure (phenytoin (PHT) reference) for individuals with partial seizures 329
Figure 52: Time-to-withdrawal of allocated treatment and time-to-first seizure (topiramate (TPM) reference) for individuals with partial seizures 330
Figure 53: Time-to-withdrawal of allocated treatment and time-to-first seizure (sodium valproate (VPS) reference) for individuals with partial seizures 331
Figure 54: Time-to-withdrawal of allocated treatment and time-to-first seizure (zonisamide (ZNS) reference) for individuals with partial seizures 332
Figure 55: NMA results of IPD reduced to summary statistics: individuals with partial seizures 333
Figure 56: NMA results of IPD combined with AD: individuals with partial seizures 334
Figure 57: NMA results of IPD reduced to summary statistics: individuals with generalised seizures 335
Figure 58: NMA results of IPD combined with AD: individuals with generalised seizures 336
Figure 59: NMA results of IPD reduced to summary statistics for all individuals regardless of epilepsy type 337
Figure 60: NMA results of IPD combined with AD for all individuals regardless of epilepsy type 338

Abbreviations

AD	Aggregate data
AD-MA	Aggregate Data Meta-Analysis
AED	Antiepileptic drug
CBZ	Carbamazepine
CI	Confidence Interval
CONSORT	Consolidated standards of reporting trials
CSDR	Clinical Study Data Request.com
EQUATOR	Enhancing the Quality and Transparency of Health Research
GBP	Gabapentin
GSK	Glaxo Smith Kline
HR	Hazard Ratio
ILAE	International League against Epilepsy
IPD	Individual Participant Data
IPD-MA	Individual Participant Data Meta-Analysis
IPD-NMA	Individual Participant Data Network Meta-Analysis
ITT	Intention to treat
J&J	Johnson & Johnson
KM	Kaplan-Meier
LEV	Levetiracetam
Ln (HR)	Logarithm of the Hazard Ratio
LTG	Lamotrigine
NICE	National Institute for Health and Care Excellence
NMA	Network Meta-Analysis
OR	Odds Ratio
OXC	Oxcarbazepine

PH	Proportional Hazards
PHB	Phenobarbitone
PHT	Phenytoin
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	Randomised Controlled Trial
RMST	Restricted Mean Survival Time
RR	Risk Ratio / Relative Risk
SD	Standard Deviation
SE	Standard Error
SOAR	Supporting Open Access to Researchers
TPM	Topiramate
TTE	Time-to-event
UK	United Kingdom
US	United States
VPS	Sodium Valproate
YODA	Yale University Open Data Access
ZNS	Zonisamide

Initials of researchers involved in the work in this thesis

AGM	Anthony G Marson
BD	Becky Davie
CTS	Catrin Tudur Smith
JW	Jennifer Weston
LS	Laura Sutton
LW	Lisa Williams
MS	Maria Sudell
SJN	Sarah J Nevitt (née Nolan)
SR	Sally Reynolds

Publications and presentations of work in this thesis

Chapter 2

Early findings of the literature review contained in Chapter 2 were presented at a conference:

Nolan SJ, Marson AG, Tudur Smith C. A systematic review of methodology developed for meta-analyses with time-to-event outcomes. Conference proceedings. Survival analysis for junior researchers. University of Leicester, United Kingdom, April 2012; page 29.

Chapter 3

The work contained in Chapter 3 has been presented at several conferences and meetings:

Nolan SJ, Marson AG, Tudur Smith C. *Meta-analysis of time-to-event data: consistency of outcome and statistical reporting in epilepsy trials.* Conference proceedings. Survival analysis for Junior Researchers. University of Warwick, United Kingdom, 3rd-4th April 2014; page 26

Nolan SJ, Marson AG, Tudur Smith C. *Meta-analysis of time-to-event outcomes: Issues in epilepsy trials*. Methods in Meta-Analysis. London, United Kingdom, December 2013.

Nolan SJ, Sutton L, Marson AG, Tudur Smith C. *Consistency of outcome and statistical reporting of time to event data: the impact on Cochrane reviews and meta-analyses in epilepsy*. Conference proceedings. 21st Cochrane Colloquium, Université Laval. Quebec City, Canada, 19th -23rd September 2013, page 99.

Nolan SJ, Marson AG, Tudur Smith C. *Methodology developed for meta-analyses with time-to-event outcomes and consistency of outcome reporting.* Conference proceedings. Survival analysis for Junior Researchers. University of Liverpool, United Kingdom, 25th-26th March 2013; page 28

Chapter 4 and Chapter 5

The work contained in Chapter 4 and Chapter 5 has been published in the British Medical Journal:

Nevitt SJ, Marson AG, Davie B, Reynolds S, Williams L, Tudur Smith C. Exploring changes over time and characteristics associated with data retrieval across individual participant data meta-analyses: systematic review. BMJ 2017;357:j1390

The work and themes within these Chapters have also been presented at many conferences and meetings (invited speaker presentations indicated with *)

*Nolan SJ. *Data Sharing: a young statistician's perspective.* Centre for Biostatistics Seminar, University of Manchester, Manchester, United Kingdom; 18 March 2016

*Nolan SJ. *Data Sharing: a young statistician's perspective*. Royal Statistical Society Merseyside Local Group meeting. University of Liverpool, Liverpool, United Kingdom; 26 February 2016

Nolan SJ, Marson AG, Tudur Smith C. *The changing world of data sharing and data transparency: what does this mean for individual participant data reviews?* Conference

Proceedings, 23rd Cochrane Colloquium, Messe Congress Center, Vienna, Austria. 3rd-7th October 2015

*Nolan SJ. *Data Transparency: To Infinity and Beyond*. PSI (Statisticians in the Pharmaceutical Industry) annual conference. London, United Kingdom; 14 May 2015

*Nolan SJ. *ClinicalStudyDataRequest.com and the SAS Data Access System: An Academic Researcher's Experience*. The 5th Clinical Trial Data Transparency Forum. SAS Germany Headquarters; Heidelberg, Germany; 23 April 2015

Nolan SJ. *Data Sharing: is it getting easier to access Individual Participant Data? Experiences from the Cochrane Epilepsy Group*. North West Epilepsy Group Meeting. Aintree Racecourse, Liverpool. 5th November 2014.

*Nolan SJ and Tudur Smith C. *Understanding that Data Transparency matters: The academic view*. The 10th Annual Conference: Data Transparency. Pharmaceutical Users Software Exchange (PhUSE), London, United Kingdom; 12-15th October 2014.

Nolan SJ, Marson AG, Tudur Smith C. *Data sharing: Is it getting easier to access Individual Participant Data? Experiences from the Cochrane epilepsy group*. Conference Proceedings, 22nd Cochrane Colloquium, Hyderabad International Convention Centre, Hyderabad, India. 21st – 26th September 2014

Nolan SJ, Marson AG, Tudur Smith C*. Meta – analysis of time to event data: what can we do about missing individual participant data?* Conference Proceedings. The International Conference of the Royal Statistical Society. Sheffield, United Kingdom, 1-4th September 2014, page 94.

*Nolan SJ and Tudur Smith C. *An academic researcher's perspective on data transparency*. Clinical Trial Disclosure and Transparency, Pharmaceutical Users Software Exchange (PhUSE), London, United Kingdom; 22nd May 2014.

*Nolan SJ and Tudur Smith C. *An academic researcher's perspective on data transparency*. The 3rd Clinical Trial Data Transparency Forum. SAS UK Headquarters, Marlow, Buckinghamshire, United Kingdom; 29 April 2014.

Chapter 6

The protocol of the Cochrane Individual Participant Data Network Meta-Analysis which has inspired much of the work of this thesis (and results of this Cochrane IPD-NMA are presented in Chapter 6) is published on the Cochrane Database of Systematic Reviews:

Nolan SJ, Sudell M, Weston J, Tudur Smith C, Marson A. *Antiepileptic drug monotherapy for epilepsy: an overview of systematic reviews and network meta-analysis (protocol).* Cochrane Database of Systematic Reviews 2014, Issue 12, Art. No.: CD011412. DOI: 10.1002/14651858.CD011412.

Nevitt SJ, Sudell M, Weston J, Tudur Smith C, Marson A. *Antiepileptic drug monotherapy for epilepsy: an overview of systematic reviews and network meta-analysis.* Cochrane Database of Systematic Reviews 2017 2014, Issue 6, Art. No.: CD011412. DOI: 10.1002/14651858.CD011412.

The following presentations have been given on the clinical findings of the Cochrane IPD-NMA (invited speaker presentations indicated with *)

Nevitt, SJ, Sudell M, Weston J, Tudur Smith C, Marson A. *Antiepileptic drug monotherapy: a network meta-analysis of 10 AEDs*. Annual meeting of the Association of British Neurologists 2017. Liverpool, United Kingdom. 3 May 2017

*Nevitt SJ. *Antiepileptic Drug monotherapy: A Cochrane Review and Network Meta-Analysis of 10 AEDs*. Walton Neuroscience Forum. Aintree, Liverpool, United Kingdom. 25 January 2017.

Related Cochrane Individual Participant Data Meta-Analyses are also published on the Cochrane Database of systematic reviews:

Nevitt SJ, Marson AG, Weston J, Tudur-Smith C. *Carbamazepine versus phenytoin monotherapy for epilepsy: an individual participant data review*. Cochrane Database of Systematic Reviews 2017, Issue 2. Art No.: CD001911. DOI: 10.1002/14651858.CD001911.pub3

Nolan SJ, Sudell M, Tudur Smith C, Marson A. *Topiramate versus carbamazepine for epilepsy: an individual participant data review*. Cochrane Database of Systematic Reviews 2016, Issue 12, Art. No: CD012065. DOI: 10.1002/14651858.CD012065.pub2

Nolan SJ, Tudur Smith C, Weston J, Marson AG. *Lamotrigine versus carbamazepine monotherapy for epilepsy: an individual participant data review*. Cochrane Database of Systematic Reviews 2016, Issue 11, Art No. CD001031. DOI: 10.1002/14651858.CD001031.pub3

Nolan SJ, Marson AG, Weston J, Tudur-Smith C. *Phenytoin versus valproate monotherapy for partial-onset seizures and generalised-onset tonic-clonic seizures: an individual participant data review.* Cochrane Database of Systematic Reviews 2016, Issue 4. Art. No: CD001769 DOI: 10.1002/14651858.CD001769.pub3.

Nolan SJ, Sudell M, Tudur Smith C, Marson AG. *Topiramate versus carbamazepine for epilepsy: an individual participant data review (Protocol).* Cochrane Database of Systematic Reviews 2016, Issue 1. Art. No: CD012065. DOI: 10.1002/14651858.CD012065.

Nolan SJ, Marson AG, Weston J, Tudur Smith C. *Carbamazepine versus phenobarbitone monotherapy for epilepsy: an individual participant data review*. Cochrane Database of Systematic Reviews 2015, Issue 7. Art. No.: CD001904. DOI: 10.1002/14651858.CD001904.pub2.

Nolan SJ, Muller M, Tudur Smith C, Marson AG. *Oxcarbazepine versus phenytoin monotherapy for epilepsy.* Cochrane Database of Systematic Reviews 2013, Issue 5. Art. No: CD003615 DOI: 10.1002/14651858.CD003615.pub3.

Nolan SJ, Tudur Smith C, Pulman J, Marson AG. *Phenobarbitone versus phenytoin monotherapy for partial-onset seizures and generalised-onset tonic-clonic seizures*. Cochrane Database of Systematic Reviews 2013, Issue 1. Art. No: CD002217 DOI: 10.1002/14651858.CD002217.pub

Chapter 1: Introduction

1.1 Evidence synthesis

Evidence synthesis is a general term used to describe techniques to combine sources of quantitative evidence. The formulation of the research question of interest in a clinical setting for any evidence synthesis requires careful attention; a research question must be specific enough for results to be clinically useful but not too specific so that inadequate amounts of evidence are available [1]. A commonly applied analogy to this decision is the choice of whether to 'lump' or to 'split' [2]; in other words, whether to take a broad approach to a wide variety of settings and participant groups or whether to narrow a research question into a homogenous evidence base [3].

In a clinical setting, where interventions and treatment effects are of interest, clinical assumptions underlying a synthesis must be considered as closely as statistical assumptions [4]. It is unlikely that a treatment effect would be replicated exactly in two clinical studies due to variations in participant populations and settings. However if an intervention does provide true benefit over another then one would expect the direction of effect to be the same in a range of heterogeneous situations [5]. This true direction of treatment effect is more likely to stand out in a synthesis when a number of studies are considered together.

The techniques of evidence synthesis of relevance to this thesis are systematic review, meta-analysis and network meta-analysis which are introduced in the following sections.

1.1.1 Systematic reviews and meta - analysis

Systematic reviews are commonly used as a means of summarising the results of all independent sources of evidence which address the same or similar questions in a systematic way [6, 7]. Systematic reviews of randomised controlled trials are widely accepted to provide the highest quality inferences in evidence based medicine [8]. However the quality of a systematic review or any synthesis is dependent on the completeness of the evidence [1].

Meta-analysis is a statistical technique used to synthesise the results of each study included in the systematic review to obtain a single pooled result which gives an overall relative treatment effect of one treatment to another [9]. The application of this technique increases sample size and may increase precision and power, minimising the likelihood of a chance

result and providing more information regarding treatment effects which single studies do not have the power to detect [7, 10].

In general terms, meta-analyses can be performed with fixed-effects assuming a baseline risk of μ_i and a common fixed underlying treatment effect δ across all i studies and within-study error ε_i or with random-effects assuming the same baseline risk μ_i and within-study error ε_i of the fixed-effects model but random systematic differences δ_i in between trial results due to study heterogeneity τ^2. In other words, pooled treatment effect Y can be estimated as:

Fixed-effects: $$Y = \mu_i + \delta + \varepsilon_i \quad \text{(Equation 1)}$$

Random-effects $$Y = \mu_i + \delta_i + \varepsilon_i \qquad \delta_i \sim (\delta, \tau^2) \quad \text{(Equation 2)}$$

where δ_i is sampled from a distribution with mean δ and variance τ^2. A comprehensive guide to meta-analytic methods for different data types and in both Frequentist and Bayesian settings is provided by Sutton *et al* [11].

Origins of heterogeneity in meta-analysis include variation in study design and methodology, clinical settings and participant characteristics including baseline risk between trials within the meta-analysis leading to variation in sampling error [11, 12]. More specifically, within a time-to-event context (see Chapter 1.1.4 for further details of time-to-event data), sources of heterogeneity include time-dependent (non-constant) treatment effects and variation in length of follow-up time across trials [13].

1.1.2 Indirect comparisons and network meta - analysis

The framework of a traditional 'pairwise' meta-analysis can consider only two interventions (or classes of interventions) head-to-head. However, within clinical settings for which a large range of intervention options are available, some of which may never have been compared directly in a clinical trial, pairwise meta-analysis cannot provide an adequate estimate of the relative effectiveness of all interventions of interest to aid medical decision making [14].

Network meta-analysis (also referred to a multiple treatment meta-analysis or mixed treatment comparison; referred to as 'network meta-analysis' (NMA) herein for consistency of terminology) provides a framework for the synthesis of direct evidence for interventions *A* and *B* from their head-to-head comparison within clinical trials and indirect evidence for the same interventions *A* and *B* deduced via a direct comparison to a common intervention *C* (see Figure 1).

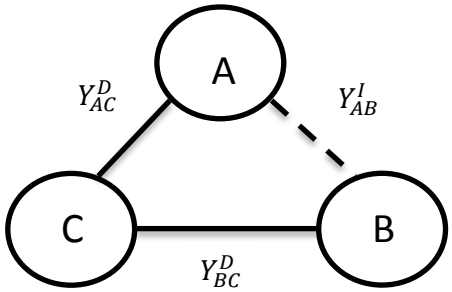


Figure 1: Direct and indirect evidence from the network of interventions A, B and C

In other words, if Y_{AC}^{D} and Y_{BC}^{D} denote the pooled estimates from the synthesis (via fixed or random-effects analyses) of all direct evidence of *A* versus *C* and *B* versus *C* respectively, the indirect estimate for *A* versus *B* (Y_{AB}^{I}) is calculated as [15]:

$$Y_{AB}^{I} = Y_{AC}^{D} - Y_{BC}^{D} \qquad \text{(Equation 3)}$$

An additional benefit of indirect comparison under the framework of a connected 'network' of interventions, such that each intervention in the network has been compared to at least one other intervention in the network directly, is that an estimate for every pairwise comparison within the network can be calculated via a combination of direct and indirect evidence without the necessity for a common 'control' intervention (e.g. direct evidence for each intervention versus placebo) across all trials [16, 17].

Indirect comparisons are also valuable where a limited amount of data is available to inform a direct comparison or where evidence informing a direct comparison is of poor methodological quality. The power and precision of a treatment effect estimate can be increased by "borrowing strength" from the indirect evidence within the network [15].

An underlying assumption of indirect comparisons and NMA is that any intervention effect is 'exchangeable' across all included trials [14]; in other words, the indirect comparison between two interventions is a feasible one to make (known as the transitivity assumption) and that the indirect evidence is consistent with the direct evidence where a comparison exists (known as the consistency assumption). Transitivity requires that all treatments are "jointly randomisable;" in other words, all interventions within a network could feasibly be randomised in the same trial and those which are not treatment arms in any given trial are "missing at random"[17]. Such an assumption cannot be formally tested statistically; transitivity must be judged by careful consideration of trial settings and characteristics, treatment mechanisms and participant demographics to investigate if any differences would be expected to modify relative treatment effects.

The consistency assumption can be evaluated statistically over a closed 'loop of evidence' where both direct and indirect evidence exists for a comparison. Inconsistency may be present in NMA for a number of reasons [18]; a common source of inconsistency is thought to be an imbalance in treatment effect modifiers across comparisons as randomisation does not hold across a network of studies [19, 20]. Inclusion of treatment-covariate interaction terms within NMA models may reduce this confounding bias and in turn reduce inconsistency within the network [19-21].

Further statistical concepts and models for NMA, including models to account for inconsistency are discussed in Salanti *et al* [3] and Efthimiou *et al* [18].

1.1.3 Aggregate data and individual participant data synthesis

The most common approach to quantitative synthesis is undertaken using aggregate data (AD); an approach where summary statistics such as mean differences, event counts, risk ratios, odds ratios, hazard ratios etc. are extracted from published literature and can be supplemented with unpublished information provided by the original trialists.

The alternative approach, an individual participant data (IPD) analysis, where participant-level data containing detailed demographic, baseline and outcome data is retrieved and re-analysed, is widely regarded as the gold standard approach to the synthesis of study results [22, 23]. Analysis of participant-level data has many advantages over traditional AD meta-analysis (AD-MA); allowing a more standardised, comprehensive and potentially more methodologically complex approach to target clinical questions within participant subgroups via treatment-covariate interactions [24, 25]. Theoretically, an IPD approach should also reduce publication, reporting and ecological biases often associated with AD-MA [26-28]. IPD meta-analyses (IPD-MAs) have been shown to directly influence the design and conduct of clinical trials [29] and in some contexts to impact upon clinical practice guidelines [30]. Furthermore in some settings, particularly where the necessary published information to perform AD-MA is not reported or is reported inconsistently [6, 31], a re-analysis of IPD may be the only feasible approach to synthesis (see Chapter 3 for further discussion).

IPD syntheses have been performed since the 1980s [23], with most early IPD-MAs limited in their statistical methodology to estimation of overall treatment effect [32]. Recent years have shown a sharp increase in the number of IPD-MAs [33-35], with an average of 49 published per year between 2005 and 2009 [34] and recent estimates suggest an increase

of around four published IPD-MAs per year [35]. Development of methodology for the synthesis of IPD has also increased [36]. For example, in addition to models for the estimation of overall treatment effect with and without treatment covariate interactions [24], methodology is now available to explore heterogeneity within and between studies [37], examine within-study and across-study associations[25], to combine IPD with AD [38-41] and to perform prognostic or diagnostic modelling [42].

AD-MA make up the vast majority of the meta-analysis literature. Previous work has indicated that up to 2004, less than 10% of published meta-analyses per year used IPD (in fact, for most years the figure was less than 5%) [43]. An IPD approach is also still relatively rare within NMA compared to an AD approach [18, 19, 44-48]. A basic MEDLINE search conducted in August 2016 by SJN found that out of 71539 records indexed as 'meta-analysis,' only 1073 (1.5%) included 'individual patient data' or 'individual participant data' from a title, abstract and keyword search. This figure should be treated as approximate; but it does imply that up to 2016, despite the increase in the number of published IPD-MAs, such an approach is still only used in a very small proportion of the meta-analysis literature.

The aim of any systematic synthesis is to obtain all relevant information from all eligible participants. However in practice for an IPD approach to synthesis, retrieving all participant data can require a considerable amount of time, cost and personnel and can be computationally intensive in the case of large individual participant datasets [22, 49]. Furthermore, statistical expertise will often be required to perform more methodologically complex analyses proposed in IPD-MAs. Lack of time, resources and statistical expertise may explain the apparent preference for an AD approach to meta-analysis over an IPD approach; interface based meta-analysis software, such as Review Manager from the Cochrane Collaboration [50], facilitates AD-MA techniques for non-statisticians while undertaking an IPD-MA is likely to require some level of statistical programming.

Additionally, retrieval of all relevant data is not always possible. IPD may have been destroyed or lost, trialists may be unwilling to collaborate due to confidentiality of data etc. and only a proportion of IPD is available for re-analysis. Synthesis with published AD may be required for a complete analysis or studies may have to be excluded from synthesis entirely [38-40, 44-46, 51]. Conversely, combining IPD with AD estimates in a synthesis may not be deemed acceptable to trialists who put in effort to provide IPD [22], but excluding studies which do not provide IPD violates a major assumption of systematic syntheses that all

relevant evidence has been included. Further discussion of retrieval of IPD for evidence synthesis and the impact of missing IPD on analysis is described in Chapters 4 to 7.

Previous work has demonstrated that meta-analyses based on aggregated results from published literature can give different results to meta-analysis using IPD from the same studies [23, 43, 52-54]. IPD and AD approaches to meta-analysis, when based on identical data from homogeneous studies, should produce theoretically identical results [55, 56], however it is unlikely the data used for each approach would be identical and factors contributing to any observed differences in results are publication bias, patient exclusions, length of follow-up and method of analysis. Conclusions of previous work recommend that if IPD-MA and AD-MA can be shown to be mathematically equivalent, and adequate AD is available, the resource savings associated with an AD-MA would make this the approach of choice [52, 53], but in all other settings, wherever possible an IPD approach is favourable and will provide the least biased and most reliable means of addressing the clinical question [23].

1.1.4 Time-to-event data

Time-to-event (TTE) data, also referred to as survival data (often within an oncology setting) or failure-time data (often within an engineering setting), arise when interest lies not only in whether an event of interest occurs but also the time taken for that event to occur from a well-defined time origin. Examples of outcomes measured using TTE data include time-to-death following diagnosis of cancer, time-to-remission of epileptic seizures after randomisation into a clinical trial and time-to-conception following fertility treatment.

Specific statistical methodology is required for the analysis of TTE data for two reasons. Firstly, TTE data is non-symmetrical, usually positively skewed, therefore methodology for continuous data which assumes a normal approximation is violated. Secondly, TTE data is frequently censored in that an event is not observed for an individual therefore they cannot contribute an event time towards an analysis; however a censoring time (i.e. the length of time that individual was known to be event free for) can contribute to analysis.

Theoretically, the survival and hazard functions are used in the statistical analysis of TTE data. The hazard function $h(t)$ is defined as the instantaneous risk of an event at time t, given that an individual has been event free up to time t, $H(t)$ is the cumulative hazard function, defined as the sum of instantaneous hazards up to time t, and the survival function $S(t)$

describes the probability of being event-free up to time t, essentially a cumulative history of events. The functions are directly linked as follows (where T represents the event time) [57]:

$$S(t) = \Pr(T > t) = \exp[-H(t)] = \exp\left[-\int_0^t h(u)du\right] \quad \text{(Equation 4)}$$

The hazard functions $h_T(t)$ and $h_C(t)$ and the corresponding survivor functions $S_T(t)$ and $S_C(t)$ of the treatment (*T*) and control (*C*) groups respectively are linked by the hazard ratio parameter θ:

$$h_T(t) = \theta h_C(t) \quad \text{(Equation 5)}$$

$$S_T(t) = S_C(t)^\theta \quad \text{(Equation 6)}$$

The nature of the hazard function makes it more flexible for modelling than the survival function [57]. The hazard function can be related multiplicatively to participant characteristics via a semi-parametric proportional hazards model as follows for i individuals, with common baseline hazard $h_0(t)$, $x_1 \ldots x_n$ characteristics and $\beta_1 \ldots \beta_n$ as follows [58]:

$$h_i(t) = h_0(t)exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n) \quad \text{(Equation 7)}$$

A comprehensive guide to modelling of TTE data in parametric, semi- and non-parametric settings, under the assumptions of proportional or time dependent hazards is provided by Collett [57] and further details of synthesis methods for TTE data are presented in Chapter 2.

1.2 Individual participant data meta-analyses and network meta-analyses in Epilepsy

To date, the Cochrane Epilepsy Group have performed nine IPD-MAs of antiepileptic drug (AED) monotherapy trials [59-67] and an IPD-NMA of 10 different AEDs [68, 69].

1.2.1 Clinical Setting

Epilepsy is a common neurological condition in which recurrent, unprovoked seizures are caused by abnormal electrical discharges from the brain. Epilepsy is a disorder of many heterogeneous seizure types and accounts for approximately 0.75% of the global burden of disease [70]. It is believed that with effective drug treatment, up to 70% of individuals with active epilepsy have the potential to become seizure-free and go into long-term remission shortly after starting drug therapy and that around 70% of these individuals can achieve seizure-freedom using a single AED in monotherapy [71].

Cochrane IPD reviews have considered two epilepsy types for which monotherapy is indicated: generalised-onset (generalised tonic-clonic) seizures in which electrical discharges begin in one part of the brain and move throughout the brain, and partial-onset seizures in which the seizure is generated in and affects one part of the brain (the whole hemisphere of the brain or part of a lobe of the brain). The reviews examine ten AEDs which are currently licensed and used in clinical practice for use as monotherapy for partial and generalised seizures in at least one country [72, 73]:

Carbamazepine (CBZ), Phenytoin (PHT), Phenobarbitone (PHB), Sodium Valproate (VPS), Lamotrigine (LTG), Oxcarbazepine (OXC), Topiramate (TPM), Gabapentin (GBP), Levetiracetam (LEV), Zonisamide (ZNS).

Current guidelines from the National Institute for Health and Care Excellence (NICE) for adults and children recommend carbamazepine or lamotrigine as first-line treatment for partial-onset seizures and valproate for generalised-onset seizures, on the condition that females of childbearing age are made aware of the potential teratogenic effects of the drug [74, 75]. Clinical profiles and mechanisms of action of these ten drugs are detailed in the Cochrane Epilepsy IPD-NMA [68, 69].

1.2.2 Rationale for individual participant data approach

With evidence that up to 70% of individuals with active epilepsy have the potential to go into long-term remission of seizures shortly after starting drug therapy [71], the correct choice of first-line AED therapy for individuals with newly diagnosed seizures is of great importance. There are currently over 30 drugs (over 50 generic and branded formulations) available worldwide for the treatment of various seizure types [76], therefore it is important that the choice of AEDs for an individual is made using the highest quality evidence regarding potential benefits and harms of treatments appropriate to given seizure types.

The design and choice of outcome for epilepsy monotherapy trials is discussed at length in Chapter 3.1.2. In summary, the important efficacy outcomes defined in epilepsy monotherapy trials and used in Cochrane reviews of epilepsy monotherapy often require analysis of TTE data (for example, time-to-first seizure after randomisation or time-to-withdrawal of allocated treatment).

Methods have been developed to synthesise TTE data using summary information [6, 13] (also see Chapter 2.3.2 for further discussion of these methods); however, the appropriate

statistics to perform a direct synthesis or indirect estimation are not commonly published in epilepsy monotherapy trials (see Chapter 3 for further discussion).

Furthermore, although seizure data have been collected in most epilepsy monotherapy trials, there is little uniformity in the definition and reporting of outcomes (see Chapter 3 for further details). For these reasons, an IPD approach has been taken for pairwise meta-analysis and proposed for network meta-analyses in Cochrane reviews of epilepsy monotherapy. This approach helps to overcome issues around inconsistent reporting in trials and is considered to be the 'gold standard' approach to synthesis of TTE data [6].

Each Cochrane pairwise meta-analysis provides high quality evidence for each pair of drugs but does not inform a choice between the whole evidence base of appropriate drugs for decision makers, clinicians or individuals with epilepsy. Furthermore, direct evidence from randomised controlled trials (RCTs) is not available between some of the commonly used AEDs (such as between OXC and PHB) and due to current first-line treatment recommendations, it is unlikely that an RCT will be designed in the future that will make a direct comparison between such drugs [74], so it is not possible to make pairwise comparisons of treatment effects between all drugs.

Use of NMA allows for indirect pairwise comparison of all drugs in the network, in other words indirect treatment effect estimates can be calculated for all 45 pairwise comparisons of the ten AEDs of interest in the Cochrane reviews of epilepsy monotherapy. Indirect evidence from the network can also increase the power and precision of direct treatment effect estimates as discussed in Chapter 1.1.2.

1.3 Individual participant data sharing: recent initiatives

Recent years have seen change in attitudes to sharing of clinical trial data with many calls for improved data transparency and data sharing initiatives introduced across the research community as a whole [77-83], in addition to the publication of data transparency policies by the Institute of Medicine [84], the European Medicines Agency [85], and the International Committee of Medical Journal Editors [86, 87]. Support of clinical trial data sharing seems to be improving in recent years, with a reported increased willingness of authors of published trials to share data in surveys conducted in 2011 [77, 88] compared to an empirical study conducted in 2009 [89]. A recent survey of patients within a U.S. emergency department has also shown that the majority of patients were in favour of de-identified clinical trial data

sharing in general, however 25% of patients indicated that they would be less likely to participate in a clinical trial if data were shared, demonstrating the increased importance of the informed consent process during trials [90].

The launch of the first data sharing platform, led by GlaxoSmithKline (GSK) in May 2013 [80], marked the beginning of a new era of data transparency within the Pharmaceutical Industry. At the time of writing (July 2017), thirteen pharmaceutical sponsors have committed to the sharing of IPD from nearly 3500 of their trials with independent researchers via multi-sponsor platform Clinical Study Data Request (CSDR)[91]. CSDR provides a structured format for requesting data, including a step-by-step diagram, user guide, supporting guidance videos and the opportunity to communicate with the sponsor throughout the process. The data sharing process for requests submitted to CSDR is as follows:

- Selection of studies of interest from a list by the sponsor(s) or submit an enquiry for any studies of a relevant sponsor not currently listed. Multiple studies from a single sponsor or from different sponsors can be included in a single request (single sponsor request or multi-sponsor request).
- Submission of a research proposal for the studies required for the research, including statistical analysis plan, publication plan and conflicts of interest for review by an independent review panel.
- Signing of data sharing agreements by the researcher and sponsor for approved proposals.
- Sponsor undertakes the de-identification of participant-level data and related documentation (e.g. protocols, case report forms) in preparation for sharing of data (see Hrynaszkiewicz *et al* [92] for a summary of the principles of data de-identification).
- Remote access to requested de-identified datasets and all related documentation is provided via a secured SAS analytic environment.

Other sponsors have opted for a single sponsor environment in contrast to the multi-sponsor format of CSDR. For example, Johnson & Johnson (J&J) announced an agreement with Yale University School of Medicine's Open Data Access (YODA) project in January 2014 in which YODA acts as an independent review panel for research proposals requesting access to 191 J&J datasets (correct to July 2017) in a similar format to CSDR [93, 94]. More recently in 2016, Bristol Myers Squib announced a collaboration with the Clinical Research Institute of Duke University for the Supporting Open Access to Researchers (SOAR) initiative, a similar data

sharing model to the YODA process [95]. Both CSDR and YODA provide metrics related to submitted research proposals on their respective websites [91, 93]; see Chapter 5.5.2 for further discussion of these metrics.

These changes and initiatives may go some way to reducing some barriers to conducting IPD-MAs (as described in Chapter 1.1.3 and further discussed in Chapters 4 to 7). However, concerns have been raised regarding practical challenges of data sharing and the potential impact of data sharing on clinical trial participation [96] and recent editorials have suggested that current initiatives do not go far enough in the commitment to data sharing [97, 98].

However, the benefits and impact of such initiatives and reported changes in attitude towards data sharing on IPD analyses may not become clear for some time; a recent study has shown that out of over 3000 trials available for data request via CSDR, YODA or SOAR, only 15.5% of these studies have been requested by a limited number of researchers [99].

Furthermore, at the time of writing, out of over 350 research proposals submitted to and reviewed by CSDR and YODA up to July 2017, results from only ten data requests have been published to date [69, 100-108]. This slow publication rate may reflect the originality and complexity of the research hypotheses proposed. While the provision of access to original datasets may have previously been associated with re-analysis to confirm validity of trial results [109], the focus of the majority of approved research proposals seems to be original research. Published titles of approved CSDR and YODA research proposals range from IPD-MAs, development of prognostic, pharmacokinetic and genetic models, the development of novel statistical methodology, investigation of adverse drug events and the design of new randomized controlled trials. Such projects are likely to take several years to reach final publication stage.

Furthermore, the researcher's perspective of using CSDR has received mixed reviews highlighting that access to IPD could be taking longer due to the additional safeguards that have been put in place [96, 110-113]. Restrictions relating to the remote analysis of IPD has prevented the completion of some projects, as highlighted via the reported metrics of the status of data requests of the YODA project [93]. The full extent of such restrictions and the impact of the provision of highly de-identified data on the range of clinical questions that can be addressed and statistical methodology which can be employed within data sharing platforms is currently unknown [114].

1.4 Thesis objective and structure

Evidence syntheses are commonly used and highly regarded techniques for the quantitative summary of evidence from a number of sources [8, 34]. The undertaking of an IPD synthesis can be methodologically complex, time consuming and open to sources of bias if not conducted rigorously [51]. While IPD syntheses published as journal articles will present clinical results and implications and some level of detail regarding methodology, it is difficult to translate the complete rationale and the practical challenges of such an approach to synthesis within the structure and word limit of a journal manuscript.

The aim of this thesis is to document the practical aspects and challenges of conducting IPD-MAs and an IPD-NMA of AED monotherapy and to provide a wider context to these syntheses in the literature of AD and IPD evidence synthesis. Such experiences and challenges are of particular relevance in current research environment with changing functionalities of and attitudes towards data sharing.

Chapter 2 summarises methodology for meta-analysis of TTE data from both an AD and IPD perspective. Chapter 3 summarises previous reviews of the reporting of TTE data in the published literature and presents a novel systematic review of the reporting of aggregate TTE outcomes and analyses specifically in epilepsy AED monotherapy trials published to 2012.

Chapter 4 then presents a novel systematic review of previously conducted IPD-MAs to explore data retrieval rates and characteristics associated data retrieved for all published IPD-MAs from 1987 to 2015. Chapter 5 then reflects upon the experiences of the Cochrane Epilepsy Group in previous data requests made between 1995 and 2005 compared to data requests made for current projects between 2011 and 2015; a time frame which has seen substantial changes to the way clinical trial data, particularly pharmaceutical sponsored trial data, is shared with independent researchers.

Chapter 6 presents statistical methodology and results of several approaches to modelling the association between treatment effect and epilepsy type in the IPD-NMA and Chapter 7 then presents further results and clinical implications of the Cochrane IPD-NMA of AED monotherapy described in Chapter 1.1.2. Chapter 8 then presents statistical methodology, for combining published AD with IPD in NMA as well as results and clinical implications of incorporating additional published AD with IPD in this example of AED monotherapy.

The final chapter summarises the findings of the previous chapters, reflects upon the implications of these findings in the clinical context and in the context of IPD and AD evidence synthesis and provides discussion of further research needed.

Chapter 2: Meta-analysis of time-to-event data: a methodology review

2.1 Introduction

As introduced in Chapter 1.1.4, in many clinical settings, the outcome of interest is the time to an event. The use of time-to-event (TTE) outcomes is particularly common in the field of oncology, where the aim of a clinical trial is often to detect a modest treatment effect [1, 4, 115]; so unless a single study is very large in sample size, it is unlikely to detect the small treatment effect due to insufficient statistical power. As cancer is a relatively common condition, a moderate treatment effect could have substantial benefit for public health. Meta-analyses are commonly used in the field of oncology. For example, the Early Breast Cancer Trialists Collaborative Group have collected IPD from over 450,000 women from 400 randomised trials from the last 30-40 years and found modest but highly significant survival benefit for a range of chemotherapy, hormonal and radiotherapy regimens [4, 116, 117]. Meta-analysis of TTE data is also important in other settings; as outlined in Chapter 1.2, the Cochrane Epilepsy Group have performed or are in the process of performing nine IPD-MAs and an NMA of AED monotherapy trials [59-67, 69], within which important efficacy outcomes such as 'time-to-withdrawal of allocated treatment' and 'time-to-12-month remission from seizures' required the analysis of TTE data.

A range of methods for the meta-analytic synthesis of TTE data have been proposed over several decades, with the applicability of specific methods depending on the clinical question of interest, the assumed hazard distribution and the type of data available for each study for meta-analysis [11].

The objective of this Chapter is to provide a summary of existing methodology developed for meta-analysis of TTE data with a particular focus on the level of data (individual level or aggregate-level) required for that method. Where possible, applications of developed methods (via illustrative examples or simulation studies) will also be discussed to highlight the relative advantages and disadvantages of developed methods within given clinical and methodological contexts.

2.2 Analysis of TTE data in a single trial

2.2.1 Direct calculation of the Hazard Ratio

Chapter 1.1.4 introduces the theoretical concepts in the analysis of TTE data of the survival and hazard functions which are linked by a parameter known as the Hazard Ratio (HR). The HR can be directly calculated in two general approaches and several methods have been suggested for the indirect estimation of HRs; see Chapter 2.3.2 for further discussion. The first direct approach, via comparison of observed and expected numbers of events across treatment groups, defines the log hazard ratio ($\ln(HR)$) as [6, 118]:

$$\ln(HR) = \ln(\frac{O_T/E_T}{O_C/E_C}) \quad \text{(Equation 8)}$$

And the associated variance of this log hazard ratio ($var(\ln(HR))$) is defined as:

$$var(\ln(HR)\ ) = \frac{1}{E_T} + \frac{1}{E_C} \quad \text{(Equation 9)}$$

Where $O_T =$ Observed number of events in the treatment group.

$O_C =$ Observed number of events in the control group.

$E_T =$ Log-Rank Expected number of events in the treatment group (under null hypothesis).

$E_C =$ Log-Rank Expected number of events in the control group (under null hypothesis).

An alternative direct calculation of $\ln(HR)$, where O_T and E_T are defined as above and $1/V_T$ is the Mantel Haenszel (log-rank) variance of $\ln(HR)$, is [6, 118, 119]:

$$\ln(HR) = \frac{O_T - E_T}{V_T} \quad \text{(Equation 10)}$$

And the associated variance of this log hazard ratio ($var(\ln(HR))$) is defined as:

$$var(\ln(HR)\ ) = \frac{1}{V_T} \quad \text{(Equation 11)}$$

HRs derived from either of the above calculations are often referred to as 'log-rank' HRs due to the use of log-rank methodology in the calculation of expected number of events [10, 118]. The two $\ln(HR)$ estimates from Equation 8 and Equation 10 (and their respective variances) are approximately equivalent, and will only differ markedly if the total number of observed events in a trial is small [6, 120].

Secondly, the coefficient of a treatment indicator variable from a proportional hazards (PH) regression model provides a direct estimate of $\ln(HR)$. For example, given the PH model (Equation 7) defined in Chapter 1.1.4, hazards $h_T \; and \; h_C$ in the treatment and control groups respectively and coefficient θ associated with treatment indicator variable x [7]:

$$\ln(HR) = \theta = \ln(\frac{h_T(t)}{h_C(t)}) \qquad \text{(Equation 12)}$$

The variance $(var(\ln(HR)))$ is provided by the associated variance of the regression coefficient. If the treatment indicator is the only variable included in the regression model, the 'unadjusted' $\ln(HR)$ derived from the regression coefficient can be considered conceptually equivalent to the $\ln(HR)$ derived from log-rank methods as described above. However, often regression modelling is performed due to the desire to adjust for potential confounders of interest, the coefficient associated with the treatment indicator variable will be 'adjusted' for the additional variables. Therefore the difference between the 'adjusted' and 'unadjusted' HRs will depend on the influence of any confounding variables. While in a meta-analytic context, combination of adjusted and unadjusted estimates is likely to introduce bias into pooled effect estimates [121], methodology has been developed which allows the combination via meta-regression of Cox proportional hazards models with different covariate adjustment or no covariate adjustment [122].

HRs derived from regression models are often referred to as 'Cox' HRs, in reference to the semi-parametric Cox PH model commonly used for the analysis of TTE data[58]. It should be noted that HRs from parametric PH models (such as Weibull or Exponential PH models) also provide a direct estimate of $\ln(HR)$ and associated variance as described above.

It has been shown that the 'log-rank' and 'Cox' HRs are approximately equivalent for modest treatment effects and fairly balanced treatment arms; for larger treatment effects, the 'log-rank' approach becomes biased and confidence intervals show a lack of coverage [123-125].

2.2.2 Use and interpretation of the Hazard Ratio

The HR, or $\ln(HR)$, and associated variance are considered, generally under the PH assumption, to be the most appropriate summary statistics to present when reporting on randomised TTE data as the HR is the only summary statistic which takes account of both the proportion of individuals experiencing an event in question and the time at which the event occurs (or the censoring time for individuals not experiencing the event in question) [6, 118].

The HR is mathematically linked to the two well-known summary statistics used to express the relative difference of a dichotomous outcome at a fixed point in time (e.g. the end of a clinical trial); the Odds Ratio (OR) or Relative Risk (RR). In fact, in the absence of censoring, the HR and RR can be considered to be mathematically equivalent [118]. In the presence of censoring, it has been demonstrated that approaches such as comparing the number of events across patient subgroups or calculating the odds of an event at fixed time points across follow-up are inefficient and could lead to inappropriate conclusions [6, 23, 126, 127], especially among trials with varying lengths of follow-up and therefore variable stages of maturity [128, 129]. In a meta-analytic context, such methods could introduce bias into meta-analysis if fixed time points are selectively chosen by trialists to demonstrate a maximal or minimal difference between treatment groups [128]. Appropriate adjustment for censoring in meta-analysis of TTE data is essential; censored patients provide less information than those who experience an event to the overall distribution of event times, therefore the extent of censoring will affect the variance of survival proportions, hence the relative weights of individual results in meta-analysis and the precision of combined estimates [130].

The HR is often interpreted in a similar way to the RR due to the similarities between the notions of hazard and risk. It must be noted for the interpretation of this measure that 'hazard' is a conditional and dynamic measure which may change continuously (e.g. an individual's hazard of death increases as they cross the road), while 'risk' is assumed to remain fairly constant across the follow-up period of a study [121]. It has been argued that the HR doesn't have an intuitive interpretation for non-statisticians and is easily misinterpreted [131, 132] and that ORs or RRs are more commonly presented in clinical trial publications due to perceived statistical complexity related to HRs [128, 133].

2.3 Meta-analysis of TTE data

2.3.1 Literature review methods

A literature review was conducted to identify methodological publications (whether full-text articles, abstracts or conference proceedings published in English) describing meta-analytic methods developed for or applied to TTE data. MEDLINE electronic database (from the earliest date up to January 2017) and Cochrane Methods Methodology Register (from the earliest date up to the last update of the register in July 2012) were searched using keywords such as 'meta-analysis,' 'time-to-event,' 'survival,' 'failure time' and 'statistical methods.' See Appendix 1 for search strategies of the two electronic databases.

Electronic searches were constructed and carried out by SJN, with advice from an information specialist (GC). References of relevant publications were also consulted and methodological journals (Statistics in Medicine, Statistical Methods in Medical Research, BMC Research Methodology and Research Synthesis Methods) were hand-searched using the keywords as described above via the 'Search' or 'Advanced Search' function of the journal website.

The inclusion criteria of this methodological review were broad; publications describing a novel methodology for meta-analysis of TTE data or publications describing the application of existing methods (with methodological discussion) were considered. Publications describing methodology for meta-analysis in general, for TTE in general or publications describing an application only without methodological discussion were not considered.

Over 100 methodological publications were identified by the search methods described above. A list of references of all publications identified from the searches is available in Appendix 1. This methodological review was not intended to be systematic and the results section of this Chapter focus on the methodology related to the topic of this thesis; particularly on methods developed for AD and for IPD of TTE outcomes. This review was also not intended to 'compare' methods; as outlined in Chapter 2.1, the applicability of most of the methods is dependent on many factors. Rather, the relative advantages and disadvantages of specific methods within given contexts are discussed where appropriate.

2.3.2 Methods for aggregate data

Whitehead and Whitehead [7] were among the earliest authors to develop a general comprehensive methodological framework for the meta-analysis of RCTs including those with censored TTE data. The methodology involves combining efficient score statistics for the HR estimates based on the efficient score statistics and Fisher Information of an assumed PH model.

Assuming that for K trials, an estimate of $\ln(HR_i)$ and $var(\ln(HR_i))$ for each trial $i = 1 \dots K$ is available; a pooled $\ln(HR)$ and its variance ($var((\ln(HR)))$ can be calculated by the inverse-variance approach to meta-analysis [6, 7];

$$\ln(HR) = \frac{\sum_{i=1}^{K} \frac{\ln(HR_i)}{var(\ln(HR_i))}}{\sum_{i=1}^{K} \frac{1}{var(\ln(HR_i))}} \quad \text{(Equation 13)}$$

$$var(\ln(HR)) = \left[\sum_{i=1}^{K} \frac{1}{var(\ln(HR_i))}\right]^{-1} \quad \text{(Equation 14)}$$

Meta-analysis of relative effects (i.e. ratios) via an inverse-variance approach is performed on the natural logarithm of the effect size and variance due to the beneficial mathematical properties for meta-analysis of converting ratios to differences when a log scale is applied. A further benefit for TTE data and the HR is that the additive difference in log hazards of two treatments does not depend on the observation times and is therefore more appropriate when combining studies with different lengths of follow-up in meta-analysis [134-136].

Even at this early stage of methodological development, Whitehead and Whitehead [7] express doubt over the chances of finding sufficient details, of adequate quality, in study publications to allow meta-analysis. The authors also warn of potential variability in conventions and terminology used across the published literature, which may make identification of the appropriate statistics difficult.

As introduced in Chapter 1.1.3, the use of IPD is widely regarded as the 'gold-standard' approach to synthesis; however, in the event that IPD is not available for a proportion of eligible studies and the required effect size is not reported in the published literature, these studies may have to be excluded from analysis.

A potential solution as an alternative to excluding studies from meta-analysis entirely is to make use of the summary statistics available in the published literature. This can be done in two ways. Firstly by pooling an alternative statistic to the HR, where HRs are not available (a summary of some methods taking this approach is outlined in Chapter 2.3.4.2). Secondly, and arguably more appropriately [6, 13, 128], a range of methods have been developed which allow the indirect estimation of HRs and their associated variances from commonly published summary statistics such as p-values, numbers of events and from published survival curves [6, 13, 128, 137-139].

2.3.2.1 Indirect estimation methods: numerical

Parmar *et al* [6], Williamson *et al* [13] and Tierney *et al* [128] provide full derivation and illustrative examples of commonly used indirect estimation methods; the latter online publication also providing a macro-enabled Excel Spreadsheet to facilitate the indirect calculation of HRs and associated variances. In summary, the indirect methods outlined within these papers are as follows:

1) Indirect estimation of variance $var(ln(HR))$

The following methods assume that HR or $\ln(HR)$ is reported and $var(ln(HR))$ must be estimated. Note that for indirect methods b to d, $var(ln(HR))$ is the inverse of V_T as outlined in Equation 11.

a. Confidence intervals are reported

$$var(\ln(HR)) = \left(\frac{UCI-LCI}{2\Phi^{-1}(1-\frac{\alpha}{2})}\right)^2 \quad \text{(Equation 15)}$$

where LCI and UCI are the lower and upper bounds of the confidence interval (CI) respectively, α is the significance level (usually 5% corresponding to a 95% CI) and Φ is the cumulative distribution function of the Normal distribution ($\Phi^{-1}\left(1-\frac{\alpha}{2}\right) = 1.96$ if $\alpha = 5\%$).

b. Total number of events in both groups reported (randomisation ratio 1:1)

$$V_T \approx (O_N)/4 \quad \text{(Equation 16)}$$

where O_N is the total number of events in both treatment groups.

c. Number of events in each group reported (randomisation ratio 1:1)

$$V_T \approx (O_T\ O_C)/O_N \quad \text{(Equation 17)}$$

where O_T, O_C and O_N are the observed number of events in the treatment group, control group and total number of events respectively.

d. Total number of events in both treatment groups and number randomised to each group reported (randomisation ratio not 1:1)

$$V_T \approx \frac{O_N\ R_T\ R_C}{(R_T+R_C)^2} \quad \text{(Equation 18)}$$

Where O_N is the total number of events in both treatment groups, R_T and R_C are the number of participants randomised to treatment and control groups respectively.

2) Indirect estimation of $ln\ (HR)$ and $var(ln\ (HR))$, see also Equation 10

e. Log-rank p-value and total number of events in both treatment groups reported (randomisation ratio 1:1)

$$O_T - E_T \approx {}^1/_2\sqrt{(O_N)} \times \Phi^{-1}\left(1-\frac{p}{2}\right) \quad \text{(Equation 19)}$$

f. Log-rank p-value and number of events on each treatment group reported (randomisation ratio 1:1)

$$O_T - E_T \approx \sqrt{(O_T\ O_C)/O_N} \times \Phi^{-1}\left(1-\frac{p}{2}\right) \quad \text{(Equation 20)}$$

g. Log-rank p-value, total number of events in both treatment groups and number randomised to each group reported (randomisation ratio not 1:1)

$$O_T - E_T \approx \sqrt{\frac{O_N R_T R_C}{R_T + R_C}} \times \Phi^{-1}\left(1 - \frac{p}{2}\right) \quad \text{(Equation 21)}$$

Where for indirect methods e to g:

$O_T =$ Observed number of events in the treatment group.

$O_C =$ Observed number of events in the control group.

$O_N =$ Observed number of events in both treatment groups.

$E_T =$ Log-Rank Expected number of events in the treatment group (under null hypothesis).

$R_T =$ Number randomised in the treatment group.

$R_C =$ Number randomised in the control group.

$\Phi =$ Cumulative distribution function of the Normal distribution.

$p =$ Two-sided log-rank p-value (or Cox regression p-value as an alternative if log-rank p-value is not reported [128]).

Note that V_T can be estimated using any of indirect methods b to d, then $\ln(HR)$ can be estimated using Equation 10.

Indirect variance estimation methods b to g are derived from the χ^2 statistic of the log-rank test and have been shown to provide reasonable estimates of variance when treatment effect 'is not too large' [134, 140]. Illustrative examples have shown reasonably good agreement between these indirect estimates and also between these indirect estimates with direct estimates in a range of clinical settings [6, 141-144]. However, a simulation study has shown that indirect method 2 (Equation 16), may systematically overestimate the true log-rank variance and all three approximations improve as the proportion of censoring increases [142]. Therefore, in the context of oncology trials where treatment effects are likely to be modest and event rates low these indirect methods are likely to perform well, however for other contexts where diseases have a poor prognosis and therefore a high event and low censoring rate, the variance estimates from the methods are likely to be biased.

2.3.2.2 Indirect estimation methods: Survival curves

Survival curves (usually Kaplan-Meier (KM) curves) are commonly presented graphical representations of TTE data; previous reviews examining studies with TTE outcomes have shown that around 90% presented at least one survival curve (see Chapter 3.1.1 for further details and references). Given the popularity of such graphical figures, many methods have been proposed for extracting and combining data from published KM curves in the context of meta-analysis. Such methods generally aim to reconstruct or estimate trial-specific or pooled survival functions, rather than to estimate $\ln(HR)$ and associated variance specifically [6, 13, 126, 130, 145-155].

Earle *et al* [126] present an overview and comparison of five of these techniques [145-149]. Firstly, two similar methods which make use of the published survival curve to iteratively estimate the survival function via iterative Generalised Least Squares and iteratively reweighted least squares algorithms respectively [145, 147]. Also, two methods which derive 'Log RR' and 'Weighted Log RR' indices respectively as parameters to represent long-term survival from the survival curves of a selection of trials [126, 146]. These pooled indices can then be back transformed to pooled survival functions [149]. Finally, a method for combining data from survival curves while adjusting variability participant covariates across studies to produce an effect estimate which represents the variability in participant subgroups [148]. This method does rely on published results being stratified by the covariates of interest so may not be feasible in practice [126]. Earle *et al* [126] assess the accuracy of the five techniques by comparison of a reconstructed pooled curve to a curve constructed using IPD and show no more discrepancy from the IPD curve is present than would be expected by chance alone. The authors conclude that while not every technique would be appropriate for every context, in general, all five techniques could accurately reproduce summary survival curves from published literature and the choice of technique would depend on data characteristics the context and the aim of analysis.

Parmar *et al* [6] present a method for estimating $\ln(HR)$ and associated variance from published survival curves. A summary of this method is as follows:

- Split the time axis into T arbitrary non-overlapping intervals,
- Estimate $\ln(HR)$ for each interval based on estimated numbers at risk in the time interval (assuming uniform censoring across the entire follow-up period) and numbers of events in the time interval. Such information may be read directly from

the curve or using additional published information where provided (e.g. numbers of participants remaining in the trial at various time points).

- Combine estimates over the intervals in a stratified way to obtain an overall $\ln(HR)$ for the trial.

The full derivation of this method is provided in Parmar *et al* [6]. An illustrative example performed by the authors using their derived method shows that the assumption of uniform censoring is somewhat 'idealized' and underestimates the true variance of the estimate, resulting in the trial being given too much weight in meta-analysis. It has been shown that the assumptions made regarding censoring in a meta-analysis of TTE data can change pooled effect sizes by between 1-9% and often has an impact on statistical heterogeneity [156].

Williamson *et al* [13] propose an extension to the Parmar *et al* [6] methods following an actuarial approach that assumes the rate of censoring is constant within each interval but can vary across intervals, with the aim of improving variance estimates. These methods also extend to examining differences in treatment effect over time (i.e. non-proportional hazards) as a source of heterogeneity in meta-analyses of trials with variable lengths of follow-up. The methods assume that if the PH assumption holds, then $\ln(HR)$ estimates should be approximately constant across time intervals.

More recently developed methods have made use of digital software to extract data from KM curves at selected time intervals or, where possible, entire curves in an attempt to reconstruct IPD via iterative algorithms [157-159]. In theory, such methods allow a better determination of censoring distributions, avoiding the assumptions regarding censoring made in the methods of Parmar *et al* [6] and Williamson *et al* [13] and therefore producing more accurate results than any previous method [157]. Use of reconstructed IPD also allows the use of more advanced methods for further analysis; illustrative examples include modelling of fractional polynomials and flexible parametric modelling to account for non-proportional hazards in pairwise and network meta-analysis [158, 159]. Guyot *et al* [157] have demonstrated a high degree of reproducibility and accuracy for survival probabilities and median survival time, even in the absence of additional information regarding numbers at risk and numbers of events; however HR estimates became less accurate as less information was available to inform censoring patterns to becoming 'unusable' when no additional information was available regarding number at risk or number of events.

2.3.2.3 Reliability of indirect methods

Many studies have investigated the reliability and practicality of applying indirect estimation methods [6, 126, 141-144, 160-163], particularly the methods of Parmar *et al* [6].

It must be emphasised that the reliability of these estimations, particularly graphical estimations, are dependent on the quality of the published survival curves and the precision of other published information [6, 13]. It has also been shown that estimating the number of events from a published KM curve is often an overestimation of the true number of events and the extent of overestimation increases as an increasing number of patients are censored [164, 165]. Further, an illustrative example shows that the graphical methods of Parmar *et al* [6] , which were developed within the context of randomised clinical trials, may produce biased estimates when applied to survival curves of non-randomised data, particularly where non-randomised data has been adjusted in analysis to account for confounding or selection bias [162]. Further discussion of analytic methods for non-randomised and observational TTE data and the implications for meta-analysis are discussed by Bennett [166].

Parmar *et al* [6] also discuss the difficulty in selecting the best time intervals such that the event rate within each time interval is appropriate. Their suggested rule of thumb is that the event rate within each time interval should be no more than 20% of those at the beginning of the time interval. However, this is dependent on the event rate and follow-up time of an individual trial and a simulation study has shown that treatment effect may be under-estimated in trials with small sample sizes and/or low event rates [144].

A general recommendation from all work within this area is that as much summary information should be extracted from published literature as possible so that several indirect estimates can be calculated and compared [6, 128, 141]. Due to the assumptions involved (e.g. choice of time intervals, censoring rate), estimation from survival curves should only be used in the absence of sufficient information for the direct or indirect methods [6]. However, if multiple estimates can be calculated, it may be reasonable to take an average of these estimates to include within meta-analysis [6].

2.3.3 Methods for individual participant data (IPD)

Two approaches can be taken when conducting an IPD-MA, either a one-stage or two-stage approach [32, 34, 36, 167-169]. Two-stage methods reduce IPD to study-specific treatment effect estimates, allowing use of standard meta-analytic methodology to obtain a pooled treatment effect estimate. Such methods are commonly used in practice and are considered conceptually less complicated than one-stage methods [32, 36, 123].

One-stage methods simultaneously analyse IPD from all studies, while accounting for the separate trials within the one-stage model, to obtain estimates of pooled treatment effect and between-study heterogeneity by fitting a single hierarchical (fixed, random or mixed effects) statistical model. Such models, while more computationally complex, offer additional flexibility to incorporate covariates, interaction terms or heterogeneity parameters [32, 37, 169]. Generally within meta-analysis, heterogeneity between trials may arise from the treatment effects themselves; i.e. the intervention may have worked better in some trials than others, and treatment-by-trial interactions may arise due to differences in implementation of treatment protocols (i.e. treatment doses, dose scheduling etc.), study participant characteristics and their individual baseline risks [136], which may be interpreted as unmeasured characteristics [170]. One-stage approaches allow for simultaneous or separate modelling of heterogeneity of treatments effects and of baseline risk. Baseline risk may be measured on a trial level and modelled as a fixed-effect or the addition of a random-effect to a hierarchical model (also referred to as a frailty parameter within a TTE setting [171]) can account for the level of risk associated with an individual which could account for heterogeneity among individuals not explained by the covariates [37, 172].

Specifically within a TTE setting, advantages of the availability of individual TTE data for meta-analysis include a more thorough and flexible investigation of treatment effect over time [135], treatment-covariate interactions [13, 173] and patient characteristics and clinical factors as potential causes of statistical heterogeneity between trials in a meta-analysis [37, 174]. Specific sources of heterogeneity in TTE data include time-dependent (non-constant) treatment effects and variation in length of follow-up time across trials [13, 37]. Methodology has also been developed for IPD-MA of TTE data in the presence of competing risks [175].

2.3.3.1 One-stage IPD-MA

A simple assumption in IPD-MA of TTE data from RCTs would be that the hazard function differs only by the allocated treatment, with all other risk factors potentially influencing the hazard balanced by randomisation [135]. To achieve a better insight into potential treatment effect modifiers, this assumption can be relaxed via stratification of baseline hazard by trial [37], or by the addition of a treatment-study or treatment-time interaction [135, 176, 177].

Tudur-Smith *et al* [37] present a series of one-stage hierarchical Cox PH regression models designed to explore heterogeneity in meta-analysis and examine the relationship between participant-level covariates and TTE data via the addition of a treatment-study interaction indicator variable to each model for use in IPD meta-analysis.

Without stratification by study, it is assumed that the hazards within each study are proportional to a single fixed baseline hazard function, common to all studies within the meta-analysis. Stratification of the model allows for different baseline hazard functions for each study (assuming that hazards are proportional within each study), a less restrictive assumption when synthesising studies of variable settings and participant populations. The models can extend to either fixed or random treatment effects [172, 178-180] and to either a Frequentist or Bayesian hierarchical framework [181, 182]. Furthermore, Bennett *et al* [183] provide a comparison of three Cox PH based approaches to meta-analysis of TTE data, two from a frequentist framework and one from a Bayesian framework, with an additional consideration of the performance of such methods in the context of low event rates.

The general form of one-stage hierarchical Cox PH models are as follows; for the i^{th} participant, $i = 1 \dots n_j$, in the j^{th} study, $j = 1 \dots J$, the following models for the hazard function at time t (without participant-level covariates) are proposed [37] :

a. Fixed treatment effects with fixed (proportional) study effects

$$h_{ij}(t) = h_0(t)\exp(\beta_{0j} + \beta_1 x_{ij}) \quad \text{(Equation 22)}$$

$h_0(t)$ is the baseline hazard function in the reference study (e.g. study $j = 1$ so β_{01} is constrained to be zero), β_{0j} is the proportional effect on the baseline hazard due to the j^{th} study, $j = 2 \dots J$, x_{ij} is the treatment group indicator and β_1 is the $\ln(HR)$ for the treatment group compared to control group, assumed constant (fixed-effects) across all studies.

b. Fixed treatment effects with baseline hazard stratified by study

$$h_{ij}(t) = h_{0j}(t)\exp(\beta_1 x_{ij}) \quad \text{(Equation 23)}$$

h_{0j} is the baseline hazard function in the j^{th} study. Other parameters are defined as in Equation 22.

c. Random treatment effects with fixed (proportional) study effects

$$h_{ij}(t) = h_0(t)\exp(\beta_{0j} + \beta_{1j}x_{ij}) \quad \text{(Equation 24)}$$

$$\beta_{1j} = \beta_1 + b_{1j} \qquad b_{1j} \sim N(0, \tau^2)$$

β_1 is now interpreted as the mean $\ln(HR)$ for a distribution of treatment effects with a deviation from the population mean of b_{1j} in the j^{th} study, assuming deviations b_{1j} follow a Normal distribution with mean zero and variance τ^2 (where τ^2 represents between-study heterogeneity). Other parameters are defined as in Equation 22 and Equation 23.

d. Random treatment effects with baseline hazard stratified by study

$$h_{ij}(t) = h_{0j}(t)\exp(\beta_{0j} + \beta_{1j}x_{ij}) \quad \text{(Equation 25)}$$

$$\beta_{1j} = \beta_1 + b_{1j} \qquad b_{1j} \sim N(0, \tau^2)$$

All parameters are defined as in Equation 22, Equation 23 and Equation 24.

e. Random treatment effects and random study effects

$$h_{ij}(t) = h_0(t)\exp(b_{0j} + \beta_{1j}x_{ij}) \quad \text{(Equation 26)}$$

$$\beta_{1j} = \beta_1 + b_{1j} \qquad b_{1j} \sim N(0, \tau^2)$$
$$b_{0j} \sim N(0, \sigma^2) \qquad cov(b_{0j}, b_{1j}) = 0$$

b_{0j} is the deviation of the j^{th} study from the overall baseline risk, assuming deviations b_{0j} follow a Normal distribution with mean zero and variance σ^2. Other parameters are defined as previously.

Tudur Smith *et al* [174] compare these hierarchical Cox regression models with participant-level covariates to meta-regression models with aggregate-level covariates on an empirical dataset. Results show that the stratified models with random-effects estimate a larger standard error of treatment effect and adding participant-level covariates and interactions into the models helps to explain variation and decreases the levels of between-study heterogeneity than those with fixed-effects only. Previous work has also demonstrated the benefit of random trial effects in a TTE setting from the sharing of information ('borrowing strength') across all trials [171, 178, 184], particularly when combining a large number of trials with treatment groups of small sample size [171].

Tudur Smith *et al* [174] also show that evidence of treatment-covariate interaction is weaker in the aggregate meta-regression models than the IPD models. Also within these aggregate-level models, effect sizes seem to be dependent on the estimation approach, precision of estimation is poor and there is scope for the identification of false treatment-covariate associations due to multiple testing potential correlations between aggregate variables. The

authors conclude that meta-regression with AD can be accurate if there is evidence for a within-study treatment-by-covariate interaction and sufficient between-study variation for the aggregate value of the covariate, however such information is unlikely to be reported in practice and the stratified IPD approach is preferred. More recent work emphasises the importance of the appropriate specification of one-stage IPD models with treatment-covariate interactions by separating within-study and across-study interactions to avoid inadvertent ecological bias [25, 38, 41, 185], particularly within a TTE setting [25].

Michiels *et al* [186] also propose Cox PH regression models, which can be fitted in both Frequentist and Bayesian frameworks, to investigate heterogeneity from both variation in treatment effect and from difference in baseline hazard rates with via random (frailty) effects. The method applies random treatment-trial interaction terms and with adjustment for study variations such as region or population variability across studies. An assumption of a common baseline hazard function shape is made across the trials, allowing for varying magnitude of the hazard function due to systematic variation across trials. The authors apply their models to a large meta-analysis of 65 trials, originally analysed using a two-stage approach (with methods developed by Peto *et al* [187] for TTE data), with results showing a beneficial treatment effect with significant heterogeneity present. Application of the new models to the meta-analysis resulted in similar pooled HRs and the addition of between-trial variance increased the relative weight of small trials for the overall pooled result.

Katsahian *et al* [188] compare the performance of four one-stage Cox PH models; fixed-effects, random-effects (frailty), stratified and marginal. The fixed-effects, frailty and stratified models are all modelled on trial-specific hazard functions which can incorporate heterogeneity (conditional models). The marginal model assumes a multivariate structure with the advantage of allowing estimation of a population averaged treatment effect [189], a benefit for studies of varying sample size, but this model cannot incorporate heterogeneity. The authors perform a simulation study based on three separate assumptions; no heterogeneity present, heterogeneity in baseline risk present, heterogeneity in baseline risk and treatment effect present. The results of the simulation study show that standard errors are consistently underestimated by the marginal model, particularly for a small number of trials in a meta-analysis. Results also show that if heterogeneity in treatment effect is present, models without a random treatment-trial interaction perform poorly and the models with interactions tend to perform better with large numbers of trials and large sample sizes. The authors note that population averaged treatment effect may be of value where substantial heterogeneity is present in meta-analysis.

Rondeau *et al* [136] present a one-stage additive Cox model, similar to the model of Vaida and Xu [190] for clustered data, to jointly account for heterogeneity in meta-analysis of both treatment effects and baseline risk via a random treatment effect and a random interaction between trial and treatment effect. The authors also investigate the relationship between the two random-effects under the assumption that magnitude of effect is related to underlying risk. Rondeau *et al* [136] perform a simulation study, using a semi-parametric penalised likelihood approach to model fitting, and show that most accurate results are obtained for meta-analyses with large numbers of trials or large sample sizes and when a non-zero covariance of the two random-effects is specified (i.e. a value for the correlation between the two random-effects is assumed).

The complexity of one-stage random-effects models, particularly the semi-parametric form of the stratified Cox models proposed by Tudur-Smith *et al* [37] can lead to computational difficulties and problems with convergence [49, 172, 185, 191], making their use within practice difficult [123]. To allow the implementation of these random-effects models in standard statistical software, Simmonds *et al* [191] propose an approach which treats the random-effects as missing data. The authors apply the expectation-maximisation algorithm, approximating the expected values of the random-effects in the expectation step using shrinkage estimators. A simulation study and application to an example of post-operative radiotherapy for non-small-cell lung cancer show that this approach can provide estimates of random-effects without bias or loss of precision.

Crowther *et al* [49] propose an alternative flexible modelling approach for hierarchical models a to d above using Poisson generalised linear models via a piecewise exponential model. Such models extend to a Frequentist or Bayesian framework, can incorporate treatment-covariate interactions and non-proportional hazards. Massonnet *et al* [192] also remark that frailty models are often likelihood based and such a model structure can be reformulated into a linear mixed model assuming clustered data structure with a random cluster effect and random treatment effect. This transformation is applicable to meta-analysis under the assumption that clusters are analogous to studies. Such alternative methods are practical as linear mixed-effects models is more accessible in standard statistical packages than procedures for fitting conditional random-effects to TTE models [49, 192].

A simulation study shows that this Poisson approach proposed by Crowther *et al* [49] produces near identical estimates to the hierarchical Cox approach, including for all parameters when covariates are added to the model and the method is extremely

computationally efficient with coverage of Poisson models taking between 5-65 seconds compared to up to 29 hours for the stratified Cox models [37]. A further simulation study (on a cluster level) shows that model parameters (treatment effect and heterogeneity) are well estimated following transformation to linear mixed-effects for large numbers of clusters and large sizes of clusters. However this simulation was conducted assuming equal cluster sizes; an assumption which may not translate to meta-analysis (i.e. equal study sizes).

Crowther *et al* [193] present a series of multilevel mixed effects parametric models as an alternative, flexible approach to previous hierarchical semi-parametric Cox model approaches. The authors extend parametric frailty PH models and accelerated failure time models, in addition to the flexible parametric model of Royston and Parmar [194], to incorporate any number of normally distributed random-effects estimated via adaptive or nonadaptive Gauss–Hermite quadrature. The authors demonstrate the application of these models to IPD-MA via a simulation study and a re-analysis of a previous IPD-MA in breast cancer, showing similar results to their previous work [49].

See a recent review by Debray *et al* [36] for an additional summary of IPD-MA methodology for all data types; including TTE data.

2.3.3.2 Two-stage IPD-MA and comparison to one stage IPD-MA

Burke *et al* [169] provide a tutorial of key statistical methods for two-stage IPD-MA and one-stage IPD-MA and note that most differences between the approaches arise due to different modelling assumptions, rather than the choice of one-stage or two-stage itself. While two-stage methods for IPD-MA generally allow use of 'standard' methodology; i.e. 'standard' methods of analysing TTE data and 'standard' methodology for meta-analysing study-specific treatment effects, a wide variety of approaches within these two stages may still be used [195]. A review of IPD-MAs by Simmonds *et al* [32, 195] found use of methods developed by Peto *et al* [187], log-rank methods [196] and Cox PH regression methods [58]. The review authors also discuss further methodology which could be used in two-stage IPD-MA such as modelling via interval censored logistic models [32]. An additional application of a two-stage approach is the derivation and synthesis of risk prediction models as described by Pennells *et al* on behalf of the Emerging Risk Factors Collaboration [197].

Haines and Hill [53] question whether it is ever appropriate to combine summary results produced by a variety of methods given that differences in methodology are generally reflective of different aims of analysis. The authors provide an illustrative example of meta-

analysis of repeated measures survival data of accidental falls where a range of analysis approaches were taken; Cox PH regression models, Andersen – Gill recurrent events models [198], negative binomial regression, linear regression or by reducing IPD to dichotomous data. The authors also make an empirical comparison to AD-MAs of published summary statistics only. Results show that estimated SEs seemed to be dependent on the methodology used, therefore by combining estimates from these different methods, the relative weightings in meta-analysis would be influenced by the methodology used in the studies rather than the actual precision of the effect estimate, resulting in biased pooled effects.

Tudur Smith and Williamson [125] compare three methods of fixed-effects meta-analysis for TTE outcomes; stratified log-rank analysis (two-stage), inverse-variance weighted average of Cox model estimates (two-stage) and stratified Cox regression (one-stage). Theoretically, the three methods should produce similar estimates of the pooled $\ln(HR)$ and its variance when the underlying treatment effect is close to the null and the degree of heterogeneity is minimal. Also, the stratified log-rank analysis should in theory have the maximal statistical sensitivity for the detection of modest treatment effects. Both the simulation study and an illustrative example show that the methods are approximately equivalent for modest treatment effects and low levels of heterogeneity and for large treatment effects, the stratified log-rank analysis overestimates and the inverse-variance weighted average underestimates treatment effect. The stratified Cox regression model is the most consistent for varying levels of effect size and heterogeneity. The authors conclude that in practice the choice of the most appropriate method depends on study size, meta-analysis size, censoring distributions and deviation from PH assumptions.

Bowden *et al* [123] compare the performance of two-stage approaches combining 'log-rank' and 'Cox' HRs respectively (see Chapter 2.2.1 for further details) via DerSimonian and Laird random-effects method [12] to a one-stage random-effects Cox model [172] fitted using Restricted Maximum Likelihood [199]. Via a simulation study and illustrative example, Bowden *et al* [123] demonstrate a small amount of bias in the pooled 'log-rank' HR as magnitude of treatment effect increases compared to negligible bias in two-stage and one-stage Cox model estimates. However in absolute terms, the estimates of the two-stage and one-stage methods are very similar. The authors also demonstrate decreased model coverage and more conservative effect estimates with increasing sample size in all three methods due to increased HR variance under the random-effects model; an effect which would not be observed in fixed-effects analyses [125].

2.3.4 Other meta-analytic methods

2.3.4.1 Multivariate and network meta-analysis

Data reported at a series of time points or over multiple correlated outcomes has a multivariate structure [200, 201]. In general, the methods described above in Chapters 2.3.2 and 2.3.3 make the assumption of non-informative censoring; in other words, that an individual's time-to-event is independent of any mechanism leading to censoring. However, where two-or-more outcomes of interest are correlated, it is likely that the censoring distributions of the outcomes will also be correlated; violating the assumption of non-informative censoring [202]. For example, in the context that disease progression within a randomised trial is defined as a treatment failure; measured outcomes of 'time-to-treatment failure' and 'time to progression' would be correlated and the censoring of an individual for one of these outcomes would inform the other outcome.

Multivariate meta-analysis allows the joint synthesis of correlated endpoints from multiple trials, taking account of both between-study correlation (i.e. association between the within-study estimates of underlying effect sizes between studies due to differences in individual-level and study-level characteristics) and within-study correlation (i.e. association between the outcomes in question) [170, 203-206]. The benefits of such a joint approach to synthesis have been widely discussed [145, 170, 202-206]. A particular advantage in the context of TTE data is the use of multivariate meta-analysis to identify and validate surrogate markers such as progression-free survival as a surrogate marker of overall survival [202, 207-212].

Arends *et al* [170] present a general linear mixed model for the joint analysis of two-or-more TTE outcomes in random-effects meta-analysis with an illustrative comparison of their multivariate analysis to the results of separate univariate meta-analyses. The authors demonstrate that multivariate meta-analysis has advantages over univariate analysis such as investigating associations between event-free survival and length of follow-up (short or long-term) and investigating how both treatment difference and heterogeneity are influenced by population baseline risk.

The use of multivariate models in meta-analysis will generally require a substantial amount of information regarding the correlation structure of the outcomes which may not be provided in sufficient detail within published studies [145, 201, 213] so an IPD approach may need to be taken [208, 214]. However, methods have been developed [145, 155, 170, 215-217] which make use of published or indirectly estimated aggregate TTE data in multivariate meta-analysis (see Chapter 2.3.2 for further discussion of indirect estimation).

Dear [145] presents an iterative generalised modified least-squares algorithm for the joint fixed-effects meta-analysis of a single TTE outcome reported repeatedly over time in several trials, while allowing a different set of between-trial and within-trial covariates to be modelled for each outcome. However, in the absence of IPD, the algorithm requires standard errors of survival probabilities are published (or reconstructed indirectly) and that sufficient information regarding the correlation structure of the data is published.

Arends *et al* [155] present a model for multivariate aggregate TTE data which generalises the fixed-effects models of Dear [145] and a random-effects model proposed by Berkey *et al* [218] which reduces to random-effects meta-analysis model of DerSimonian and Laird [12] for a fixed time point. The generalised model also allows extensions to include proportional or non-proportional hazards and time, trial and treatment interactions.

Fiocco *et al* [215-217] take a different, hazard based approach to the multivariate meta-analysis of published survival curves under the assumption of heterogeneity between studies. Using extracted or indirectly estimated information on number of events, proportion censored and effective numbers at risk within given time intervals, the authors construct piecewise hazard functions by treatment arm, constant within time intervals, via a Poisson correlated gamma frailty model assuming negative binomial marginal distributions and that the correlation structure is described by a multivariate gamma process derived by the authors [217]. The methodology can be used to estimate mean survival, correlation over time, within-studies, within and between treatment arms and the degree of heterogeneity between trials. The methods can be extended to incorporate study-level covariates to explain heterogeneity via meta-regression and can accommodate PH or non PH.

Jackson *et al* [219] propose an alternative multivariate approach to AD-MA of TTE outcomes which models the probability of the event at multiple time points using exact binomial within-study distributions, thus avoiding assumptions regarding hazard functions. The approach also extends to modelling covariates and accounting for censoring and the authors provide an application to an AD meta-analysis of critical leg ischemia data.

Multivariate meta-analysis methods extend also to NMA [220]; see Chapter 1.1.2 for an introduction to NMA and Efthimiou *et al* [18] for a review of methodology for NMA. Within a TTE setting, tutorials have been provided on performing NMA on a $\ln(HR)$ scale or using other summary statistics such as mean or median time-to-event [221, 222]. Additionally, methodology has been developed to investigate the impact of non PH, parametric modelling

of the survival function, and extension to study-level covariates on the consistency of NMA results [20, 158, 159, 223]. Methodology has also been developed for the simultaneous synthesis of IPD and AD in a Bayesian framework [45]; this methodology also extends to the synthesis of IPD (censored TTE data) with summary-level count data (event count within a given follow-up time) where published summary TTE statistics are not available [44].

2.3.4.2 Alternatives to the hazard ratio

As introduced in Chapter 2.2.2, the HR is recommended as the most appropriate relative summary statistic of TTE data under the PH assumption and the use of RRs or ORs as an alternative or approximation to the HR can result in inappropriate conclusions. However, when the PH assumption is violated, the HR is dependent on the length of participant follow-up and may not necessarily have an intuitive interpretation [132, 224, 225]. Furthermore, it has been demonstrated that even if PH violations are not a concern within individual trials, the PH assumption may still be violated in meta-analysis across multiple trials, which has implications for the interpretation of a pooled HR [224]. Various alternatives to the HR have been proposed which may be more appropriate under non PH such as the ratio or difference of medians [132, 226, 227], percentiles of survival [228, 229], survival rates [229], and most recently difference in restricted mean survival time (RMST) [224, 225, 229-231].

In a meta-analytic context, Simes *et al* [226] proposed the pooled log 'median ratio' ($log\ (MR)$) by taking an average of median survival times weighted by the sample size in each treatment arm. However, via an empirical comparison using IPD, Michiels *et al* [160] demonstrate that this $log\ (MR)$ can over or under-estimate treatment effect, with discrepancies mainly occurring at trial level with even more pronounced biases for low event rates. The authors argue that use of published median values only to calculate $log\ (MR)$ is likely to reduce in further biases and lack of statistical power where participant-level information regarding attrition is not available and the authors do not recommend the use of $log\ (MR)$ as a surrogate for $ln(HR)$.

Moodie *et al* [232] present a non-parametric procedure for the evaluation of treatment effect in the meta-analysis of TTE which uses the log (-log) survival function difference, as an alternative measure when published HRs are not available. The resulting pooled effect estimate is interpreted as the 'weighted average on the natural log scale of hazard ratios within interval $[0, t]$ in a trial.' The authors also discuss various weighting schemes for meta-analysis of log (-log) survival function difference to account for precision of study-specific estimates and for study quality.

Siannis *et al* [228] argue that the imposition of the PH assumption across multiple trials in the context of meta-analysis is particularly restrictive and a more flexible parametric representation of treatment effects allowing for shape parameters to vary between trials may be preferable. The authors propose use of the pooled 'percentile ratio' (a continuous function of survival percentile) which reduces to the 'median ratio' for the 50^{th} survival percentile, estimated via a parametric accelerated failure time model as an alternative to the semi-parametric PH model. Such a model would be fitted in a one stage-approach (see Chapter 2.3.3.1) to IPD-MA and can be fitted with fixed or random-effects in a Frequentist or Bayesian framework. Barrett *et al* [233] extend this method to the two-stage estimation of pooled percentile ratios using only KM estimates of the survival function, removing the need to make any distributional assumptions.

Most recently, the difference in RMSTs, defined as the difference in areas under two survival curves to time t, has been proposed as an alternative to HR [224, 225]. Such a measure has direct applications to cost-effectiveness analysis [230] and the attractive properties of not requiring an assumption of PH and is interpreted on the scale of the time-to-event which is arguably a more intuitive interpretation than that of relative hazards [224, 225]. Wei *et al* [224] present three parametric and non-parametric estimation methods of RMST and describe the calculation of the effect size difference in RMST and associated variance as an alternative to HR for two-stage IPD meta-analysis of TTE outcomes. The authors compare different estimation methods via a simulation study and conclude that the three methods of estimating RMST perform similarly well.

Lueza *et al* [225] extend the methods of Wei *et al* [224] to consider the use of difference in RMST in meta-analysis from addition estimation methods and across a wider range of simulated meta-analysis parameters such as variations in heterogeneity in baseline risk and treatment effect, fixed or random-effects, number of trials and number of participants. The authors conclude that pooling of trial-specific KM curves under DerSimonian-Laird random-effects [12] provides the best compromise across all scenarios of estimating difference in RMST for an IPD-MA. Both Wei *et al* [224] and Lueza *et al* [225] also demonstrate the use of RMST in meta-analysis via reanalysis of IPD-MA in cancer and both author groups conclude that that difference in RMST is a useful and interpretable effect measure for IPD-MA.

Combescure *et al* [234] note that interest may not always lie in a relative comparison of two interventions and the authors propose an approach for the meta-analysis of the published survival curves of single treatment arms to obtain a distribution free summary survival curve

assuming random-effects via product-limit estimation. The method allows for the estimation of pooled mean and median survival times, as well as estimates of heterogeneity between the published survival curves. The authors demonstrate the method via a simulation study and an application to an AD-MA of graft survival following kidney transplant.

2.4 Discussion

Meta-analyses of TTE data are commonly performed, particularly in the field of oncology. A range of methods have been developed to allow synthesis of TTE data depending on the type of data available (IPD or summary (AD) level only), use of fixed or random-effects, the desire to explore heterogeneity of treatment effect or baseline risk via the addition of participant-level or study-level covariates or treatment covariate interactions and whether the proportional hazards can be assumed. Many such meta-analytic methods also extend other synthesis approaches such as multivariate and network meta-analysis.

This chapter presents a methodological review of many important methods used in the synthesis of TTE data as well as a summary of the reliability and applicability of many of these methods in practice.

It is widely accepted that an IPD approach to meta-analysis is the 'gold-standard'[24]; particularly for TTE data [6]. This is reflected in many methods which have been developed in the last decade require an IPD approach, particularly the development of one-stage approaches allowing clinical questions of growing complexity to be addressed via the addition of random trial and participant effects or by participant-level covariates and treatment by covariate interactions [36]. The practical use of some of these methods has, however, been questioned due to the complexity of some modelling assumptions leading to problems with convergence and various alternative, more accessible, approaches have been suggested [49, 123].

It must not be forgotten that within many settings, a complete IPD approach to synthesis may not be feasible and a partial or complete AD approach to meta-analysis may be required [6]. It is also well documented that within a TTE setting, summary statistics required for meta-analysis are rarely published (see Chapter 3 for further discussion), therefore a range of methods have been developed which make use of more commonly reported summary statistics and published survival curves to indirectly estimate hazard ratios and associated variances. The methods proposed by Parmar *et al* [6], later translated into 'plain language'

and implemented to a macro enabled spreadsheet for use by non-expert users by Tierney *et al* [128], are perhaps the best known of these methods. While such methods are theoretically useful and accessible to statisticians and non-statisticians alike, it has been questioned whether such methods are useful in practice as often the summary statistics required to make use of such methods aren't reported either, or published graphical figures are of too poor quality to adequately make use of graphical estimation methods [6, 141, 144]. Furthermore, some applications including simulation studies have shown that some of the indirect estimation methods proposed become biased in certain scenarios and if used, would introduce bias into meta-analysis [6, 141, 142].

The information presented within this chapter is a review of methodological literature with relevance to this thesis, rather than a systematic review of all methods developed for meta-analysis of TTE. While this is a limitation, the search techniques employed to inform this methodological review were broad (see Chapter 2.3.1), therefore it is unlikely that any important methodology in relation to this thesis was missed.

In summary, meta-analytic techniques of TTE data have been proposed for and applied to a wide range of clinical and methodological scenarios. While the availability of IPD allows for more complex meta-analytic modelling of TTE data and a wide variety of indirect methods have been suggested where IPD and published AD are not available, it is important that as this research field continues to develop, the applicability and accessibility of new methodology is kept in mind.

Chapter 3 and Chapter 8 of this thesis make an assessment of published summary statistics for TTE outcomes epilepsy monotherapy trials and demonstrates the applicability of these indirect methods within this setting. Chapter 6 and Chapter 7 present the methodology and results of an IPD-NMA of antiepileptic drugs with four TTE outcomes.

Chapter 3: Aggregate time-to-event (TTE) data: a systematic review of reporting of outcomes and statistical analyses in epilepsy monotherapy trials

3.1 Introduction

An individual participant data meta-analysis (IPD-MA) is considered to be the 'gold-standard' approach to data synthesis for many reasons (see Chapter 1.1.4 for a more detailed discussion) [23, 24]. These reasons include the ability to comprehensively undertake time-to-event (TTE) analyses and the standardisation of outcomes and analyses across studies. Inconsistency of definitions, reporting and presentation of outcomes, effect sizes and statistical analyses for TTE outcomes have been documented since the early 1990s [31, 144, 160, 235-237]. Due to this inadequate reporting, an IPD analysis is often the only approach that can be taken for TTE data. The consistency and quality of reporting of TTE data is particularly important in an evidence synthesis context where aggregate data meta-analyses (AD-MA) can only use published information. As earlier outlined in Chapter 2.3.2, methods have been developed for the indirect estimation of TTE measures required for AD-MA from other published statistics [6, 13]. However, in practice, it is uncommon for the statistics required for indirect estimation to be reported either [141, 144, 160, 161].

This chapter summarises previous work on the reporting of aggregate TTE data and extends this methodology to a systematic review of the reporting of TTE outcomes and associated statistics in epilepsy monotherapy studies.

3.1.1 Summary of previous reviews of aggregate TTE data

Table 1 and Table 2 summarise the findings of previous reviews of aggregate TTE data.

The first review to investigate the reporting of TTE analyses was conducted in the 1990s [31]. The authors considered 132 studies with TTE endpoints published in five oncology journals.

In summary, 11 out of the 132 (8%) papers failed to state how many participants were analysed, almost half of the papers (48%) did not give any summary of length of follow-up and in 62% of papers at least one end point was not clearly defined. Results were often reported with p-values only; 63 out of 84 papers performed log-rank analyses (75%) and 22 out of 47 (47%) performed 'multivariate' analyses reported p-values. Less than a sixth of

papers using log-rank analyses (10 out of 84, 12%) and only half of those using 'multivariate' analyses (25 out of 47, 53%) reported a survival estimate or effect size such as hazard ratio (HR) or odds ratio (OR) and even fewer reported an associated measure of precision such as a standard error or confidence interval.

Survival plots were presented in 95% of the papers; however, the quality of survival plots was deemed poor in 42 out of 117 (37%) papers with the most common issues being censored observations not marked, poor or unhelpful numerical axis, survival curves of two groups for comparison not clearly distinguished, inadequate or no legend reported and inconsistency between curves and other reported results.

Overall, the presentation of analyses and graphs was deemed adequate in only 21% of papers (28 out of 132) of [31]. The authors report an appendix of guidelines for the reporting of survival analyses and a later series of tutorial papers for the conduct of such analyses [31, 238-240].

Table 1: Characteristics of previous reviews on the reporting of aggregate TTE data and current systematic review of epilepsy monotherapy trials

Review number	Reference	Number of studies (outcomes)	Journals	Study publication dates	Type of study
1	Altman *et al* 1995	132 studies	5 oncology journals	October to December 1991	Any study design reporting a survival analysis
2	Michiels *et al* 2005	131 comparisons	Not stated	Not stated	Comparisons of chemotherapy drugs in metastatic lung cancer
3	Hirooka *et al* 2009	129 studies	2 oncology journals	January 2004 to December 2005	Phase III RCTs
4	Mathoulin-Pelissier *et al* (2008)	125 studies (267 outcomes)	4 general medical journals, 4 oncology journals	2004	RCTs
5	Abraira *et al* 2013	104 studies published in 1991; 240 studies published in 2007	13 high impact journals (cardiology, internal medicine, nephrology and oncology)	1991 and 2007	Any study design reporting a survival analysis
6	Baston *et al* 2016	32 studies	5 oncology journals	April to July 2015	Phase III RCTs
7	Current review	54 studies (98 outcomes)	No restriction	1978 - 2012	Epilepsy monotherapy RCTs

Table 2: Summary of the reporting of TTE outcomes and statistical analysis; previous work and current systematic review of epilepsy monotherapy studies

Information or summary statistic	Review Number[1]									
	1	2	3	4	4	5	5	6	7	7
	Study level	Study level	Study level	Outcome level	Study level	Study level (1991)	Study level (2007)	Study level	Outcome level	Study level
Median follow-up time reported	39%	NR	60%	NA	57%	26%	41%	78%	NA	13%
Minimum and maximum follow-up time reported	NR	NR	25%	NA	NR	6%	2%	NR	NA	19%
Any summary of follow-up time reported	55%	NR	NR	NA	57%	69%	76%	91%	NA	76%
Method for calculating the median follow-up reported	12%	NR	NR	NA	2%	NR	NR	NR	NA	0%
Sample size calculation reported	93%	NR	NR	NA	78%	14%	28%	NR	NA	50%
Outcome(s) clearly defined	38%	NR	NR	42%	52%	NR	NR	66%	49%	55%
Time origin clearly defined	52%	NR	NR	76%	78%	NR	NR	NR	60%	65%
Event of interest clearly defined	NR	NR	NR	70%	79%	NR	NR	NR	87%	78%
Censoring clearly defined	NR	NR	NR	47%	58%	30%	39%	NR	49%	55%
Analysis methods for losses to follow-up reported	26%	NR	NR	NR	NR	NR	NR	NR	48%	56%
Number of participants analysed	92%	NR	NR	NR	NR	100%	99%	NR	95%	93%
Number of events reported	45%	NR	74%	65%	72%	71%	75%	72%	73%	63%
Observed and expected number of events	NR	NR	1%	NR	NR	NR	NR	NR	0%	0%
Multivariable analysis used	36%	NR	NR	NR	NR	NR	NR	59%	45%	65%
Cox regression used	41%	NR	NR	NR	51%	45%	75%	87%	43%	63%

Hazard Ratio presented	NR	3%	52%	NR	52%	NR	NR	87%	32%	31%
Hazard Ratio and CI presented	NR	NR	51%	NR	NR	NR	NR	87%	32%	31%
Any effect size presented	NR	NR	NR	93%	95%	NR	NR	NR	45%	44%
Any effect size and CI presented	NR	NR	NR	61%	67%	53%	94%	NR	45%	44%
Adjusted effect size presented (from a multivariable analysis)	12%	NR	NR	NR	NR	NR	NR	NR	21%	31%
Adjusted effect size and CI presented (from a multivariable analysis)	19%	NR	NR	NR	NR	NR	NR	NR	21%	31%
Log-rank p-value presented	57%	37%	97%	NR	NR	NR	NR	NR	40%	48%
Cox model p-value presented	27%	NR	NR	NR	NR	NR	NR	NR	14%	15%
Median survival presented	2%	73%	59%	NR	46%	NR	NR	NR	19%	26%
Subgroup analyses presented	52%	NR	NR	NR	NR	NR	NR	56%	30%	39%
Survival probability presented	2%	43%	71%	NR	NR	NR	NR	NR	61%	72%
Survival curve presented	95%	NR	94%	NR	92%	95%	86%	100%	76%	93%
Number at risk presented	8%	NR	38%	45%	53%	NR	NR	66%	22%	17%

Abbreviations: CI – Confidence Interval; NA – not applicable (items are applicable only to study-level reporting); NR – Not reported

1. See Table 1 for listings of review numbers and characteristics of reviews

Pocock *et al* [241] have also made recommendations regarding the presentations of survival plots in published literature. The authors emphasise the difference in graphical interpretation depending on the direction of the survival curve and event rate in terms of absolute and relative treatment difference. Pocock *et al* [241] also argue that an effect size and related measure of precision should be presented alongside a survival curve to aid with visual interpretation.

A literature review of 131 comparisons of chemotherapy drugs in metastatic lung cancer conducted by Michels *et al* [160] shows the median survival time to be the most commonly reported summary statistic (73% of studies). In their review, only 3% of studies reported a HR, 37% reported a p-value and 43% reported one year survival in the text of the published report or clearly on a survival plot.

Mathoulin-Pelissier *et al* [236] and Arkenau *et al* [237] conducted systematic reviews of 125 and 144 oncology RCTs respectively and noted inadequacies in relation to the definitions and reporting of important recommended TTE outcomes such as overall survival and progression free survival. Arkenau *et al* [237] note that binary response rates are reported more frequently than TTE outcomes (i.e. proportion surviving rather than survival time) and note on a lack of standardisation of definition of TTE intervals. For example, from the colorectal cancer studies included in their review, time origin for survival time was defined as date of disease diagnosis, date of enrolment into trial or date of first visit to oncologist. Arkenau *et al* [237] note that a lack of a uniform definition could potentially impact on the ability to combine trials with differently defined endpoints in meta-analysis.

Mathoulin-Pelissier *et al* [236] also assessed several areas related to reporting of TTE outcomes and statistics. Out of 125 oncology studies; time origin was defined in 98 (78%), the event of interest was defined in 99 (79%) and the number of events were reported in 90 (72%), censoring events were defined in 73 (58%) and the number of patients at risk were reported in 66 (53%). Further, median follow-up was reported in 71 (57%) and an estimation of survival or effect size was reported in 119 (95%); HR in 65 (52%) of studies and difference in median survival in 57 (46%) and survival curves were presented in 115 (92%). Mathoulin-Pelissier *et al* [236] conclude that standardised definitions of survival endpoints, events, censored events and follow-up should be developed to improve the reporting and interpretation of results in cancer clinical trials.

Hirooka *et al* [144] performed a similar review to that of Altman *et al* [31], of phase III RCTs published between January 2004 and December 2005 from two oncology journals with the

objective of determining whether the indirect estimation methods of Parmar *et al* [6] could be used in practice (see Chapter 2.3.2 for further details). From 129 included articles, only 66 (51%) reported a HR and confidence interval that could be directly used in an AD-MA. Only 2 studies (1%) reported observed and expected numbers of events required for use of the direct Peto method [10]. While 125 studies (97%) reported log-rank p-values and 96 studies (74%) report the number of events, only 35 (27%) reported both statistics required for indirect estimation [6]. Also while 121 (94%) of the studies presented KM plots only 49 (38%) provided the numbers at risk required for the estimation method of Williamson *et al* [13] and only 32 (25%) provided minimum and maximum follow-up times required for the Parmar *et al* [6] indirect estimation method. Hirooka *et al* [144] note that median survival time is the most common effect size reported in 76 studies (59%).

More recent work conducted by Abraira *et al* [242] comparing survival analyses published in 1991 to those published in 2007 in 13 high impact internal medicine, cardiology, nephrology and oncology journals has shown a large increase in the number of published analyses but little improvement in the quality of reporting in these analyses between 1991 and 2007. Abraira *et al* [242] emphasise the lack of reporting of numbers of events and follow-up time in around 30% of studies and highlight the lack of articles reporting any checks or validations of important statistical assumptions, such as the PH assumption [58]. Batson *et al* [243] also highlight the lack of details reported in oncology trials published in 2015 regarding validation of the statistical models in terms of the PH assumption and the potential impact on the interpretation of such unvalidated data for meta-analysis, NMA and health technology assessment. Both author groups call for recommendations to improve the reporting of survival analyses in journal articles and Abraira *et al* [242] propose an appendix of minimum requirements for the reporting of survival analyses.

While some reviews of this topic [31, 160] were conducted and published before the introduction of the CONSORT statement for improving the quality of reporting of randomised controlled trials first published in 1996 [244], revised in 2001 [245] and updated in 2010 [246]; more recent work published after 2008 [144, 236, 237, 242, 243] demonstrates levels of reporting similar to that reported over 10 years earlier.

Variability of outcomes and reporting is not restricted to single studies. Floriani *et al* [235] reviewed 15 AD-MAs of comparative cancer studies published after 1985 and found a large variation in the summary statistics reported and methods used for outcome evaluation. Four of the meta-analyses evaluated the total number of deaths, five compared survival at a single fixed time point, three compared survival at multiple time points, two compared median

survival time and three compared the hazard rate. In terms of pooled summary statistics reported, nine used an OR, three used HR, two used RR, one used risk difference, and two used difference in log median survival. Only four of the 15 reviews accounted for censoring in analysis. Floriani *et al* [235] raise their concerns over this heterogeneity of reporting which may compromise the reliability and interpretation of results.

Guyot *et al* [157, 247] argue that the poor reporting of studies and reviews with TTE outcomes may be due to lack of clarity in guidelines of what should be reported. The Cochrane Handbook of Systematic reviews [248] advises that effect size should be expressed as a HR and the CONSORT statement recommends that "the measure could be the HR or difference in median survival time." [245]

Neither recommends the reporting of a measure of precision (standard error, variance, confidence interval) which is required for meta-analysis even though the CONSORT statement recommends the reporting of a measure of precision for continuous and dichotomous outcomes. Furthermore, the HR or difference in median survival gives no indication of average prognosis in each trial arm which is generally required for cost-effectiveness analysis. Guyot *et al* [157, 247] also recommend the reporting of numbers of events and effective numbers at risk (potentially in the form of a life table) to facilitate evidence synthesis and economic evaluations of TTE event data.

3.1.2 TTE outcomes in published epilepsy monotherapy trials

Previous reviews of the reporting of TTE analyses have mainly considered trials in the field of oncology where the focus of the trials is usually 'survival' (i.e. time-to-death) [31, 144, 160, 235-237]. As outlined in Chapter 1.1.4, there are other clinical areas where important outcomes require the analyses of TTE data; one of these conditions is epilepsy.

The majority of people with epilepsy can achieve remission of seizures following treatment with a single antiepileptic drug (AED monotherapy) [71, 249] and recruitment populations for such a design include newly diagnosed patients, patients previously intolerant to or inadequately treated with a single drug, and patients previously successfully treated who relapse on discontinuation of their previous regimen [249].

A monotherapy trial has one of two designs; a 'complete' monotherapy in which all participants are AED naïve and receive only the allocated drug throughout the whole trial period or a withdrawal / conversion to monotherapy design in which any pre-study AEDs are discontinued before randomisation and then all participants receive only the allocated drug

throughout the whole trial period. The conversion to monotherapy design is thought to be useful as a proof-of-principle investigation (i.e. proof of efficacy) but of limited use in routine clinical practice with a complete monotherapy design being more pragmatic [250, 251].

Epilepsy monotherapy trials are usually designed to measure efficacy, tolerability and overall effectiveness of AEDs [249]. The Commission on AEDs of the International League Against Epilepsy (ILAE) [249, 252, 253] defines 'efficacy' as the ability of a medication to produce seizure-freedom and 'tolerability' is related to the 'incidence, severity and impact' of AED-related adverse events, most importantly those which lead to the discontinuation of a drug. It is recommended that 'retention time,' defined as time-to-withdrawal of allocated treatment after randomisation due to inadequate efficacy and/or poor tolerability [249] over a "long-term" treatment period of at least 48 weeks [252, 253], should be used as the primary outcome for monotherapy trials as this is a combined 'effectiveness' outcome reflecting both efficacy and tolerability. This is the outcome of greatest clinical utility and relevance [254] and retention is an outcome to which the individual makes a contribution. This is the outcome adopted as the primary outcome for systematic reviews of monotherapy studies performed by the Cochrane Epilepsy Group [59-67, 69]. It is also recommended that the primary 'effectiveness' outcome should also be supported by secondary outcomes of efficacy and/or tolerability [249, 251, 254]. The secondary efficacy outcomes adopted in Cochrane Epilepsy reviews of monotherapy studies are time-to-12-month remission, time-to-6-month remission and time-to-first seizure, in addition to a summary of the tolerability of the treatments in terms of reported adverse events [59-67, 69].

Guidelines proposed by the European Medicines Agency Committee for Proprietary Medicinal Products (CPMP) in 2000 [255] and updated in 2010 [251] make generally similar recommendations to those of the ILAE guidelines [249, 252, 253], but some differences relating to recommended outcomes in monotherapy studies. These guidelines recommend that for epilepsy monotherapy studies [251, 255]:

> *"…in newly diagnosed patients the primary efficacy variable should be based on the proportion of patients remaining seizure-free for at least six months (excluding the dose escalation period). However the trial should have a minimum duration of one year in order to assess safety and maintenance of efficacy…"*

and that possible secondary efficacy variables may concern;

> *"…a treatment retention time, measuring the combination of failed efficacy and tolerability, enables to assess the global clinical effectiveness of the drug."*

Despite guidelines from the ILAE [249, 252], International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use [254] and European Medicines Agency [251, 255] that a primary outcome of effectiveness or efficacy should be clearly defined in monotherapy trials, little uniformity in the definition of outcomes has been shown. For example, while the majority of trials will measure and report an efficacy outcome relating to seizures, some trials will record time-to-first seizure following randomisation [256] while others will record seizure-freedom at a specific time point (e.g. 12 months after randomisation) [257] or change in seizure frequency over a period of time [258].

There is also a lack of uniformity of statistical analyses performed and reported in monotherapy trials. While the majority of monotherapy trials aim to compare one AED compared to another in terms of efficacy and tolerability, statistical analyses performed in trials range from reporting counts of seizure and adverse event frequency only [259], to multivariable Cox regression modelling to adjust for relevant prognostic factors [260], to complex and innovative trial designs to demonstrate non-inferiority of a newer AED compared with a "standard" treatment [261, 262].

An ILAE investigation [252] of epilepsy monotherapy RCTs and systematic reviews conducted up to 2006 and updated in 2013 [253] concluded a paucity in high quality evidence due to *"alarming lack of well-designed, properly conducted epilepsy RCTs,"* and therefore insufficient information to "answer important clinical questions" and better inform treatment policy. However, this investigation places a high level of importance on the blinding of studies which some investigators believe may not be possible or ethical in epilepsy monotherapy and not representative of pragmatic clinical practice [256, 263-265] despite previous ILAE guidelines recommending that comparative monotherapy studies *"should aim to replicate usual clinical practice as closely as possible* [249]." The investigation is also critical of studies conducted for *"regulatory and marketing-driven"* purpose with protocol driven endpoints and populations that may *"bias the results in favour of the sponsors' product*" and would not *"reflect routine clinical care, meaning that results may not be fully generalizable to routine practice* [252]."

Less emphasis is given in the investigation to the accuracy and quality of statistical analysis and reporting in studies; these aspects are essential for the synthesis of trial results to inform clinical practice and future research. The investigation concludes that demonstrating differences statistically between drugs in terms of efficacy and tolerability *"has been very hard to show, except in a few studies* [252],*"* and furthermore it has not been possible to demonstrate many convincing differences between AEDs in systematic reviews and meta-

analysis conducted by the Cochrane Epilepsy Group [59-67]. However, it is difficult to know whether this apparent lack of convincing differences between AEDs is due to a true absence of differences between treatments or whether individual results may been biased due to inappropriate study designs and analysis of *"methodologically flawed"* studies.

The choice of study design in epilepsy has also been widely debated in the literature. An argument against the superiority design [252] if a study fails to show significant differences in seizure control between treatments is that seizure remission rates could be related to the natural history of the disease rather than efficacy of administered drugs, particularly in newly diagnosed individuals [71, 266]. Therefore in order to obtain a licence for monotherapy, it is now necessary to demonstrate non-inferiority of an experimental drug to an established comparator at its optimum dose [251, 267]. However, a finding of equivalence or non-inferiority does not exclude that in the particular population and under the specific conditions in which the trial was undertaken, both treatments could have been similarly ineffective [250], particularly if trials of a statistically complex non-inferiority or equivalence design have not been adequately powered or results and conclusions interpreted appropriately under the assumptions of equivalence or non-inferiority [252, 253].

The variation in the designs, definitions and type of outcome reported as well as variable approach to statistical analysis makes undertaking a meta-analysis of epilepsy monotherapy trials difficult without performing an IPD review. While such an approach is considered 'gold-standard' for TTE outcomes [23, 24], obtaining and re-analysing IPD is time consuming and resource intensive. Often, potentially relevant information has to be excluded from meta-analysis if IPD is not available and summary information for an outcome has not been reported adequately or not reported at all [59-67, 69]. Some updates of Cochrane reviews have taken many years to complete or updates are still ongoing [67], due to the time required to obtain, prepare and analyse IPD (see Chapter 5 for further discussion).

The choice of initial AED for an individual should be based on the highest quality evidence from randomised controlled trials and systematic reviews regarding the potential benefits and harms of all available treatments. If potentially important evidence is excluded from the evidence base of a systematic review due to lack of standardisation of outcome reporting across trials or insufficient, inadequate quality of outcome and statistical reporting, implications for clinical practice and medical decision making are inevitable.

At the time of writing, we are not aware of a review which systematically considers the reporting of TTE data in all trials within a context outside of oncology or within a context of TTE data without a 'survival' focus.

Therefore, the aim of this chapter was to extend previous work described in Chapter 3.1.1 by systematically reviewing all epilepsy monotherapy trials, with a particular focus on the definitions and reporting of primary outcomes and the use of TTE analyses when making treatment comparisons, in order to better inform and improve reporting standards for future RCTs, systematic reviews and evidence synthesises and therefore clinical practice.

3.2 Methods

3.2.1 Systematic Search

In order to identify all epilepsy monotherapy trials, a systematic search of the Cochrane Epilepsy Group's specialised register was carried out by the group's information specialist (GC, see Appendix 2 for the search strategy).

3.2.2 Eligibility Criteria

3.2.2.1 Inclusion criteria

- Randomised, controlled trials of adults and/or children of all parallel designs (e.g. superiority, non-inferiority, equivalence etc.) reported in a full-text journal article.
- Study participants with epileptic seizures of any kind except those requiring emergency treatment in hospital settings (see Exclusion criteria).
- Monotherapy design studies with drug naive participants and withdrawal/conversion to monotherapy studies (e.g. all participants current AED treatment tapered off during titration phase of study) are included if all participants are converted to monotherapy without an add-on treatment period of any length.
- Two-or-more active treatments are compared, dose-controlled or placebo-controlled designs. ILAE guidelines debate ethical issues in the use of placebo controls [249, 252]. However, such designs provide the majority of the evidence base for some epileptic syndromes of childhood (such as benign epilepsy of childhood with centro-temporal spikes (BECTS))[252], therefore studies of a placebo-controlled design were included.

3.2.1.2 Exclusion criteria

- Non-randomised or observational studies or report which is not a study such as letters, comments on journal articles, clinical summaries, book chapters etc.
- Cross-over studies; such a design cannot adequately measure long-term TTE outcomes of interest and do not replicate routine clinical practice in monotherapy treatment [249].

- Full-text not originally published in English to allow for an assessment of outcome and statistical reporting by English speaking reviewers.
- Studies with an add-on or poly-therapy phase of any length.
- Withdrawal of monotherapy comparisons (withdrawal of treatment to no treatment) as such a design does not aim to evaluate AED efficacy.
- Pharmacokinetic studies (e.g. comparison of two preparations of the same drug) as such studies consider the chemical effect of the drug rather than medical effect.
- Studies in which the randomised comparison made is not between AED treatments (e.g. randomisation of methods of treatment delivery (fast vs. slow titration)).
- Studies of emergency IV treatment with AEDs (e.g. status epilepticus, infantile spasms) as such studies are too short-term to measure TTE outcomes of interest.
- Other types of non-epileptic seizures (e.g. post traumatic, alcohol withdrawal, febrile).
- Studies with healthy controls (no epilepsy) or participants with a single seizure (epilepsy is defined as the occurrence of two-or-more unprovoked seizures)[268].

3.2.3 Screening of Studies

All studies identified in the systematic search of the Cochrane Epilepsy Group Specialised Register were screened for eligibility by SJN. Title and abstract screening was first performed, followed by full-text screening and reference lists of included studies were also screened for further eligible studies. If a full-text manuscript of an abstract could not be found, the abstract was excluded. Any uncertainty over eligibility of studies was discussed with AGM and CTS and a decision was made whether to include or exclude the study. Secondary analyses or multiple publications of the same subset of participants were included if different outcomes were measured and treated as separate 'studies' (i.e. data extraction for each study was performed using only information from a single publication; any online supplementary material linked to the single publication was considered but no external information from related publications was used).

3.2.4 Data Extraction

Data extraction was performed in four stages using a piloted data extraction form (see Appendix 3) converted into a Microsoft Access database which was used to create a database of all extracted data. A screenshot of the database is provided in Appendix 4.

The content of the data extraction form was based on the recommendations for the reporting of outcomes in epilepsy monotherapy studies from the ILAE [249, 252], guidelines

for the reporting of survival outcomes and analyses in Altman *et al* [31], recommendations for the presentation of survival plots in published literature from Pocock *et al* [241] and the summary statistics required to use the indirect estimation methods for TTE analyses as described by Parmar *et al* [6] and Williamson *et al* [13]. Data extraction was performed on all studies by SJN and LS independently extracted from a subset of 10% of studies. Agreement between extractions was good and any discrepancies were resolved by discussion.

The first stage of data extraction was performed on all eligible studies identified in the search. This stage of data extraction included a question which acted as an indicator variable; "Is at least one time-to-event outcome reported?" with possible responses Yes, No or Unclear. If it was clear that no TTE outcomes were reported for the study, data extraction was considered complete. If a study did report at least one TTE outcome or if it was unclear whether a TTE outcome had been reported from the definition of all reported outcomes, further data extraction was performed including study design and characteristics, definition of outcomes, statistical analyses and presentation of results for each reported TTE outcome of the study.

It was anticipated that the following TTE outcomes would be reported in the studies: time-to-withdrawal of allocated treatment, time-to-first seizure, time-to-6, 12 or 24 month remission of seizures and time-to-exiting the study. Extraction was performed according to the definition of the outcome as described in the study publication and TTE outcomes were classified as meeting the definition of one of the pre-defined outcomes above or other.

3.2.5 Data analysis and presentation of results

Numerical results are presented as medians and ranges or numbers and percentages as appropriate. No formal statistical analyses were conducted.

3.3 Results

3.3.1 Results of the search

From an electronic search outlined in Chapter 3.2.1 and conducted on 14th September 2012, 1007 references were identified and downloaded into Endnote Software. Applying inclusion and exclusion criteria outlined in Chapter 3.2.2, 822 references were excluded from title and abstract screening and 185 full-text articles were screened for inclusion in the review. Full-text screening included fifteen conference abstracts which were linked to full-text articles that were not found in the search. Seventy-seven full-text articles were excluded resulting in

108 full-text epilepsy monotherapy studies for eligible for inclusion in the review. See Figure 2 for a study flow diagram of the screening process, including reasons for exclusion at each screening stage and see Appendix 5 for a reference list of the 108 included studies.

The 108 eligible studies were published between 1978 and 2012. The majority of studies (91 out of 108 studies, 84%) were published in speciality Epilepsy or Neurology journals such as Epilepsia (21 out of 108 studies, 19% of total studies), Neurology (13 studies, 12%), Epilepsy Research (12 studies, 11%) and Epilepsy & Behaviour (11 studies, 10%). The remaining 17 studies were published in general (and some high impact) medical journals such as the Lancet, the British Medical Journal and the New England Journal of Medicine. See Table 31 and Table 32 in Appendix 6 for further details of publication dates and journals.

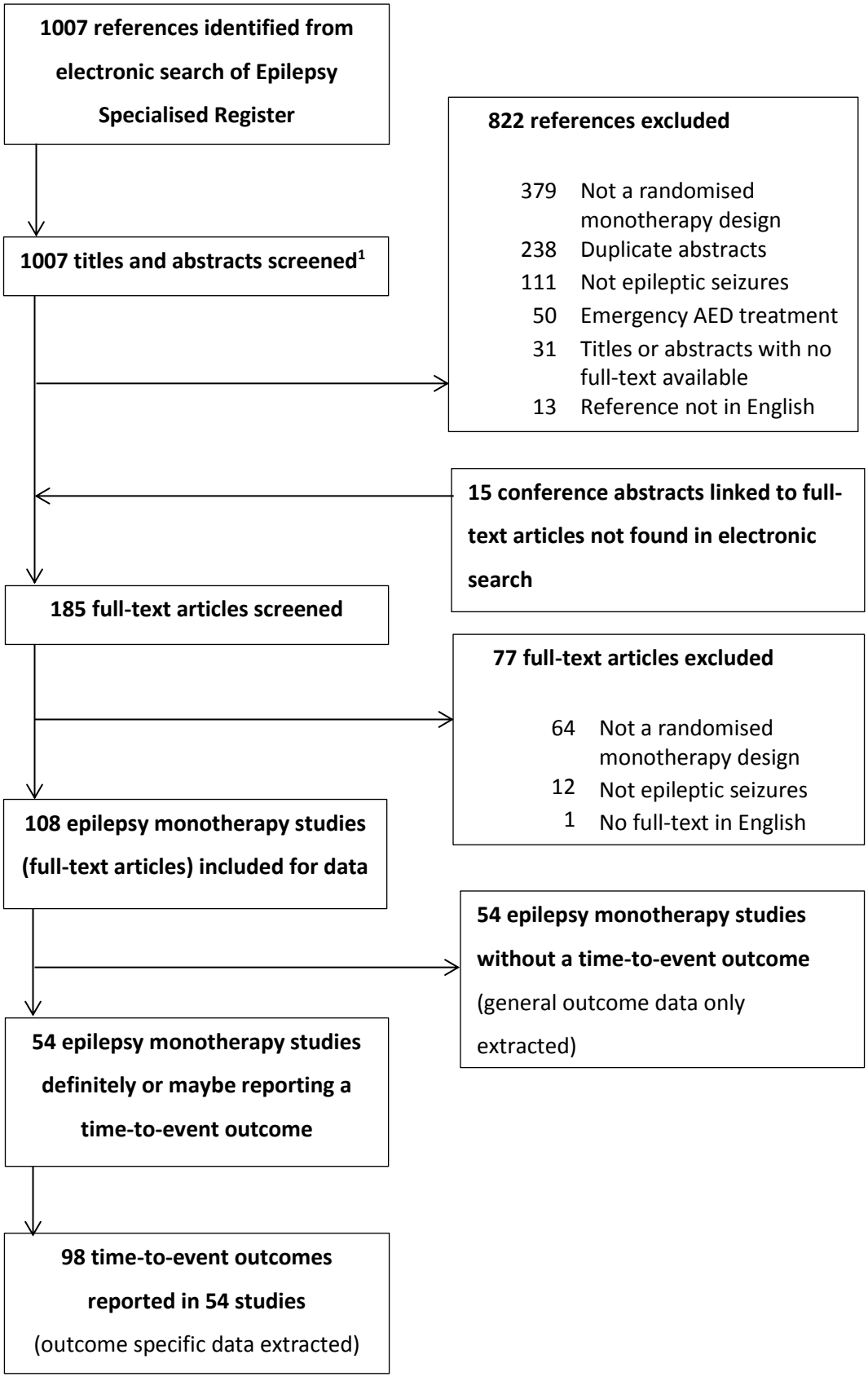
3.3.2 Outcomes reported in Epilepsy Monotherapy studies

Considering all outcomes of any data type (TTE or otherwise), the 108 studies reported a median of 4 outcomes per study (range 1 to 9 outcomes per study). Fifty-four studies out of the total 108 eligible studies (50%) were deemed as not reporting a TTE outcome and data extraction was therefore complete. Outcomes (as reported in study publications) are listed for the 54 studies without a TTE outcome in Table 31 (Appendix 6). No further data extraction was performed for these 54 studies.

Forty-six out of 108 studies (43%) clearly reported at least one TTE outcome and for eight studies (7%), it was unclear whether a TTE outcome had been reported. Uncertainty arose for two reasons:

1. In six studies the definition of the outcome implies a categorical nature (i.e. the outcome is defined as a percentage, rate etc.) or the statistical analysis section of the paper states or implies that the outcome is to be analysed with statistical methods for categorical data (chi-squared test, Fisher's exact test, logistic regression etc.). No statistical methods relating to TTE analyses are described for this outcome, however in the results section a KM curve (or unnamed "survival" plot) is presented for the outcome. As the definition of the outcome is unclear and no time-to-event statistical analyses have been specified, it is unclear whether the censored nature of the data has been properly taken into account and whether the outcome has been analysed appropriately as a TTE outcome.
2. Two studies report the "Mean time-to-withdrawal." The use of 'mean' in relation to TTE data is considered inappropriate due to the likely skewed nature of the data (see Chapter 1.1.4). Therefore it is unclear whether censoring has been taken into account properly in the analyses of these outcomes or if they have been analysed as continuous outcomes.

Figure 2: Study Flow Diagram of identification of eligible epilepsy monotherapy studies



1. A single database was searched so duplication across databases was not applicable.

Outcomes (as reported in study publications) for the 46 studies with at least one TTE outcome and the 8 studies in which a TTE outcome may have been reported (unclear) are listed in Table 32 (Appendix 6). Some of the outcomes reported in the 54 studies were vague in definition such as 'overall assessment of efficacy and tolerability,' 'seizure rating score,' seizure control' and 'adverse events reports.' Further data extraction was conducted on these 54 studies; for this purpose, the 'unclear' outcomes in eight studies (see Table 32 (Appendix 6) for details) were treated as TTE outcomes.

3.3.3 Characteristics of Epilepsy Monotherapy studies reporting at least one time-to-event outcome

See Table 1 and Table 2 for a summary of the reporting of outcomes and statistical analyses in the 54 epilepsy monotherapy studies.

3.3.3.1 Study and Participant Characteristics

See Table 3 for a summary of the study and participant characteristics. Out of the 54 studies, 42 had two treatment arms (78% of total), five had three arms (9%), six had four arms (11%) and one had five arms (2%).

Forty-one studies (76%) followed a monotherapy design and the other 13 studies (24%) followed a withdrawal or conversion-to-monotherapy design (see Chapter 3.1.2 for definitions of designs). Twenty-eight studies (52%) recruited AED-naive individuals with newly diagnosed seizures only, seventeen studies (31%) recruited both AED-naïve newly diagnosed participants and currently untreated participants who had failed a previous AED or with relapsed seizures after remission and nine studies (17%) recruited individuals with refractory / drug resistant seizures only.

The majority of studies (41 studies, 76%) had an active comparator design comparing two-or-more antiepileptic drugs (AEDs), six studies (11%) had a dose-controlled design (i.e. comparing doses of the same AED), four studies (7%) had both a dose-controlled and active comparator design and three studies (5%) were placebo-controlled.

The majority of the studies (42 studies, 78%) were of a superiority design. Seven studies (13%) were of a non-inferiority design and three studies (5%) were of an equivalence design; however, three of these studies did not describe a sample size calculation or statistics relating to an appropriate non-inferiority boundary or equivalence range. One study was of a 'double triangular sequential' design (2%) and the sample size calculation for this design was

described and one study (2%) reported that "the study was not powered to detect statistical differences in efficacy."

Table 3: Study designs and participant characteristics of 54 epilepsy monotherapy studies

Design or characteristic	Number of studies (% of 54 studies)
Therapeutic Design	
Monotherapy	41 (76%)
Conversion / withdrawal to monotherapy	13 (24%)
Participant population	
Antiepileptic drug (AED) naïve participants only	28 (52%)
AED naïve and currently untreated participants	17 (31%)
Refractory / drug resistant seizures	9 (17%)
Type of control	
Active comparator	41 (76%)
Dose-controlled	6 (11%)
Active comparator and dose-controlled	4 (7%)
Placebo-controlled	3 (5%)
Statistical Design	
Superiority	42 (78%)
Non-inferiority	7 (13%)
Equivalence	3 (5%)
Double triangular sequential	1 (2%)
Descriptive only	1 (2%)
Blinding	
Double-blind	36 (67%)
Single-blind	2 (4%)
Open label	13 (24%)
Unclear or not stated	3 (5%)
Funding	
Pharmaceutical funded	29 (54%)
Public funding	9 (17%)
Pharmaceutical and Public funded	5 (9%)
No funding source stated	11 (20%)
Role of the funding source	
Funding source involved	6 (11%)
Funding source not involved	4 (7%)
Funding source involvement not stated	44 (82%)

Thirty-six studies (67%) described a double-blind design, however only 15 of these studies provided any details of who was blinded and how the blinding was achieved. Two studies (4%) described a single-blind design of outcome assessors and 13 studies (24%) had an open-label design. Out of the remaining three studies, one made no mention of blinding and two

described 'partially' blinded designs due to variation in titration periods of the study drugs; it was unclear exactly how this blinding was achieved.

Twenty-nine studies (54%) were described as funded, sponsored or supported by a pharmaceutical industry; specifically, the manufacturing industry of the investigational AED in the trial. Nine studies (17%) were completely publically funded; funding sources were Department for International Development (UK), the Wellcome Trust (UK), International League against Epilepsy (worldwide), National Institute of Health (US) and Veterans Administration Medical Research Service Cooperative studies program (US). Five studies (9%) were part-publically funded and part-industry funded (9%); as above industry funding was provided from the manufacturer of the investigational AED in the trial and public funding sources were the Medical Research Council (UK), the Health Promotion Trust (UK), the Wellcome Trust (UK), the Health Technology Assessment programme (UK) and the Royal Melbourne Hospital Neuroscience Foundation (Australia).

For the remaining eleven studies (20%), no source of funding was disclosed. For eight of these studies, at least one author was affiliated to the manufacturing industry of the investigational AED suggesting pharmaceutical involvement in the trial. Regarding the design, data collection and analysis of the trial, six studies (11%), all pharmaceutical funded, reported that the funding source was directly involved in at least one area and four studies (7%), all publically funded, reported that the funding source had no involvement. The remaining 44 studies (82%) did not describe any involvement of the funding source.

3.3.3.2 Disposition of participants in the study

Twenty-five out of 54 studies (46%) presented a flow diagram of disposition of participants throughout the study. All except one study (98%) stated how many participants were randomised to each treatment arm; the single study reported the number of participants included in analysis but not the number originally randomised to each treatment arm.

Forty-two studies (78%) stated how many randomised participants completed the study and 35 studies (65%) specified the number of participants completing by treatment arm. The remaining 12 studies (22%) did not state how many participants completed the study, however six of these studies had long-term follow-up and no fixed duration, therefore it would be difficult to quantify how many participants actually "completed" the study.

Frequently, different populations (e.g. intention to treat, per-protocol, safety) were included in analyses but it was not always clear how such populations were defined and how many participants were included in each population. Thirty-four studies (63%) reported that participants were excluded from one-or-more analyses in the study and 25 studies (46%) reported exclusions made by treatment arm. The remaining 20 studies (37%) did not report any exclusion from analyses.

There were inconsistencies in participant numbers throughout different sections of the study publication for three studies (5%). In nine studies (17%), the numbers or statistics presented in at least one of tables didn't correspond with results specified in the text.

3.3.3.3 Time frame of the study and extent of follow-up

Thirteen out of 54 studies (24%) reported a diagram of study periods or phases and their respective lengths. Twenty-eight studies (52%) reported the length of the period in which participants were recruited into the study and 20 studies (37%) reported the length of a pre-randomisation baseline or screening phase. Post – randomisation, 41 studies (76%) reported a titration or dose-escalation period and 34 studies (63%) reported a maintenance period and their lengths in the design. Sixteen studies (30%) reported an open label or double-blind extension phase of the study; however only three of these 16 reported the length of extension phase. Forty-five out of 54 studies (83%) reported the frequency and/or times of scheduled follow-up visits during the study.

Forty-seven out of 54 studies (87%) reported the duration of the study; 37 of these studies had a fixed duration (ranging from 28 days to 4.5 years), ten studies reported a maximum study duration (ranging from one to seven years). Four studies (7%) reported that the duration was 'variable' or 'approximate' and three studies (6%) did not report any details regarding study duration.

Twelve studies (22%) reported the minimum length follow-up of the participants in the study, 38 studies (70%) reported the maximum follow-up and 10 of these studies (18.5%) reported both the minimum and maximum follow-up. Seven studies (13%) reported mean follow-up and seven studies (13%) reported median follow-up time. Thirteen studies (25%) reported no extent of follow-up at all. Frequently, it was not clear if extent of follow-up was summarised for all participants or for those completing the study only and summary values of follow-up were rarely reported by treatment arm.

There were inconsistencies in the described time frame or extent of follow-up throughout different sections of the study publication for 8 studies (15%), mainly where follow-up time indicated on a survival plot was different to the extent of follow-up described in the text of the publication (see Chapter 3.3.4.4 for further details).

3.3.3.4 Time-to-event and primary outcomes

In total, 98 time-to-event outcomes were reported in the 54 studies; 23 studies (43%) reported one time-to-event outcome, 22 studies (41%) reported two outcomes, five studies (9%) reported three and four studies (7%) reported four.

According to pre-specified definitions of outcomes of interest for this review (see Chapter 3.2.4), the TTE outcomes reported in the studies were time-to-withdrawal of allocated treatment (reported by 35 studies, 65%; one study reports this outcome twice with slightly different definitions), time-to-first seizure (reported by 27 studies, 50%), time-to-exiting the study (reported by 10 studies, 19%), time-to-withdrawal of allocated treatment due to adverse events (reported by 8 studies, 15%), time-to-6-month remission (reported by five studies, 9%), time-to-12-month remission (reported by seven studies, 13%) and 24-month remission (reported by 5 studies, 9%). Further definitions of these outcomes are discussed in Chapter 3.3.4 and see Table 32 (Appendix 6) for other non TTE outcomes reported.

Thirty-five studies (65%) defined a single primary outcome; 16 of which were TTE outcomes (time-to-exit in seven studies, time-to-withdrawal of allocated treatment in five studies and time-to-first seizure in four studies, see Table 32 (Appendix 6) for other primary outcomes). Four studies defined two primary outcomes, all of which were TTE outcomes (time-to-withdrawal of allocated treatment and time-to-12-month remission for all studies) and three studies defined three primary outcomes, two of which in each study were TTE outcomes (time-to-withdrawal of allocated treatment and time-to-withdrawal of allocated treatment due to adverse events).

Out of 35 studies which define a single primary outcome, 27 studies (77%) described a sample size calculation relating to the primary outcome and for the seven studies defining more than one primary outcome the sample size related to all primary outcomes. Three studies reported a sample size calculation relating to one-or-more outcomes other than the defined primary outcome and three studies without a defined primary outcome report a sample size calculation relating to 'all outcomes.'

The primary recommended outcome for epilepsy monotherapy studies from ILAE guidelines first published in 1998 [249] originates from the outcome of 'retention time' defined in Mattson *et al* 1985 [269] and Mattson *et al* 1992 [270]. Therefore, the reference lists of all studies published after 1985 (53 out of 54 studies, one study was published in 1981 [271]) were checked for reference to either of the Mattson *et al* studies [269, 270] in relation to the outcome of time-to-withdrawal of allocated treatment. Reference lists of the 37 studies published after 1998 were also checked for reference to the ILAE guidelines [249](or updated versions from 2006 [252]) and reference lists of all studies were checked for reference to any other citation relating to the definition of choice of outcomes in the study.

Twenty-three studies (43%) cited one or both of the Mattson *et al* studies [269, 270] and ten studies cited the ILAE guidelines from 1998 [249] or 2006 [252], however only five of these 33 citations seemed to be in relation to the definition of the outcome 'time-to-withdrawal of allocated treatment.' Citations were related to the results of the Mattson *et al* studies [269, 270] or other aspects of study design or conduct. Three studies cited guidance published by the European Medicines Agency Committee for Proprietary Medicinal Products (CPMP) in 2000 [255] or an updated version published in 2010 [251]. Recommendation of primary outcomes was slightly different in this guidance to ILAE guidance [249, 252, 253] (see Chapter 3.1.2 for further details) and the three studies citing these guidelines have employed the recommended primary efficacy outcome of proportion of patients remaining seizure-free for at least six months.

Seventeen studies (31%) reported one or more outcomes in the results section of the study publication which are not defined in the methods section. Furthermore, three studies (5%) defined outcomes in the methods section which were not reported in the results section. Ten studies were not consistent in the definition of outcomes throughout the publication; e.g. use of terms 'time-to-withdrawal of allocated treatment' and 'time-to-exit' interchangeably.

3.3.4 TTE outcomes reported in Epilepsy Monotherapy studies

As stated in Chapter 3.3.3.4, 98 TTE outcomes were reported in 54 studies. Results are now expressed according at the level of outcomes rather than at the level of studies; all proportions are out of 98 outcomes (also see Table 2 for a summary of the reporting of outcomes and statistical analysis at the outcome level and at the study-level).

3.3.4.1. Number of participants and time origin

For 93 out of 98 outcomes (95%), the numbers of participants contributing to the outcome was reported. For 72 outcomes (73%) the number of participants experiencing events for the outcome was reported and for 36 outcomes (38%) the number of participants censored was reported. For 53 outcomes (54%) the number of participants lost to follow-up were reported and for 25 outcomes (26%) the number of participants censored and lost to follow-up were reported separately. In 22 out of 98 outcomes (22%) the numbers of events, censored participants and those lost to follow-up were reported by treatment arm.

For 59 outcomes (60%) the time origin of the outcome was reported; for 43 of these 59 outcomes (73%) this was randomisation, for eleven outcomes (19%) this was the end of the titration period / start of the maintenance period, for three outcomes (5%) this was the first dose of study medication and for two outcomes (3%) this was enrolment in the study.

3.3.4.2. Definition of events and censoring

For 85 out of 98 outcomes (87%), the event of the outcome was clearly defined and for 48 outcomes (49%), the definition of a censored observation was clearly defined. All outcomes which clearly defined censoring also clearly defined an event. For 41 outcomes (42%), loss to follow-up was a censored observation, for 7 outcomes (9%) it was classed as an event and in 23 studies (23%) it was unclear whether those lost to follow-up were treated as events or censored observations. For 27 outcomes (28%) censoring was not mentioned at all. Reporting of the definitions of events and censoring by outcome type is summarised in Table 4.

Under the definition of the outcome 'Time-to-withdrawal of allocated treatment' as defined by the ILAE [249], if a participant withdraws from the study due to lack of efficacy (e.g. recurrence of seizures), poor tolerability (e.g. occurrence of adverse events) or a combination of the two then the withdrawal is classed as an event. If a participant withdraws for other reasons, including reasons not related to the study drug and losses to follow-up, or the participant completes the study without withdrawal then these participants are classed as a censored observation. Although for 29 out of 36 outcomes (78%), the event was well defined, only eight of the outcomes defined treatment withdrawal as in the ILAE definition [249]. The eight outcomes (22%) which did not clearly define an event referred to withdrawals from the study but did not specify which reasons for withdrawal were classified as events. An event for the outcome 'Time-to-withdrawal due to adverse events' was generally defined as "the occurrence of an adverse event leading to treatment withdrawal or premature discontinuation from the study" or similar.

An event for the outcome of 'Time-to-first seizure' was generally defined as the "occurrence of first seizure during the study" or similar. Twelve of these outcomes also specified the seizure type in the definition of the event (i.e. 'occurrence of first generalised tonic-clonic seizure during the study'). An event for the outcomes of 'Time-to-6, 12 or 24 month remission' was generally defined as the "achievement of 6, 12 or 24 months of remission from seizures during the study" or similar. An event for 'Time-to-exit from the study' was generally defined according to listed protocol-defined exit criteria (e.g. occurrence of status epilepticus, increase in seizure rate, emergence of a more severe seizure type or intolerable adverse experience). Outcomes which enforce exit criteria based on seizure recurrence are thought to be of little clinical relevance and do not reflect routine clinical practice [252]. Three outcomes didn't fall under any of our pre-specified definitions (see Chapter 3.2.4) and for two of these outcomes it was unclear if they had been analysed as TTE outcomes (see Chapter 3.3.2). One outcome seemed to be defined as 'time-to-first, second, fifth and tenth seizure' but analysed as 'time-to-seizures;' the definition of this outcome was not clear.

Table 4: Definition of events and censoring in 98 TTE outcomes reported in 54 epilepsy monotherapy studies

Outcome	Event defined	Censoring defined	Unclear how loss to follow-up was defined	Censoring not mentioned
Time-to-withdrawal of allocated treatment (n=36)	29 (81%)	15 (42%)	12 (33%)	9 (25%)
Time-to-first seizure (n=27)	25 (93%)	13 (49%)	6 (22%)	8 (30%)
Time-to-6, 12 or 24 month remission (n=14)	13 (93%)	11 (79%)	1 (7%)	2 (14%)
Time-to-exit from the study (n=10)	9 (10%)	5 (50%)	2 (20%)	3 (30%)
Time-to-withdrawal due to adverse events (n=8)	7 (88%)	4 (50%)	2 (25%)	2 (25%)
Other (n=3)	2 (66%)	0 (0%)	0 (0%)	3(100%)
All outcomes (n=98)	**85 (87%)**	**48 (49%)**	**23 (23%)**	**27 (28%)**

Overall, in only 28 outcomes (29%) was enough information provided regarding events and censoring to allow statistical analysis of the outcome to be replicated (i.e. numbers of events and censored observations provided and clear definitions for both). For three outcomes, the definition of the outcomes presented on the plots was not the same as the definition of the outcome in the text (two outcomes of time-to-first seizure and one outcome of time-to-exiting the study).

3.3.4.3. Statistical analysis and time-to-event statistics

All 54 studies reporting at least one TTE outcome included details of statistical analysis within the methods section of the study publication, however the methods of TTE analysis used and the level of detail reported to describe statistical analysis conducted for each outcome in each study was variable. Furthermore, eight of the studies reporting more than one TTE outcome analysed the outcomes using different statistical methods.

For 48 out of 98 outcomes (49%) due to lack of clear detail given in statistical analysis sections, it would not be possible to replicate statistical analysis. For 11 out of 98 outcomes (11%) no statistical analyses of any kind were reported at all for the outcome and for five outcomes (5%) no statistical analyses of a TTE nature were reported. It was reported that these five outcomes were to be analysed by methods for categorical data (see Chapter 3.3.2 for further details). For 17 outcomes (17%) it was stated only that "methods of survival analysis" (or a similar description) were used and no further details were given. Also, for 16 outcomes, analyses were presented in the results sections which were not specified in the methods sections.

For the 98 TTE outcomes, the following TTE statistics were reported (also see Table 2):

- Survival probability is reported for 60 outcomes (61%)
- Median survival time is reported for 19 outcomes (19%)
- Hazard ratio and 95% confidence interval is reported for 31 outcomes (32%)
- Odds or Risk Ratios and 95% confidence intervals are reported for 14 outcomes (14%)
- Log-rank p-value is reported for 39 outcomes (40%)
- Number of events and a log-rank p-value is reported for 27 outcomes (28%)
- Other p-values (e.g. from Cox Proportional Hazards Model, Generalised Wilcoxon, chi-squared, Fisher's Exact) were reported for 28 outcomes (29%)
- The Observed and Expected number of events is not reported for any of the outcomes
- For ten outcomes (10%), none of the above statistics were reported

Multivariable methods were used in 44 out of 98 outcomes (45%). Forty-two out of the 44 outcomes (95%) used Cox multivariable regression models and one used logistic regression. Commonly specified variables used in multivariable models were gender (10 outcomes), baseline age (18 outcomes), baseline seizure type (25 outcomes), baseline seizure frequency (16 outcomes) country, region or centre (14 outcomes) and duration of epilepsy (5

outcomes). For 6 outcomes the only information given for adjusted variables is that they were “baseline characteristics” or “important prognostic variables.” Reasons for adjusting for specific variables were provided for 31 out of the 44 outcomes (70%) but for 13 outcomes (30%) only an adjusted effect size was presented (i.e. an unadjusted effect size for treatment effect only was not presented). Out of 42 outcomes analysed via Cox PH models, it was stated for only 8 of the outcomes (19%) that the assumption of PH had been checked. In fact, for three outcomes, the PH assumption was found to be violated yet HR and 95% confidence intervals from the model were still presented.

Considering the 35 outcomes of ‘time-to-withdrawal of allocated treatment,’ for 26 outcomes (74%), separate withdrawal rates due to lack of efficacy and poor tolerability are presented and for 11 outcomes, the composite nature of this outcome is taken account of in an analysis; for eight outcomes, separate analyses or subgroup analysis are presented for withdrawal for any reason, due to lack of efficacy and due to poor tolerability and for three outcomes a formal competing risks analysis is presented. It should be noted that subgroup analyses based on post-randomisation information are not generally recommended, and a competing risks analysis is a more appropriate way to analyse this composite outcome.

For 29 out of 98 outcomes (30%) at least one subgroup analysis is reported; in total, 35 subgroups analyses are reported for 29 outcomes. Subgroup analyses performed were by seizure type or epilepsy syndrome (18 outcomes), reason for treatment withdrawal (six outcomes, also see above paragraph), drug plasma concentration (5 outcomes), age (4 outcomes), previous AED use at baseline (one outcome) and seizure frequency at baseline (one outcome). A significant difference in treatment effect between subgroups was demonstrated in 26 out of 35 subgroup analyses (74%).

For 34 out of 98 outcomes (35%) a sensitivity analysis is reported; in total, 38 sensitivity analyses are reported for 34 outcomes. Sensitivity analyses performed were per protocol population only (compared to intention-to-treat population, 13 outcomes), exclusion or re-classification of ‘uncertain’ seizure types (nine outcomes), exclusion of events in a specific period of the study (e.g. seizures during the titration period, seven outcomes), worst-case scenario analysis (i.e. all missing participants in one group are assumed to be non-responders and all missing participants in the other group are assumed to be responders, four outcomes), alternative definition of treatment failure (four outcomes), alternative variables in a multivariable model (one outcome). A significant difference in treatment effect was demonstrated in two out of 38 sensitivity analyses (5%).

3.3.4.4. Survival plots

For 74 out of 98 outcomes (75%) a survival plot was presented. Forty-nine out of 74 plots (70%) were described as 'Kaplan-Meier' plots, 6 were described as 'actuarial' plots' (8%), 5 were described as 'cumulative-incidence' plots (7%) and 14 (15%) were described as 'survival' plots or no definition was provided. For 68 of the plots (92%), a step function was used.

It has been recommended that plots with an upwards direction (i.e. a cumulative incidence plot) are more reliably informative, particularly in the situation of low event rates [241], however methods for the indirect estimation of HR from survival plots require a downwards direction [6, 13]. For 62 out of 74 plots (84%) the plot had a downwards direction and the remaining 12 (16%) had an upwards direction.

For 32 plots (43%) a HR or a p-value was displayed on the graph and in 11 (15%) a measure of precision (e.g. confidence interval of the HR) was displayed on the graph. For twelve plots (16%) effective numbers at risk were reported on or underneath the plots and for ten plots (10%), effective numbers at risk were reported or could be deduced from the text of the publication. For nine plots (12%) censored observations were clearly marked on the plot; for five of these plots, the only marked censoring was at the end of follow-up time.

For 64 plots (86%) different line types were clearly used for multiple curves. For eight plots, coloured lines were used to distinguish between curves however it would be difficult to distinguish between these line colours on a grayscale copy of the publication. For two plots, it was not possible to distinguish between the lines. For 66 plots (89%) a clear legend was provided for the graph and for seven plots no legend was provided but labels were written next to the lines or underneath the graph. For one plot, no legend or labels were provided.

It is not necessarily recommended to display the entire vertical axis of a survival plot and to do so may *"inhibit the ability to discriminate between treatments."* It is also not necessarily recommended to present the whole extent of follow-up for studies with long durations and that the horizontal axis should be *"halted once the proportion of patients free of an event, but still in follow-up, becomes unduly small* [241]" However, enlarging any differences between the lines on the graph by presenting only part of the axis could lead to erroneous conclusions regarding importance and significance of the lines, particularly at later follow-up times with fewer participants, if not also presented with an effective number at risk and / or a measure of statistical uncertainty on the plot [31].

For 17 plots (23%) the vertical axis did not extend from 0 to 1 (survival probability), the smallest proportion of the vertical axis displayed was 0 to 0.15. For nine plots from study of duration of 12 months or less, the entire study duration is not displayed on the plot and for two plots from studies of two to five years duration, only a subset of the follow-up time is displayed on the plot. Only three of these plots do present numbers at risk which would aid with interpretation of the graphs.

For eight plots (11%) the vertical axis scale was inappropriate; in other words, a plot with a downward direction and an increasing 'cumulative' axis (or vice-versa). Another seven plots (9%) had unlabelled or unclear labels on either the vertical or horizontal axis. Five plots (7%) were presented without any title or label therefore the content of the graph had to be deduced from the text of the publication.

3.4 Discussion

While an IPD approach to analysis is considered to be the 'gold-standard,' particularly for synthesis of TTE outcomes [6, 23, 24], such an approach is time consuming and resource intensive; therefore an analysis of AD may be considered as an alternative. Reviews of epilepsy monotherapy treatments have been conducted using an IPD approach by the Cochrane Epilepsy Group [59-67, 69], partly justified by the expectation that the AD required for such an approach to analysis would not be adequately and consistently reported in the relevant publications.

This chapter systematically examines all trials of an epilepsy monotherapy design (whether included in a Cochrane Epilepsy IPD review or not) in terms of the reporting of important TTE outcomes of interest in relation to their inclusion in an AD synthesis.

3.4.1 Summary of key results and implications

This systematic review considers the reporting of 98 TTE outcomes in 54 epilepsy monotherapy RCTs published between 1978 and 2012 in a range of speciality and general medical journals. In total, half of the studies considered reported to have analysed at least one TTE outcome. However, definitions and methodology for analysing such outcomes greatly varied in detail, to the extent that it was not completely clear if outcomes had been defined and analysed appropriately as TTE outcomes in 8 out of the 54 trials (15%).

The majority of studies described study and participant characteristics well relating to design, eligible population and source of funding. However, at least one reporting inadequacy in

participant numbers, time frames and definitions of outcomes in different sections of the same trial publication was identified in 38 out of the 54 studies (70%); within half of the 14 studies which were completely or partially publically funded, 20 out of 29 (69%) studies which were industry funded and all of the 11 studies without a source of funding declared.

Less than half of the outcomes considered (49%) were clearly defined, in terms of the definition of event, of censoring and of the time origin of the analysis. For most of the outcomes which were not clearly defined, it was definition of censoring which was not mentioned or not clear; in fact for 28% of TTE outcomes, censoring was not mentioned at all throughout the study publication, despite the fundamental methodological importance of censoring to TTE analysis. Although two-thirds of the studies reported the ILAE recommended outcome of 'time-to-withdrawal of allocated treatment' [249], only eight of these outcomes followed the ILAE definition and others differed in their definitions of events and censored observations in analysis (for example, whether all withdrawals were analysed as events or only those related to the allocated treatment).

Further inadequacies relating to reporting of statistical analysis methods were noted with sufficient details to replicate analyses provided for only 49% of outcomes. There was also potential indication of selective reporting of subgroup analyses, with 74% of subgroup analyses reported showing significant results compared to only 5% of sensitivity analyses.

From an evidence synthesis perspective, HRs and 95% CIs were presented for only 32% of outcomes. Considering the other 68% of outcomes, indirect estimation of HRs [6, 13] would be possible for only a small minority with observed and expected number of events presented for no outcomes, number of events and a log-rank p-value presented for 18 outcomes and a survival plot with expected number at risk and/or minimum and maximum follow-up reported for only six outcomes (see Table 2 for full details). Furthermore, given the variability in definition of events and censoring for the outcomes of interest such as time-to-withdrawal of allocated treatment as mentioned above; it likely that the synthesis of aggregate HRs from these studies would not be appropriate and standardisation of outcome definition via IPD analysis would be preferred.

In summary; findings of this review imply that screening for eligible studies for an AD-MA of epilepsy monotherapy studies would be feasible with the majority of studies clearly reporting features relating to design and participant population. However, the necessary summary statistics needed to perform an AD-MA tend to be reported for only around one in three outcomes, and it may not be appropriate to synthesise this data due to variation in outcome

definition. Overall, the results of this review confirm the assumption made by the Cochrane Epilepsy group that re-analysing IPD is the only feasible approach for meta-analysis of epilepsy monotherapy treatment.

3.4.2 Strengths and limitations

At the time of writing, we do not know of any other review which systematically considers the reporting of TTE data in a context other than oncology. Previous reviews have considered a subset of studies from a specific time period or from specific journals whereas no date or journal restriction was made in this review. Therefore results reflect the reporting standards across 35 years and 24 speciality and general medicine journals.

While the focus of this review was to examine the reporting of published TTE data in relation to inclusion of AD in epilepsy monotherapy reviews rather than to examine differences over time or across journals, no clear differences were apparent by year or journal of publication. This was not examined statistically so it cannot be ruled out that such differences in reporting by year or publication journal may exist. However given that a systematic review would generally not make exclusions based on year or journal of publications, any differences are of little relevance to this review.

Furthermore, some of the trials were reported before the introduction of CONSORT minimum reporting standards in 1996 [244]. However as discussed in Chapter 3.1.1, the CONSORT statement makes very little reference to reporting of TTE analyses and statistics specifically so while reporting of general information in the epilepsy monotherapy trials may have improved following the introduction of the CONSORT statement, it is unlikely that such guidelines will have had much impact on the reporting of TTE analyses.

To minimise the number of references to screen for this work, only a single source was searched to identify eligible studies for this review. This is potentially a limitation as multiple electronic databases and grey literature are usually searched for a systematic review. However the Cochrane Epilepsy Group Specialised register is compiled from regular electronic searches of the Cochrane Central Register of Controlled Trials (CENTRAL) and MEDLINE along with hand searches of relevant epilepsy journals and conference abstract booklets. All epilepsy monotherapy studies that had been included in Cochrane reviews and all monotherapy studies known by clinical expert (AGM) up to 2012 at were identified by the search therefore it is unlikely that any studies relevant to the review were missed.

An additional limitation is that external information to the publication of each study was not taken into account for the review (e.g. if a previous publication was referenced in relation to a sample size calculation, it was recorded that a sample size calculation was not recorded in the given publication). In reality when conducting a systematic review and meta-analysis, all relevant information would be considered and original trial authors may be contacted to request unpublished data. Therefore, some of the results presented here may be underestimated (reflecting a worst-case scenario), where information may be available from related sources. It does however seem unlikely that a TTE outcome would be defined in one (non-protocol based) publication and results presented in another; therefore it is unlikely that reporting rates of summary statistics and other results have been underestimated by not considering related publications.

3.4.3 Comparison to previous work

Chapter 3.1.1, Table 1 and Table 2 summarise the findings of previous reviews of aggregate TTE data. It should be noted that previous reviews have varied in characteristics, inclusion criteria and objective (such as examining change in reporting standards over time [242], assessing potential impact on health technology assessment [243] and determining feasibility of indirect estimation methods [144]), therefore all comparisons made between the reviews and the current review are informal and narrative and results of each should be interpreted within the context and objective of the review.

Table 2 demonstrates a wide variability in the results across all of the reviews regarding the reporting of measures of follow-up time, sample size calculations and effect sizes and measures of precision; particularly HRs and 95% CIs required for synthesis of TTE data reported in between 3% and 87% of reviewed articles. Reporting of other summary statistics such as survival probability, log-rank p-values and median survival times was also very variable; reporting rates ranged from 2% to 97% of reviewed articles. Consistently, the majority of studies reviewed (86% to 100%) presented survival curves, a useful graphical representation of TTE outcomes. However the number of studies presenting the effective number at risk, essential for properly interpreting such plots, ranged from 8% to 66% and other reporting inadequacies such as inappropriate axes or legends and poor graphical quality were noted.

Results of this review agree with previous work that the statistics required for indirect estimation of TTE measures required for AD-MA published statistics [6, 13] are not commonly

reported either [141, 144, 160, 161]. Michiels *et al* [160] argue that meta-analysis can be conducted using the difference in median survival rather than HRs as the former measure is more commonly presented. However, in this review median survival was actually reported less frequently than HRs (26 % of outcomes compared to 32% of outcomes).

Poor reporting of outcome definitions was also consistently noted across the reviews. While over 70% of reviewed studies clearly define an event of interest; between 45% and 75% specified the number of events, 30% to 58% defined which individuals would be censored in analysis and 26% to 56% described how losses to follow-up were handled in analyses. Previous reviews of oncology trials [236, 237] have commented on the impact of differently defined endpoints on the ability to conduct in meta-analysis. Arguably commonly used endpoints in oncology such as overall survival have fairly intuitive definitions (i.e. deaths during the trial are events and individuals who do not die are censored at their last follow-up), whereas the commonly used TTE outcome in epilepsy monotherapy trials such as 'time-to-withdrawal of allocated treatment,' have more complex and potentially variable definitions, resulting in further difficulties in conducting an AD meta-analysis of such outcomes is even more difficult in this context.

Recent work has also noted that the lack of statistical detail, particularly relating to validation of statistical assumptions for TTE analyses and how unvalidated summary statistics can be interpreted for secondary analyses including meta-analysis [242, 243]. This review shows similar results with confirmation of the PH assumptions for less than 20% of outcomes analysed by a Cox PH model. These findings provide further justification of why an AD approach is not feasible for reviews of epilepsy monotherapy treatment.

All previous reviews have concluded that an improvement is needed in the quality of the reporting of aggregate TTE data and both the first review [31] and one of the most recent reviews [242] have proposed guidelines for the reporting of survival analyses. Abraira *et al* [242] noted in 2013 that the EQUATOR (Enhancing the QUAlity and Transparency of health Research) network [272], an initiative to promote transparent and accurate reporting of health research publications, did not report any recommendations specific to survival analysis at the time. Up to May 2017, the EQUATOR network online database (http://www.equator-network.org) provides 320 different reporting guidelines relating to many areas and many different designs of clinical studies, but still no recommendations for the reporting of TTE analyses.

3.4.4 Concluding remarks

In conclusion, in line with all previous work conducted in this area, the current systematic review has shown concerning reporting inadequacies relating to the definition, analysis and reporting of TTE outcomes in epilepsy monotherapy trials and the results of this review are sufficient to confirm that an AD-MA based on the published information presented in these trials would not be feasible or recommended to inform clinical practice.

Results of this review are perhaps not surprising given conflicting advice within guidelines, both epilepsy specific guidelines [249, 251-253, 255] and also analysis and reporting guidelines [244-246]. Recommendations specific to epilepsy trials also place a lot of emphasis on study design features, particularly blinding of studies, rather than the accuracy and quality of statistical analysis and reporting in studies; essential aspects for the synthesis of trial results to inform clinical practice and future research.

Findings of this systematic review suggest that consistency of recommendations, supported by clinical reasoning, is needed within epilepsy specific guidelines in addition to further emphasis on the transparent reporting of published results.

Findings of this review also further highlight calls from previous reviews for the urgent development of minimum reporting standards for TTE analyses. In the continuing absence of the development of such standards, use of the suggested guidelines from previous work [31, 242] by journal editors and peer reviewers when considering study publications using TTE analyses would greatly improve reporting rates and in turn facilitate the conduct of AD-MAs and syntheses with TTE endpoints.

Chapter 4: IPD retrieval: A systematic review of individual participant data meta-analyses (IPD-MA)

4.1 Introduction

IPD-MA is widely regarded as the gold-standard approach to the synthesis of clinical trial data with many documented advantages over traditional AD-MA. Recent years have shown a sharp increase in the number of published IPD-MAs [33-35]. An average of 49 were published each year between 2005 and 2009 [34] and recent estimates suggest an increase of around four published IPD-MAs per year [35]. IPD-MAs have been shown recently to directly influence the design and conduct of clinical trials [29] and clinical practice guidelines [30].

While IPD-MAs may offer many advantages, it is well recognised that greater resources are required to conduct them [22-24] and the use of IPD in meta-analyses does not guarantee freedom from biases. IPD-MAs are subject to a risk of selection bias and 'availability bias;' in that they may only include studies for which IPD is made available, which may not be representative of the whole evidence base [24, 40, 51]. IPD-MAs may be delayed or abandoned owing to unclear data requesting procedures or barriers to accessing IPD [89, 273-275]. Review articles have shown that around a quarter of IPD-MAs published up to 2001 [32], up to 2005 [40] and even as recently as 2012 [35] obtained IPD for less than 80% of eligible participants. These reviews also reveal poor reporting particularly in regard to the amount of included IPD, with between 10 and 20% of IPD-MAs not clearly stating how many studies and participants were eligible, requested and included in analysis [32, 33, 35, 40]. The most recent of these reviews found that reasons for unavailability of IPD were reported in only 23% of a sample of 100 IPD-MAs [33].

Despite a growing increase in the popularity of IPD-MAs [34, 35]; in the context of all published meta-analyses, an IPD approach is still taken in only a small minority [33] with meta-analysts reporting lack of resources, lack of time and difficulty of such an approach as barriers to conducting an IPD meta-analysis [276]. Another barrier to IPD-MA experienced within our own research group [277], has been the direction in which to address IPD requests, particularly for studies involving a pharmaceutical sponsor.

The culture of clinical trial data sharing has changed in recent years. Authors of published trials have reported an increased willingness to share data in surveys conducted in 2011 [77, 88] compared to an empirical study conducted in 2009 [89]. The publication of data

transparency strategies and policies by the Institute of Medicine [84] and the European Medicines Agency [85], a proposed policy by the International Committee of Medical Journal Editors [86, 87] and initiatives across the wider research community as a whole [77-80] may go some way to improving the sharing of IPD. Indeed, the launch of data sharing initiatives such as Clinical Study Data Request (CSDR)[91], a platform allowing researchers to request IPD from nearly 3000 clinical trials of thirteen pharmaceutical sponsors, should make access to IPD easier and faster. However, researchers have reported mixed experiences of using data sharing portals such as CSDR suggesting that the increased safeguards may have an unintended negative impact on the conduct of IPD-MA [110-113].

The aim of this Chapter is to examine whether the shift in attitudes and awareness, and the increased number of options available for accessing IPD, is reflected by a positive impact on IPD-MA. This Chapter presents a systematic review of all published IPD-MAs to assess whether availability of IPD has improved over time, and explore characteristics associated with the retrieval of IPD. The primary aim of this systematic review was to investigate whether the success rate of retrieving IPD for the purpose of IPD-MA has increased over time. The secondary aim of the systematic review was to explore the characteristics associated with IPD retrieval.

The work contained in this Chapter has been published in the British Medical Journal [278].

4.2 Methods

4.2.1 Systematic search methods

The following databases were searched: MEDLINE, Central, SCOPUS, Web of Science, CINAHL Plus and PsycINFO. The search strategies for each database are described in Appendix 7 and were based on a systematic search strategy of an earlier review of Riley *et al* 2007 [40]. Databases were searched from June 2005 (end date of the Riley *et a*l 2007 [40] search) up to June 2014 initially and all systematic searches were updated in August 2015. The reference lists of two previous large reviews of IPD-MAs were also consulted; reference list of Riley *et a*l 2007 [40] was provided by the first author on request and the reference list of Huang *et al* 2014 [35] was available as an online appendix to journal publication of the review.

Articles identified from electronic databases and the reference lists of the previous reviews were exported to Endnote version X7, lists merged and duplicates removed. One reviewer (SJN) performed title, abstract and full-text screening of articles identified in electronic

searches according to Eligibility Criteria described in Chapter 4.2.2. The principle reason for exclusion was recorded for relevant articles. Any uncertainties were discussed with CTS and resolved. For accuracy, two authors (BD and SR) also screened a random sample of between 50 and 100 identified articles for eligibility; agreement between the independent screening (SJN and BD or SR) was good and any discrepancies were discussed and resolved.

4.2.2 Eligibility criteria

IPD-MAs of studies of all types (randomised, observational, diagnostic etc.) and all clinical areas published in English were eligible for inclusion. Articles were included if IPD was requested from original study investigators, if IPD was already available to review authors or if review authors were able to extract IPD from published articles.

Methodological articles, conference abstracts, review protocols and non-clinical reviews (e.g. engineering articles etc.) were excluded. Articles including the analysis of IPD from a single study as a supplement to an AD-MA or articles in which the primary objective of the analysis was not to estimate a pooled effect size (e.g. prognostic model validation studies, cost-effectiveness analysis) were excluded as inclusion criteria of studies in such analyses are generally selective and related to the objective of the analysis (e.g. an estimate of prognosis or cost is provided) rather than the availability of IPD. Where duplicate publications relating to the same IPD-MA were identified (e.g. identical publication across multiple journals) the most recently published article was retained. Updates of analyses (e.g. updated Cochrane Reviews) were included if at least one new eligible study was identified for the analysis.

4.2.3 Data extraction

Information was extracted from eligible IPD-MAs using a piloted data extraction form (see Appendix 8). The data extraction form was piloted by three reviewers (SJN, BD and SR) extracting information from a sample of IPD-MAs referenced in the Riley *et al* 2007 [40]; following pilot extractions, content of the data extraction form was discussed and the final data extraction form (Appendix 8) was used to extract information from all IPD-MAs identified in the searches described in Chapter 4.2.1.

Information extracted from IPD-MAs was year of publication, authorship policy, source of funding, clinical area, type of studies, type of analysis, number of eligible studies providing IPD or AD, reasons for IPD not being provided and details of any additional or sensitivity analyses performed to account for missing IPD.

Reasons for IPD not being provided and sensitivity analyses were recorded as free text by all reviewers and later classified into broad categories. Clinical area was also recorded as free text and later classified in broad categories based on the clinical areas covered by the review groups of the Cochrane Collaboration. All classification of extracted free text was performed by one author (SJN). Type of analysis was classified as follows on extraction:

- Systematic IPD-MA; where a systematic search aiming to identify all eligible studies was performed.
- Pooled or 'opportunistic' analysis [33]; where an existing IPD database or IPD of a collaborative group was analysed without an attempt to systematically identify all eligible studies. Such analyses, by definition, used 100% of eligible IPD in analysis.
- Other analysis; any other approach to IPD-MA which does not fit into either of the above definitions.

Where published articles presented multiple IPD-MAs addressing different research questions with different eligible cohorts for IPD-MAs, information was extracted for each IPD-MA. If multiple analyses were presented for the same IPD-MA (e.g. analysis of several outcomes), information was extracted on the maximum amount of IPD provided, even if all IPD provided were not used in IPD-MA.

One author (SJN) extracted information from all eligible articles and three authors (BD, SR, LW) independently extracted from a subset of around 40% of the eligible articles. Agreement between authors was good and any discrepancies were resolved by discussion.

4.2.4 Statistical analysis and presentation of results

The primary aim of analysis was to examine the IPD retrieval rate (i.e. the number of participants IPD was provided for divided by the number of participants identified as eligible for analysis) over time and the secondary aim was to explore the characteristics associated with IPD retrieval.

Multivariable logistic regression was performed to examine associations between IPD-MAs characteristics and proportion of IPD retrieved. Proportion of IPD retrieved (dependent variable of interest) was highly skewed, despite attempts at transformation, as few IPD-MAs retrieved a very small proportion of data (i.e. less than 20% of IPD retrieved). It was therefore deemed most appropriate to dichotomise this variable to:

- Complete IPD retrieval rate (100% compared to less than 100% or unknown proportion of IPD provided)
- High IPD retrieval rate (at least 80% compared to less than 80% or unknown proportion of IPD provided)

Dichotomisation cut-off (i.e. 80% compared to less than 80%) was chosen to allow comparison with retrieval rates in previous reviews [32, 33, 35, 40]. The following variables were included in the model and results for all variables included in the model are presented regardless of statistical significance; no model selection techniques were used:

- Age of publication of the IPD-MA (calculated as years before 2016, log transformed due to skew, let this variable be x_1)
- Number of participants eligible for inclusion in IPD-MA (log transformed due to skew, x_2)
- Study design (inclusion of randomised studies only in IPD-MA versus other study designs; non-randomised studies, diagnostic test accuracy studies or a combination of randomised and non-randomised studies, x_3)
- Cochrane IPD-MA (IPD-MA performed as a Cochrane Review compared to non-Cochrane IPD-MA, x_4)
- Authorship policy (individual authorship for those providing IPD or collaborative group versus no authorship policy, x_5)
- Source of funding (IPD-MA with a commercial source of funding (pharmaceutical or manufacturer) versus other funding: non-commercial sources of funding only, no funding or no information regarding funding provided, x_6)

In other words, let Z_j be the dependent variable in the j^{th} IPD-MA such that $Z_j = 1$ if 100% of IPD was retrieved and $Z_j = 0$ if less than 100% of IPD was retrieved (or equivalently, $Z_j = 1$ if at least 80% of IPD was retrieved and $Z_j = 0$ if less than 80% of IPD was retrieved) and let the probability that $Z_j = 1$ be $\hat{p}_j$. Multivariable logistic regression is performed as follows:

$$log\left(\frac{\hat{p}_j}{1-\hat{p}_j}\right) = \beta_1(log(x_{1j})) + \beta_2(\log(x_{2j})) + \beta_3 x_{3j} + \beta_4 x_{4j} + \beta_5 x_{5j} + \beta_6 x_{6j} \quad \text{(Equation 27)}$$

Where $x_{1j} \dots x_{6j}$ are the explanatory variables of the j^{th} IPD-MA (see above) and $\beta_1 \dots \beta_6$ are the resulting regression coefficients associated with each of the explanatory variables.

Results of multivariable regression are presented as odds ratios and 95% confidence intervals. Other numerical results are presented as medians and ranges or numbers and percentages as appropriate.

4.2.5 Additional and sensitivity analyses

Additional and sensitivity analyses were conducted to investigate the assumptions made in the primary multivariable logistic regression:

1) Univariable analysis was performed to examine the effect of each variable independently on IPD retrieval rate.
2) Further examination of the association of authorship policy on IPD retrieval; authorship policy redefined as no authorship policy (reference), individual authorship or collaborative group.
3) Inclusion of the variable 'Type of Study' (defined as drug or device (interventional, reference), non-drug (interventional), diagnostic test accuracy or epidemiological study) in the multivariable model.
4) Exclusion of IPD-MAs from analysis with no information regarding funding reported.
5) Assuming the following scenarios for 257 IPD-MAs for which the proportion of IPD retrieved could not be calculated:
    a. Less than 80% of IPD was retrieved
    b. 80% or more IPD was retrieved
    c. 100% of IPD was retrieved

    An additional analysis was also performed to examine characteristics associated with non-reporting of the proportion of IPD retrieved for IPD-MA.
6) Use of fractional logistic regression with proportion of IPD retrieved (dependent variable) expressed as a fraction between 0 and 1 as an alternative to logistic regression [279]. Under this approach, rather than constraining dependent variable Z_j to take a value of 1 or 0, Z_j is modelled as a fraction within the interval $(0, 1)$. Papke and Wooldridge [279] demonstrate the Bernoulli log-likelihood function for a fractional logit model:

$$\log(\boldsymbol{\beta}) = Z_j log\left(\frac{\exp(\boldsymbol{X_j\beta})}{1+\exp(\boldsymbol{X_j\beta})}\right) + (1 - Z_j) log\left(1 - \left(\frac{\exp(\boldsymbol{X_j\beta})}{1+\exp(\boldsymbol{X_j\beta})}\right)\right) \quad \text{(Equation 28)}$$

    Where $\boldsymbol{X_j}$ is a vector of explanatory variables ($x_{1j} \dots x_{6j}$ are the explanatory variables of the j^{th} IPD-MA in this example) and $\boldsymbol{\beta}$ is a vector of the resulting regression coefficients associated with each of the explanatory variables ($\beta_1 \dots \beta_6$ in this example).
7) Multivariable logistic regression of the proportion of study data retrieved (i.e. the number of studies IPD was provided for divided by the number of studies identified as eligible for analysis).

4.3 Results

4.3.1 Results of the search

Systematic searches (outlined in Chapter 4.2.1) identified 1278 eligible articles describing 1280 IPD-MAs published to August 2015. See Figure 3 for study flow diagram of the searching and screening process and a reference list of eligible articles is available as an online Appendix to Nevitt *et al* [278]).

A non-systematic method of identifying studies for inclusion (IPD-MA defined as pooled analyses or other analyses, see Chapter 4.2.4) such as collaboration of a group of researchers with available IPD, had been used in 520 IPD-MAs. Within these 520 non-systematic IPD-MAs, the number of eligible studies was reported in 516 (99%) IPD-MAs with a median number of eligible studies of 7 (range 2 to 287). The number of eligible participants was reported in 501 (96%) systematic IPD-MAs with a median of 3633 (range 16 to 2,051,158) participants.

For the remaining 760 IPD-MAs, a systematic approach was taken to identify all eligible studies. The number of eligible studies was reported in 746 (98%) IPD-MAs with a median of 14 (range 2 to 923) studies. The number of eligible participants within an IPD-MA was reported in 510 (67%) systematic IPD-MAs with a median of 2369 (range 16 to 33369) participants. In 14 (2%) of the 760 systematic IPD-MAs, it was unclear how many studies were eligible and in 250 (33%) it was unclear how many participants were eligible; mainly as the number of participants without available IPD excluded from analysis was not stated.

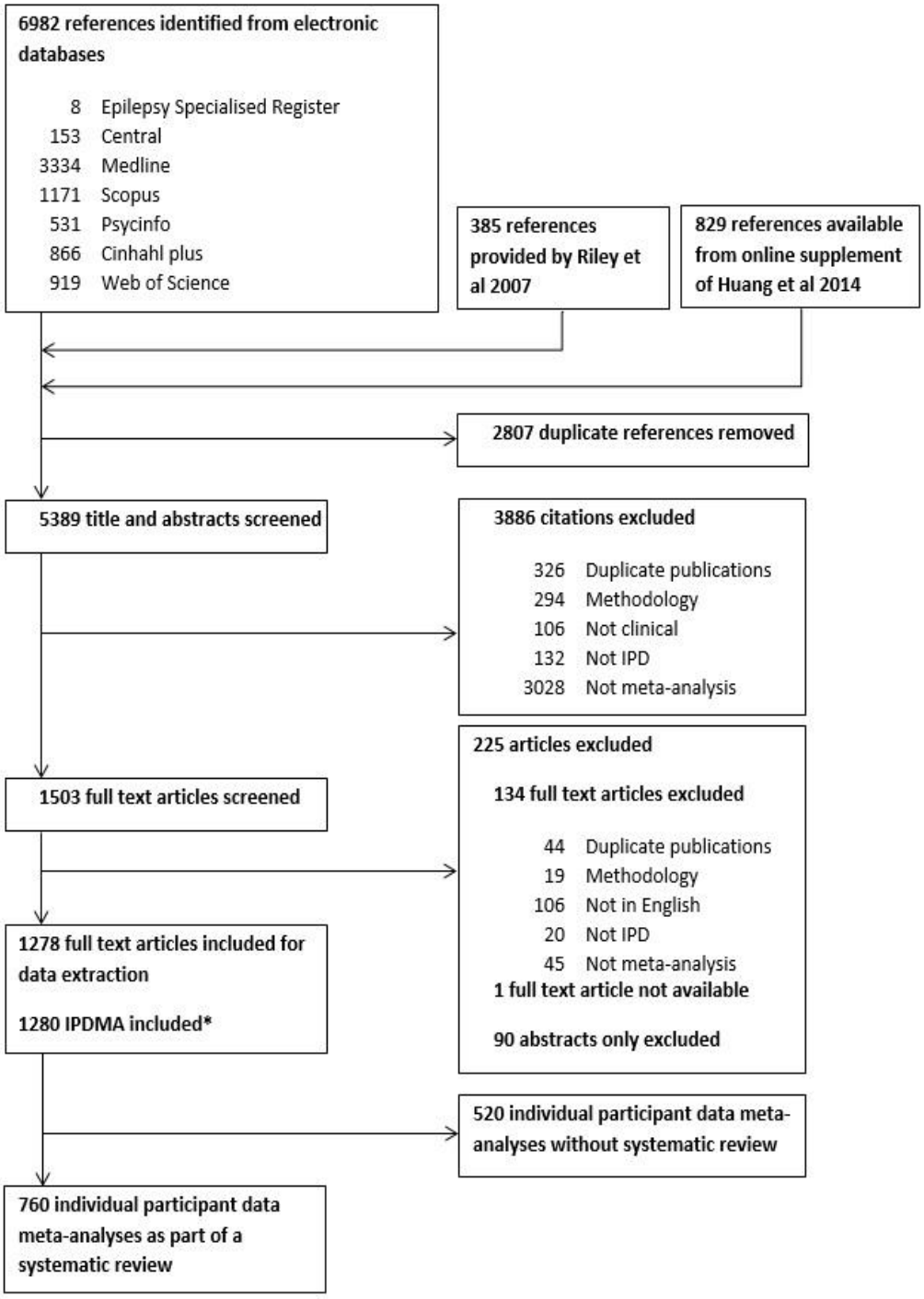
These median values imply that while systematic IPD-MAs identify more eligible studies, non-systematic IPD-MAs tend to identify and include more eligible participants.

4.3.2 Characteristics of IPD-MAs

Table 5 presents the characteristics of the 1280 included IPD-MA according to systematic or non-systematic design.

Within 85 systematic IPD-MAs, IPD was extracted from study publications rather than requested from original study authors or sponsors. Characteristics of these 85 systematic IPD-MAs are presented in Appendix 8.

Figure 3: Study Flow Diagram of identification of eligible IPD-MA



* two full-text articles each reported two IPD-MA

Table 5: Characteristics of 1280 individual participant data reviews

IPD-MA Characteristic	Type of Analysis		
	Systematic IPD-MAs (n and % of total)	Non-systematic IPD-MAs (n and % of total)	Total
Total	**760 (59%)**	**520 (41%)**	**1280**
Year of publication of IPD-MA			
1987 – 1995	20 (61%)	13 (39%)	33
1996 – 2000	72 (65%)	39 (35%)	111
2001 – 2005	116 (57%)	88 (43%)	204
2006 – 2010	195 (56%)	152 (44%)	347
2011 – 2015	357 (61%)	228 (39%)	585
Clinical area of IPD-MA			
Breast Cancer	40 (62%)	25 (38%)	65
Cancer (other)	53 (65%)	28 (35%)	81
Cardiology	105(49%)	110 (51%)	215
Central Nervous System, Neurology and Brain Injury	50 (62%)	31 (38%)	81
Cervical Cancer and Ovarian Cancer	16 (59%)	11 (41%)	27
Diabetes and Endocrinology	30 (63%)	18 (37%)	48
Gastroenterology, Colorectal and Gastric Cancer	49 (56%)	39 (44%)	88
Gynaecology, Pregnancy and Neonatology	35 (88%)	5 (12%)	40
Haematology, Leukaemia and Blood Cancer	43 (72%)	17 (28%)	60
Head and Neck Cancer	16 (64%)	9 (36%)	25
Hepatitis and Liver Disease	19 (56%)	15 (44%)	34
HIV	17 (55%)	14 (45%)	31
Infection and Infectious Diseases	31 (70%)	13 (30%)	44
Injuries and Wounds	21 (58%)	15 (42%)	36
Lung Cancer	32 (76%)	10 (24%)	42
Mental and Psychiatric Disorders	32 (48%)	35 (52%)	67
Musculoskeletal and Pain	34 (52%)	32 (48%)	66
Other[1]	26 (48%)	28 (52%)	54
Otolaryngology , Ophthalmology and Periodontology	22 (76%)	7 (34%)	29
Renal and Urology	17 (61%)	11 (39%)	28
Respiratory and Pulmonary	21 (60%)	14 (40%)	35
Stroke, Thrombosis and Hypertension	51 (61%)	33 (39%)	84
Design of included studies			
Randomised	405 (58%)	288 (42%)	693
Non-Randomised	253 (57%)	194 (43%)	447
Diagnostic Test Accuracy	34 (97%)	1 (3%)	35
Both Randomised and Non-Randomised	68 (65%)	37 (35%)	105

Type of included studies			
Diagnostic Test Accuracy	34 (97%)	1 (3%)	35
Drug or device	348 (54%)	291 (46%)	639
Epidemiology / Risk Factor	185 (52%)	173 (48%)	358
Non-drug (interventional)	193 (78%)	55 (22%)	248
Type of IPD-MA			
Cochrane Review	64 (100%)	0 (0%)	64
Non Cochrane Review	696 (57%)	520 (43%)	1216
Authorship Policy			
Individual authorship	243 (42%)	337 (58%)	580
Collaborative Group	264 (60%)	177 (40%)	441
None[2]	253 (98%)	6 (2%)	259
Source of Funding			
Non-commercial[3]	383 (64%)	218 (36%)	601
Commercial[4]	72 (42%)	101 (58%)	173
Mixed[5]	35 (36%)	62 (64%)	97
No funding	77 (73%)	28 (27%)	105
Not stated	193 (63%)	111 (37%)	304
Number of eligible studies			
2 to 5	102 (32%)	214 (68%)	316
6 to 10	174 (57%)	130 (57%)	304
11 to 15	120 (63%)	72 (37%)	192
16 to 20	87 (81%)	21 (29%)	108
21 to 30	101 (77%)	31 (23%)	132
31 to 40	50 (70%)	21 (30%)	71
41 to 50	29 (85%)	5 (15%)	34
over 50	83 (80%)	22 (20%)	105
Not stated	14 (78%)	4 (22%)	18
Number of eligible participants			
under 100	18 (86%)	3 (14%)	21
101 to 200	20 (65%)	11 (35%)	31
201 to 500	45 (56%)	35 (44%)	80
501 to 1000	67 (56%)	52 (44%)	119
1001 to 5000	198 (51%)	187 (49%)	385
5001 to 10000	62 (53%)	54 (47%)	116
over 10000	100 (39%)	159 (61%)	259
Not stated	250 (93%)	19 (7%)	269

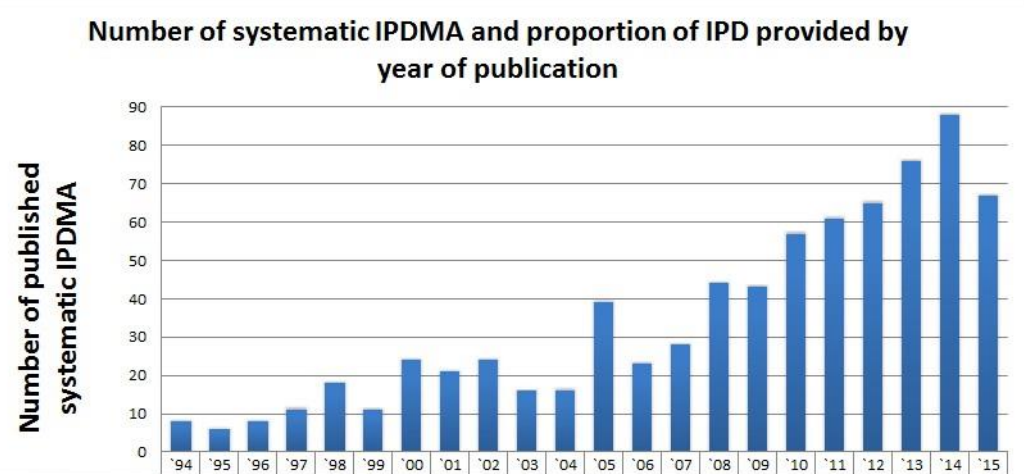
1. Other defined as lifestyle, nutrition, emergency medicine, patient care, patient preference, Pharmacokinetics and Forensics
2. Including 83 IPD-MAs where IPD was extracted from published study reports (IPD not requested from original study authors). See Appendix 8 for further details.
3. Non-commercial sources included institutional, government, charity, research council or research foundation funding.
4. Commercial sources were defined as pharmaceutical or manufacturer funding.

As 'opportunistic' IPD-MAs by design often differ in their objectives and inclusion criteria from systematic IPD-MAs [33]; no formal statistical comparison of the characteristics of these two types of IPD-MA was made. From visual comparison of Table 5, there did not seem to be any changes over time in approach (systematic or non-systematic) to IPD-MA and few clear differences by clinical area. A systematic approach was taken proportionally most often in the topics of Gynaecology, Pregnancy and Neonatology, Lung Cancer and Haematology, Leukaemia and Blood Cancer while a non-systematic approach was taken proportionally more often than a systematic approach in the topics of Cardiology, Mental and Psychiatric Disorders and other clinical areas (see Table 5). The majority of IPD-MAs of diagnostic test accuracy studies and studies of non-drug interventions took a systematic approach, as well as all Cochrane Reviews. The majority of non-systematic IPD-MAs had an authorship policy and the majority of IPD-MAs receiving commercial funding took a non-systematic approach. Table 5 also suggests that systematic IPD-MAs identify more eligible studies, while non-systematic IPD-MAs tend to identify and include more eligible participants.

4.3.3 IPD retrieval rate in systematic IPD-MA

Non-systematic IPD-MAs were mostly conducted with only the IPD which was already available to the analysts; therefore IPD retrieval rate is not relevant in these 520 IPD-MAs. The following two sections report only on the 760 systematic IPD-MAs.

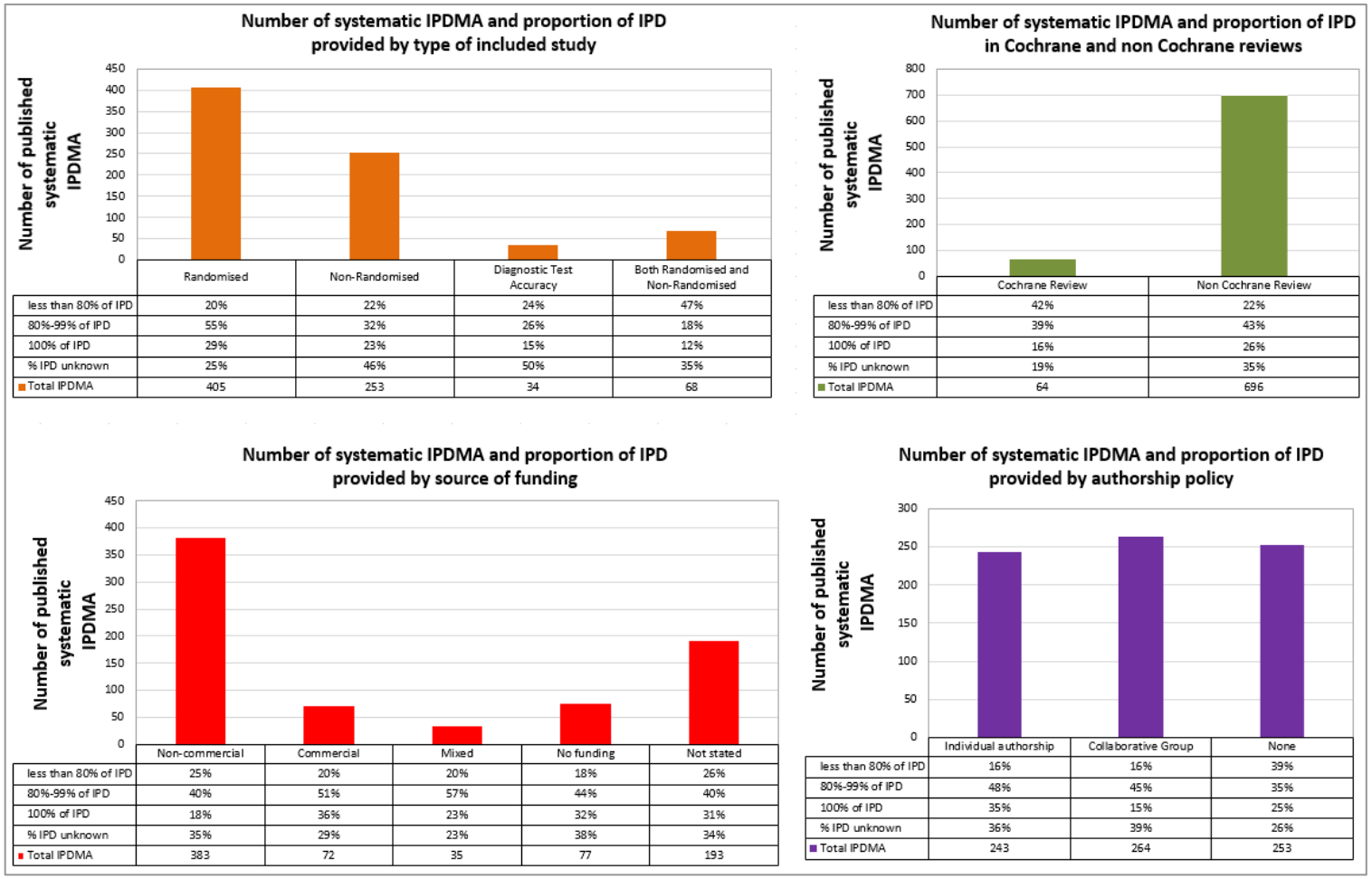
Figure 4: Number of distinct systematic IPD-MA published to August 2015 and proportion of IPD provided



	`94	`95	`96	`97	`98	`99	`00	`01	`02	`03	`04	`05	`06	`07	`08	`09	`10	`11	`12	`13	`14	`15
less than 80% of IPD	38%	17%	38%	0%	17%	27%	25%	29%	38%	63%	6%	21%	22%	11%	25%	19%	19%	20%	23%	25%	20%	33%
80-99% of IPD	25%	67%	25%	27%	33%	18%	21%	24%	13%	13%	19%	10%	17%	36%	14%	14%	18%	18%	15%	12%	17%	16%
100% of IPD	0%	0%	13%	18%	22%	27%	25%	33%	25%	13%	31%	41%	17%	21%	23%	26%	23%	28%	23%	29%	27%	22%
% IPD unknown	38%	17%	25%	55%	28%	27%	29%	14%	25%	13%	44%	28%	43%	32%	39%	42%	40%	34%	38%	34%	35%	28%
Total IPDMA	8	6	8	11	18	11	24	21	24	16	16	39	23	28	44	43	57	61	65	76	88	67

1. See Table 6 for proportion of systematic IPD meta-analyses providing 100%, 80-99%, less than 80% of IPD and the proportion of IPD not reported.
2. Six IPD-MA were published from 1987 to 1993; one was provided with less than 80% of IPD, three were provided with 80 – 99% of IPD and for two the proportion of IPD provided was not reported.

Figure 5: Characteristics of systematic IPD-MA and proportion of IPD provided.



1. See Table 6 for proportion of systematic IPD meta-analyses providing 100%, 80-99%, less than 80% of IPD and the proportion of IPD not reported.

Table 6: Characteristics of all systematic IPD-MAs according to proportion of IPD provided

IPD-MA Characteristic[1,2]	Total IPD-MAs (N)	Proportion of IPD retrieved for systematic IPD-MA (n and % of N)			
		100%	≥80%	<80%	Unknown[3]
Total	**760**	**188 (25%)**	**324 (43%)**	**179 (24%)**	**257 (34%)**
Clinical area of IPD-MA					
Breast Cancer	40	8 (20%)	22 (55%)	7 (17%)	11 (28%)
Cancer (other)	53	14 (26%)	27 (51%)	14 (26%)	12 (23%)
Cardiology	105	30 (29%)	53 (51%)	17 (16%)	35 (33%)
Central Nervous System, Neurology and Brain Injury	50	13 (26%)	20 (40%)	14 (28%)	16 (32%)
Cervical Cancer and Ovarian Cancer	16	1 (6%)	7 (44%)	1 (6%)	8 (50%)
Diabetes and Endocrinology	30	8 (27%)	13 (43%)	3 (10%)	14 (47%)
Gastroenterology, Colorectal and Gastric Cancer	49	11 (22%)	17 (35%)	23 (47%)	9 (18%)
Gynaecology, Pregnancy and Neonatology	35	13 (37%)	18 (51%)	9 (26%)	8 (23%)
Haematology, Leukaemia and Blood Cancer	43	11 (26%)	20 (47%)	4 (9%)	19 (44%)
Head and Neck Cancer	16	4 (25%)	8 (50%)	5 (31%)	3 (19%)
Hepatitis and Liver Disease	19	7 (37%)	8 (42%)	3 (16%)	8 (42%)
HIV	17	6 (35%)	8 (47%)	2 (12%)	7 (41%)
Infection and Infectious Diseases	31	6 (19%)	9 (29%)	12 (39%)	10 (32%)
Injuries and Wounds	21	2 (10%)	4 (19%)	13 (62%)	4 (19%)
Lung Cancer	32	9 (28%)	15 (47%)	3 (9%)	14 (44%)
Mental and Psychiatric Disorders	32	7 (22%)	12 (38%)	7 (21%)	13 (41%)
Musculoskeletal and Pain	34	9 (26%)	11 (32%)	5 (15%)	18 (53%)
Other	26	5 (19%)	9 (35%)	12 (46%)	5 (19%)
Otolaryngology, Ophthalmology and Periodontology	22	3 (14%)	5 (23%)	6 (27%)	11 (50%)
Renal and Urology	17	3 (18%)	6 (35%)	5 (30%)	6 (35%)
Respiratory and Pulmonary	21	7 (33%)	11 (52%)	3 (15%)	7 (33%)
Stroke, Thrombosis and Hypertension	51	12 (24%)	21 (41%)	11 (22%)	19 (37%)
Design of included studies					
Randomised	405	117 (29%)	222 (55%)	83 (20%)	100 (25%)
Non-Randomised	253	58 (23%)	81 (32%)	56 (22%)	116 (46%)
Diagnostic Test Accuracy	34	5 (15%)	9 (26%)	8 (24%)	17 (50%)
Both Randomised and Non-Randomised	68	8 (12%)	12 (18%)	32 (47%)	24 (35%)

Type of included studies					
Diagnostic Test Accuracy	34	5 (15%)	9 (26%)	8 (24%)	17 (50%)
Drug or device	348	102 (29%)	183 (53%)	73 (21%)	92 (26%)
Epidemiological	185	38 (21%)	58 (31%)	44 (24%)	83 (45%)
Non-drug (interventional)	193	43 (22%)	74 (38%)	54 (28%)	65 (34%)
Type of IPD-MA					
Cochrane Review	64	10 (16%)	25 (39%)	27 (42%)	12 (19%)
Non Cochrane Review	696	178 (26%)	299 (43%)	152 (22%)	245 (35%)
Authorship Policy					
Individual authorship	243	84 (35%)	116 (48%)	39 (16%)	88 (36%)
Collaborative Group	264	40 (15%)	119 (45%)	43 (16%)	102 (39%)
None	253	64 (25%)	89 (35%)	97 (39%)	67 (26%)
Source of Funding					
Non-commercial	383	70 (18%)	155 (40%)	94 (25%)	134 (35%)
Commercial	72	26 (36%)	37 (51%)	14 (20%)	21 (29%)
Mixed	35	8 (23%)	20 (57%)	7 (20%)	8 (23%)
No funding	77	25 (32%)	34 (44%)	14 (18%)	29 (38%)
Not stated	193	59 (31%)	78 (40%)	50 (26%)	65 (34%)
Number of eligible studies					
2 to 5	102	72 (71%)	83 (81%)	10 (10%)	9 (9%)
6 to 10	174	67 (39%)	98 (56%)	34 (20%)	42 (34%)
11 to 15	120	16 (13%)	47 (39%)	27 (23%)	46 (38%)
16 to 20	87	12 (14%)	29 (33%)	27 (31%)	31 (36%)
21 to 30	101	6 (6%)	30 (30%)	28 (28%)	43 (42%)
31 to 40	50	3 (6%)	11 (22%)	19 (38%)	20 (40%)
41 to 50	29	2 (7%)	5 (17%)	9 (31%)	15 (52%)
over 50	83	10 (12%)	19 (23%)	24 (29%)	40 (48%)
Not stated	14	0 (0%)	2 (14%)	1 (7%)	11 (79%)
Number of eligible participants					
under 100	18	14 (78%)	16 (94%)	1 (6%)	0 (0%)
101 to 200	20	13 (65%)	16 (80%)	4 (20%)	0 (0%)
201 to 500	45	21 (47%)	25 (56%)	19 (42%)	1 (2%)
501 to 1000	67	35 (52%)	45 (67%)	22 (33%)	0 (0%)
1001 to 5000	198	70 (35%)	134 (68%)	61 (31%)	3 (1%)
5001 to 10000	62	13 (21%)	37 (60%)	23 (37%)	2 (3%)
over 10000	100	22 (22%)	53 (53%)	46 (46%)	1 (1%)
Not stated	250	0 (0%)	0 (0%)	0 (0%)	250 (100%)

1. See Table 5 for full definitions of characteristics.
2. See Figure 4 for proportion of IPD provided in IPD-MAs by year.
3. Unknown as the number of eligible participants and/or the number of participants excluded from IPD analysis due to lack of IPD was not reported.

IPD was provided from 100% of eligible studies in only 189 (25%) and from 100% of participants in only 188 (25%) out of 760 systematic IPD-MAs; one IPD-MA provided with IPD from 100% of studies received an incomplete dataset for one study. IPD from at least 80% of studies was retrieved in 375 systematic IPD-MAs (49%) and from 80% of participants in 324 systematic IPD-MAs (43%). IPD was retrieved for less than 50% of studies in 136 systematic IPD-MAs (18%) and for less than 50% of participants in 71 systematic IPD-MAs (9%). One of the reviews was designed as an IPD-MA but no IPD was available [280].

For 257 IPD-MAs, the proportion of IPD retrieved could not be calculated because the number of eligible participants and/or the number of participants excluded from IPD analysis due to lack of IPD was not reported. Figure 4 shows the number of IPD-MAs published by year and the proportion of IPD retrieved.

Figure 5 and Table 6 and show the characteristics of the 760 systematic IPD-MA overall as well as separated according to IPD retrieval rate.

4.3.3.1 Characteristics associated with IPD retrieval

Table 7 presents the results of multivariable logistic regression (see Chapter 4.2.4 for further details). A total of 503 IPD-MAs were included in this analysis for which the proportion of IPD retrieved could be calculated (i.e. the number of participants eligible for analysis and the number of participants data was provided for was reported).

Table 7: Multivariable logistic regression model results: Characteristics associated with retrieving 100% of IPD or receiving more than 80% of IPD in 503 IPD-MAs

IPD-MA Characteristic	100% of IPD retrieved compared to less than 100% of IPD			At least 80% of IPD retrieved compared to less than 80% of IPD		
	OR	95% CI	P-value	OR	95% CI	P-value
Age of publication[1]	1.081	0.885 to 1.320	0.445	1.153	0.938 to 1.418	0.177
Number of eligible participants[1]	0.851	0.800 to 0.904	<0.001	0.889	0.837 to 0.943	<0.001
Includes randomised studies only	1.415	0.919 to 2.182	0.115	2.735	1.755 to 4.262	<0.001
Cochrane IPD-MA	0.402	0.189 to 0.859	0.019	0.427	0.218 to 0.835	0.013
Authorship Policy[2]	1.667	1.074 to 2.585	0.022	3.366	2.183 to 5.190	<0.001
Commercial source of funding[3]	1.291	0.762 to 2.187	0.341	1.043	0.568 to 1.914	0.892

1. Log transformation applied due to skewed distribution of data
2. Authorship policy (individual authorship for those providing IPD or collaborative group) compared to no authorship policy
3. Commercial source of funding (pharmaceutical or manufacturer) compared to non-commercial sources of funding only, no funding or no information regarding funding provided.

The odds of retrieving all IPD was significantly higher for IPD-MAs that were non-Cochrane, had a lower number of eligible participants and those with an authorship policy. The odds of retrieving a high proportion (at least 80%) of IPD were also significantly higher for these factors but in addition for IPD-MAs of randomised trials only. There was no association between the IPD retrieval rate and source of funding or the date of publication of IPD-MAs.

4.3.3.2 Additional and sensitivity analyses

A range of additional and sensitivity analyses were conducted to investigate assumptions made within the primary multivariable logistic regression analysis, see Chapter 4.2.5 for further details. Results of each additional analysis is summarised below and tables of numerical results are presented in Appendix 9.

1) Univariate logistic regression analysis (i.e. unadjusted analysis) (see Table 34)

Results of this analysis were numerically similar to those of multivariable (adjusted) analysis presented in Table 7; the only difference in conclusions was that unadjusted analysis shows no association between complete IPD retrieval rate and an authorship policy (association between authorship policy and high retrieval rate was maintained in adjusted and unadjusted analysis).

2) Association of authorship policy on IPD retrieval (see Table 35)

Results of this analysis show that odds of retrieving at least 80% of IPD were significantly increased when either individual authorship or collaborative authorship policies were used but odds of retrieving at 100% of IPD were significantly increased only when an individual authorship policy was used. Other numerical results were similar to those in Table 7 and results unchanged.

3) Inclusion of the variable 'Type of Study' in the multivariable model (see Table 36)

The variable 'Type of study' (drug or device (interventional), non-drug (interventional), diagnostic test accuracy or epidemiological study) was not included in the model due to correlation between this variable and type of study (interventional studies were significantly more likely to be randomised, chi-squared $p<0.001$) and source of funding (drug or device studies were significantly more likely to be commercially funded, chi- squared $p<0.001$).

Sensitivity analysis was conducted adding an additional variable to the multivariable logistic regression model. Results showed that this characteristic was not statistically significant; other numerical results were similar to those in Table 7 and conclusions were unchanged.

4) *Exclusion of IPD-MA from analysis with no information regarding funding reported (see Table 37)*

Numerical results were similar to those in Table 7 and conclusions were unchanged.

5) *Assuming the following scenarios for 257 IPD-MA for which the proportion of IPD retrieved could not be calculated:*
    a. Less than 80% of IPD was retrieved (see Table 38)
    b. 80% or more IPD was retrieved (see Table 39)
    c. 100% of IPD was retrieved (see Table 40)

Results of these sensitivity analyses are varied compared to those reported in Table 7; for example scenario a. and scenario b. contradict Table 7 and suggest that the odds of complete or high IPD retrieval rate are significantly higher in IPD-MA without an authorship policy. These sensitivity analyses highlight the importance of reporting the proportion of IPD retrieved in IPD-MA.

An additional analysis was also performed to examine characteristics of the 257 IPD-MA where proportion of IPD retrieved could not be calculated compared to the 503 IPD-MA were proportion of IPD-MA could be calculated (see Table 41).

Results of this additional analysis indicate that the odds of the proportion of IPD retrieved being reported are significantly higher in more recently published IPD-MA, IPD-MA including RCTs only and IPD-MA without an authorship policy. There was no association between publication as a Cochrane IPD-MA and the source of funding on whether the proportion of IPD retrieved was reported.

6) *Use of fractional logistic regression as an alternative to logistic regression (see Table 42)*

Results of this analysis indicate that odds of retrieving a higher proportion of IPD are significantly associated with older IPD-MAs, IPD-MAs including only randomised studies, non-Cochrane IPD-MAs and IPD-MAs with an authorship policy. There was no association between the number of eligible participants and source of funding on the proportion of IPD retrieved.

7) *Multivariable logistic regression of the proportion of study data retrieved (see Table 43)*

A total of 744 IPD-MAs were included in this analysis for which we could calculate the proportion of study data retrieved (i.e. the number of studies eligible for analysis and the number of studies data was provided for was reported).

Results of this sensitivity analysis are mostly similar to those reported in Table 7, however these results suggest that the odds of retrieving at least 80% of study data are significantly associated with older IPD-MA; suggesting that IPD retrieval rate on a study-level has got worse over time.

In summary, numerical results of additional and sensitivity analyses are mostly similar and conclusions mainly unchanged indicating that the results of primary multivariable regression analyses are robust to assumptions made. Some results do, however, indicate that the proportion of IPD retrieved over time (including on a study-level) has got worse over time.

The most variability in results was shown in the sensitivity analyses exploring a range of scenarios for the 257 IPD-MAs which did not report the proportion of IPD retrieved. Results of these sensitivity analyses varied and some were contradictory to the primary analysis, for example indicating that the odds of complete or high IPD retrieval rate are significantly higher in IPD-MAs without an authorship policy. These sensitivity analyses highlight the importance of a reporting the proportion of IPD retrieved in IPD-MAs.

Interestingly, when further considering the association of an authorship policy with IPD retrieval rate, results of this additional analysis show that odds of retrieving at least 80% of IPD were significantly increased when either individual authorship or collaborative authorship policies were used but odds of retrieving at 100% of IPD were significantly increased only when an individual authorship policy was used.

4.3.4 Unavailability of IPD and the impact on analysis

Out of the 571 systematic IPD-MAs that failed to retrieve 100% of the IPD, 201 (34%) had supplemented IPD with AD extracted from study publications. The additional AD had been included from a median of 5 (range 1 to 541) studies and a median of 683 (range 9 to 1,180,505) participants.

At least one study had been excluded from the meta-analysis due to lack of IPD or AD in 419 (55%) systematic IPD-MAs. Across these, a median of 4 (range 1 to 342) studies and a median of 478 (range 8 to 1,792,339) participants were excluded from IPD-MAs but 241 systematic IPD-MA (32%) failed to state how many participants were excluded from analysis.

Up to six reasons were reported for unavailability of IPD (Table 8); non-specific reasons, such as ‘data was not available for analysis’ were reported in 341 out of 571 systematic IPD-MAs

(58%). The most common specific reasons for not obtaining IPD were that investigators could not be contacted, investigators had declined to share data or that data had been lost or destroyed. In 24 systematic IPD-MAs it was reported that data was not requested for all studies; mainly due to the size or quality of these studies.

Table 8: Reasons reported for unavailability of IPD in 571 systematic IPD-MA without 100% of IPD retrieved

Reasons reported for not retrieving 100% of eligible IPD	Number of IPD-MA[1,2]
Data not available[3]	341 (60%)
No contact could be made with study authors	104 (18%)
Investigators declined but no reason given	74 (13%)
Data lost or destroyed	65 (11%)
Data could not be extracted[4]	55 (10%)
Trial was still ongoing	42 (7%)
Data quality issues	29 (5%)
Failed to provide data in time for the IPD-MA	26 (5%)
Data not requested	24 (4%)
Ethical / ownership restrictions	15 (3%)
Reason unclear	11 (2%)

1. 189 IPD-MA with 100% of IPD provided not included in the table.
2. Up to 6 reasons reported for unavailability of IPD. Therefore total number of reasons (and total percentages) sum to greater than 571 (100%)
3. IPD was not available for a proportion of studies without any specific reason quoted.
4. Applicable only in a small number of IPD-MAs where IPD were extracted from publications rather than requested.

Table 9: Approach to accounting for missing IPD in 571 systematic IPD-MA without 100% of IPD retrieved

Approach reported to account for missing IPD	Number of IPD-MA[1,2]
None stated	143 (25%)
Separate meta-analyses are conducted including IPD only and IPD plus available AD	81 (14%)
Stated that missing IPD is a limitation of the meta-analysis and / or that availability bias may be present	76 (13%)
AD included in primary analysis	61 (11%)
Sensitivity analysis with AD performed	57 (10%)
Stated that the missing IPD is unlikely to change results	56 (10%)
Results from the studies without IPD summarised narratively	48 (8%)
Stated that the majority of data is included in analysis	47 (8%)
Narrative comparison to an AD meta-analysis	18 (3%)
Intend to include data in an update	14 (2%)

1. 189 IPD-MA with 100% of IPD provided not included in the table.
2. Up to three approaches described to account for missing IPD. Therefore total number of approaches (and total percentages) sum to greater than 571 (100%)

In 143 (25%) out of the 571 systematic IPD-MAs there was no acknowledgement of potential bias resulting from missing IPD. In 199 (34%) of the systematic IPD-MAs additional analyses using AD had been performed and in a further 66 (11%) systematic IPD-MAs, a narrative description of the studies without IPD or a narrative comparison to an aggregate data meta-analysis had been provided. The remaining 183 (31%) systematic IPD-MAs make reference to the missing data, some acknowledging this may result in bias, without any further investigation of the implication on the conclusions of the review (Table 9).

4.4 Discussion

4.4.1 Summary of main results

At the time of writing, this systematic review is believed to include the largest cohort of published IPD-MAs to date. Recent years have shown an increase in development of statistical methodology for the synthesis of IPD [36] (see also Chapter 2.3.3 of this thesis) as well as a rapid increase in the uptake of methods, with the number of systematic and non-systematic IPD-MAs published per year increasing to an average of 105 published per year between 2009 and 2015 compared to 49 per year published between 2005 and 2009 [34]. However, these rapid increases do not seem to be mirrored by improved IPD retrieval rates, which may be due, in part, to the increasing uptake of IPD-MAs across a wide range of clinical areas and settings where IPD may be difficult to obtain.

The findings of this systematic review showed that Cochrane reviews were less likely to retrieve all or a high proportion of IPD than systematic non-Cochrane reviews. This may be explained by the inclusion of thorough search methods within Cochrane reviews, as well as advances in systematic searching of larger electronic databases generally, leading to the identification of larger numbers of studies including more grey literature studies where IPD may be difficult to retrieve with the resources available to review authors, such as Cochrane review authors who usually undertake systematic reviews on a voluntary basis. Furthermore, the framework of a Cochrane review requires the registration of a protocol and publication of results regardless of the IPD retrieval rate; therefore, Cochrane IPD reviews with a low IPD retrieval rate may be less subject to review-level 'publication bias' than non-Cochrane IPD reviews with a low IPD retrieval rate.

On the other hand, results also showed that IPD-MAs with an authorship policy (individual authorship or collaborative group authorship) were associated with retrieving a high proportion of IPD but it was only the IPD-MAs offering individual authorship which were

associated with 100% retrieval of IPD (see Table 35). This is an important finding as the implementation of an authorship policy as an incentive to participate in an IPD-MAs, as a feature of a well-designed project, is a factor which is in control of the IPD-MAs team; even where other characteristics such as study design and number of eligible participants for IPD-MA are constrained by the research question.

4.4.2 Strengths and weaknesses

The aim of this systematic review was to systematically identify all published IPD-MAs regardless of use of a systematic design to identify studies, resulting in a large cohort of nearly 1300 IPD-MAs. Inclusion criteria were wide and reasons for exclusion were documented for all references identified in electronic searches. Ninety abstracts which could not be matched to full-text articles, despite best efforts, were excluded from the systematic review. Due to the size of the cohort of this study, double reference screening and data extraction was performed on only a subset of the articles. Agreement between double extractions was good and all discrepancies were minor and easily resolved, therefore any errors made in screening and extraction would have been minimal and unlikely to influence the overall findings of the study.

It was not possible to systematically investigate the IPD retrieval methods employed within the IPD-MAs; such as the number of attempts to contact investigators to request data etc., due to the lack of published detail regarding such processes. Data collection methods are likely to be an important factor influencing the proportion of IPD retrieved and clearer reporting of approaches to IPD collection, would be valuable to those planning new IPD-MAs.

The primary analysis approach taken in this study involved dichotomising the dependent variable (proportion of IPD retrieved) and performing multivariable logistic regression analysis. The limitations of dichotomisation should be noted and further approaches to modelling IPD retrieval rate which takes account of the bimodal distribution of the dependent variable could be considered as future research. However, we believe that any loss of information will be reduced by the size of the cohort included in analysis and a range of sensitivity analyses have been presented to investigate all assumptions made in the primary analysis, demonstrating overall consistency and robustness of results.

This systematic review examines associations between IPD retrieval rate and characteristics of the IPD-MA. Arguably, it would have been more informative to consider the association between IPD retrieved (yes or no) and characteristics of the individual studies within the IPD-MA, particularly when considering whether there has been any changes over time in IPD

retrieval rates. However, it was not possible to systematically examine the characteristics of the studies providing or not providing IPD within the 760 IPD-MAs due to lack of specific information reported at the study-level. For example, 125 out of 760 IPD-MAs (16%) did not provide any information at all regarding the years of publication of studies not providing IPD, and many of the IPD-MAs provided only year ranges of eligible studies and/or studies providing IPD. Therefore, modelling the probability of a study providing IPD was not deemed appropriate. This is a limitation of this analysis but it should be noted for future IPD-MAs that it is essential to clearly describe the characteristics of the eligible studies which do and also eligible studies which do not contribute to the IPD analysis. Reporting of such information on a study-level allows a judgement of 'availability bias' within the IPD-MA; in other words, whether the provision of IPD may be associated with characteristics of the eligible studies.

4.4.3 Relation to other studies and implications

Present results have shown that a quarter of systematic IPD-MAs published since 1987 retrieved all IPD for analysis and only half retrieved at least 80% of relevant IPD. This latter finding is higher than previous results which reported that around 25% of IPD-MAs had included less than 80% of IPD [32, 33, 35, 40]. However previous work has been based on smaller cohorts of IPD-MAs, has mostly focused on IPD-MAs of RCTs only and has been conducted over smaller time frames.

In line with previous work [32, 33, 35, 40], present results show that important inadequacies around conduct and reporting of IPD-MAs remain. Non-systematic methods, mostly based on the known availability of IPD, had been used to select eligible studies for inclusion in 41% of the initial cohort of IPD-MAs identified. It was outside the scope of this study to further examine the design of these analyses; however, it is recommended that non-systematic pooling of IPD is conducted in the framework of a prospective meta-analysis [281] and that the conclusions of such analyses must take the inevitable selection bias into account.

Furthermore in around 5% of the systematic IPD-MAs, IPD was not requested from a subset of the eligible studies (Table 8), often due to the small size or study quality in relation to the other eligible studies. It is arguably acceptable to exclude studies of poor quality which may impact on the overall IPD-MA [282], however such exclusions should be specified *a priori* and investigated via sensitivity analysis to avoid introducing selection bias to reviews of a systematic nature [51].

Present results highlight the importance of clear reporting of study and participant numbers contributing to different stages of the IPD-MA with an adequate investigation of the reasons for lack of data and discussion of the potential for 'availability bias.' The total number of eligible participants and the total number of participants' data requested was unclear in 34% of published IPD-MAs; in 58% of the IPD-MAs that failed to retrieve 100% of eligible IPD, there were no specific reasons provided for the unavailability of data, making interpretation of IPD-MAs results and conclusions in the presence of potential 'availability bias' difficult. In a quarter of IPD-MAs unable to retrieve 100% of IPD, there was a complete lack of discussion or acknowledgement of 'availability bias'. A systematic investigation of the impact of 'availability bias' on IPD-MAs conclusions was outside the scope of this review and is specific to the clinical context in question. Despite this, further efforts are recommended by researchers conducting an IPD-MA to thoroughly investigate and report the impact of data availability [51].

Proper uptake of new PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) IPD guidelines for the conduct and reporting of IPD-MA [283], in addition to guidance on the use of IPD-MA to synthesise the results of RCTs [282], should lead to improved conduct and reporting in IPD-MAs. In particular, transparent reporting of the number of eligible studies and participants, how much data was requested and obtained with clear reasons for non-availability of IPD, preferably via a flow diagram, and data collection methods. Discussion of limitations and impact on conclusions due to missing IPD is essential.

4.4.4 Concluding remarks

IPD-MAs are resource demanding, time consuming and methodologically challenging but when conducted well [282], ideally following a registered protocol [284] and adhering to the PRISMA-IPD guidance [283], can provide more detailed and potentially more reliable results than a meta-analysis of aggregate data. Meta-analysts must carefully consider the appropriateness of an IPD analysis and demonstrate awareness of potential biases induced by missing IPD. Only one in four published systematic IPD-MAs have had access to all IPD; we hope that this proportion will grow in future years with the growing awareness of data sharing and transparency in the pharmaceutical industry and beyond [77-80, 84, 85].

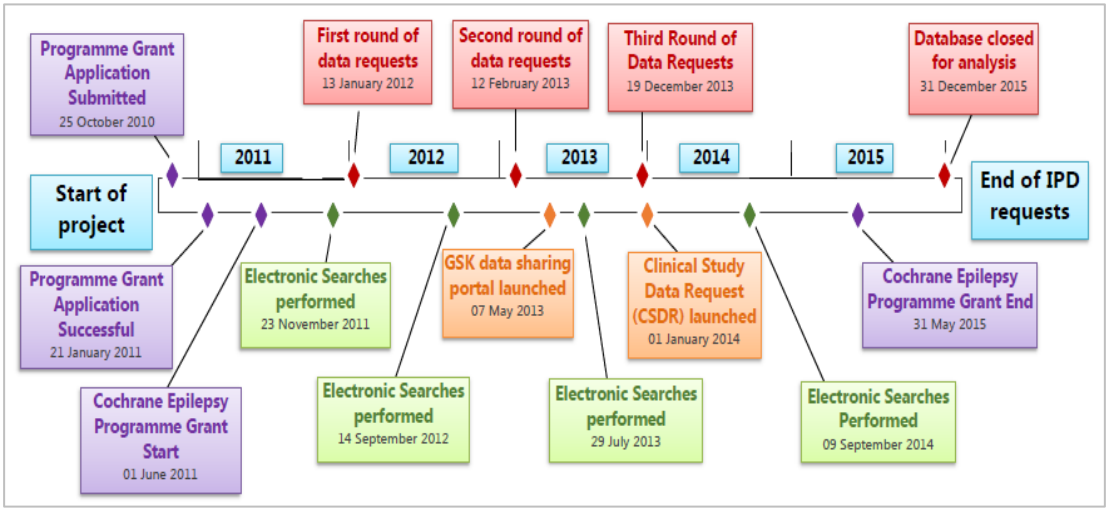
Chapter 5: Individual participant data requests: 20 years' experience of the Cochrane Epilepsy Group

5.1 Introduction

The Cochrane Epilepsy Group has been making IPD requests to the authors of AED monotherapy trials since the mid-1990s with eight reviews for IPD-MA of pair-wise AED comparisons published to date since 2000 [59-67]. The group have also previously published an IPD-NMA including participants randomised to one of eight AEDs [285]. See Chapter 1.2 for further discussion of the clinical setting and the rationale of these IPD reviews.

Since the publication of the original NMA in 2007, additional AEDs have been used in clinical practice and additional clinical trials have been conducted which has prompted the need for an updated analysis. Plans to conduct a Cochrane Review and IPD-NMA of 10 AEDs began in 2010 (see Figure 6), with the submission for project funding via a Cochrane programme grant in October 2010 and beginning in June 2011.

Figure 6: Timeline of IPD requests for Cochrane Epilepsy IPD-NMA



This chapter outlines the data requesting process for the current IPD-NMA, further details of which can be found in the Cochrane review [69]. This chapter also reflects upon the experiences of the Cochrane Epilepsy Group in requesting IPD prior to the current IPD-NMA and changes over time following 20 years of data requesting.

The work contained in this Chapter relating to IPD requesting and IPD retrieval for Cochrane Epilepsy reviews has been published in the British Medical Journal [278] and the IPD-NMA has been published on the Cochrane Database of Systematic Reviews [69].

5.2 Selection of studies for IPD-NMA

5.2.1. Inclusion criteria

5.2.1.1 Study Design

Randomised controlled trials (RCTs) of a parallel design in which the unit of analysis is the individual (i.e. cluster randomised trials were excluded). RCTs may be blinded (double-blind, single-blind etc.) or open label and may use either an adequate method of allocation concealment (e.g. sealed opaque envelopes) or a quasi-method of randomisation (e.g. allocation by date of birth).

Trials of a monotherapy design only were included; i.e. participants are randomised to treatment with a single drug throughout the trial period. Trials with an add-on, poly-therapy, transitional or withdrawal to monotherapy periods of any length were excluded.

Trials of a cross-over design were excluded as such as design is inappropriate for measuring primary outcome 'time-to-withdrawal of allocated treatment' as withdrawal during the first treatment period would prevent cross-over into the second period, resulting in incomplete outcome data. Furthermore, the use of cross-over designs is no longer recommended in epilepsy trials of a monotherapy design [286].

5.2.1.2 Participants

Children or adults with partial-onset seizures (simple partial, complex partial, or secondarily generalised tonic-clonic seizures) or generalised-onset tonic-clonic seizures (with or without other generalised seizure types) with a new diagnosis of epileptic seizures or who had had a relapse of seizures following antiepileptic monotherapy withdrawal.

Trials recruiting only participants with other generalised seizure types alone (e.g. participants recruited with absence seizures alone without generalised tonic clonic seizures) as guidelines for the first-line treatment of other generalised seizure types are different from the guidelines for generalised tonic-clonic seizures [74] and due to documented evidence that certain drugs of interest may exacerbate some generalised seizure types [287, 288].

Trials considering AEDs as treatment for conditions other than epilepsy were excluded.

5.2.1.3 Interventions

Ten AEDs currently licensed and commonly used as monotherapy in at least one country were included in the treatment network [72, 73]:

- carbamazepine (CBZ)
- phenobarbitone (PHB)
- phenytoin (PHT)
- sodium valproate (VPS)
- lamotrigine (LTG)
- oxcarbazepine (OXC)
- topiramate (TPM)
- gabapentin (GBP)
- levetiracetam (LEV)
- zonisamide (ZNS)

Clinical profiles and mechanisms of action of these ten drugs are detailed in the Cochrane Epilepsy IPD-NMA [68, 69].

Included trials must make at least one pairwise comparison between at least 2 of the 10 antiepileptic drugs included in the network. For trials with three treatment arms or more, only treatment arms of the ten AEDs listed above are included and any treatment arms not included in the network were excluded from analysis. Trials with multiple arms (doses) of the same drug were included as long as at least one arm of another drug from our network was included (e.g. multiple doses of GBP compared to CBZ) [289]. Multiple dose arms of the same drug are pooled in analysis; dose comparisons were outside the scope of this analysis.

5.2.2 Study selection

5.2.2.1 Systematic Search methods

The following databases were searched with no language restrictions: the Cochrane Epilepsy Group Specialised Register, the Cochrane Central Register of Controlled Trials (CENTRAL) via the Cochrane Register of Studies Online (CRSO), MEDLINE, SCOPUS, ClinicalTrials.gov and World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP). Search strategies are published in the Cochrane Epilepsy IPD-NMA [69].

Systematic searches were performed on the following dates:

- 23rd November 2011
- 14th September 2012
- 29th July 2013
- 9th September 2014
- 27th July 2016

Hand-searching was performed of relevant conference proceedings and reference lists of retrieved trials. Experts in the field were also contacted for details of any ongoing or unpublished trials.

5.2.2.2 Screening of studies

One reviewer (SJN) screened all titles and abstracts of all records identified by the electronic searches according to the pre-specified inclusion criteria (see Chapter 5.2.1). Subsequently, two reviewers (SJN and AGM) independently assessed full-text publications according to the same inclusion criteria. Disagreements were resolved by discussion or by consulting a third reviewer (CTS).

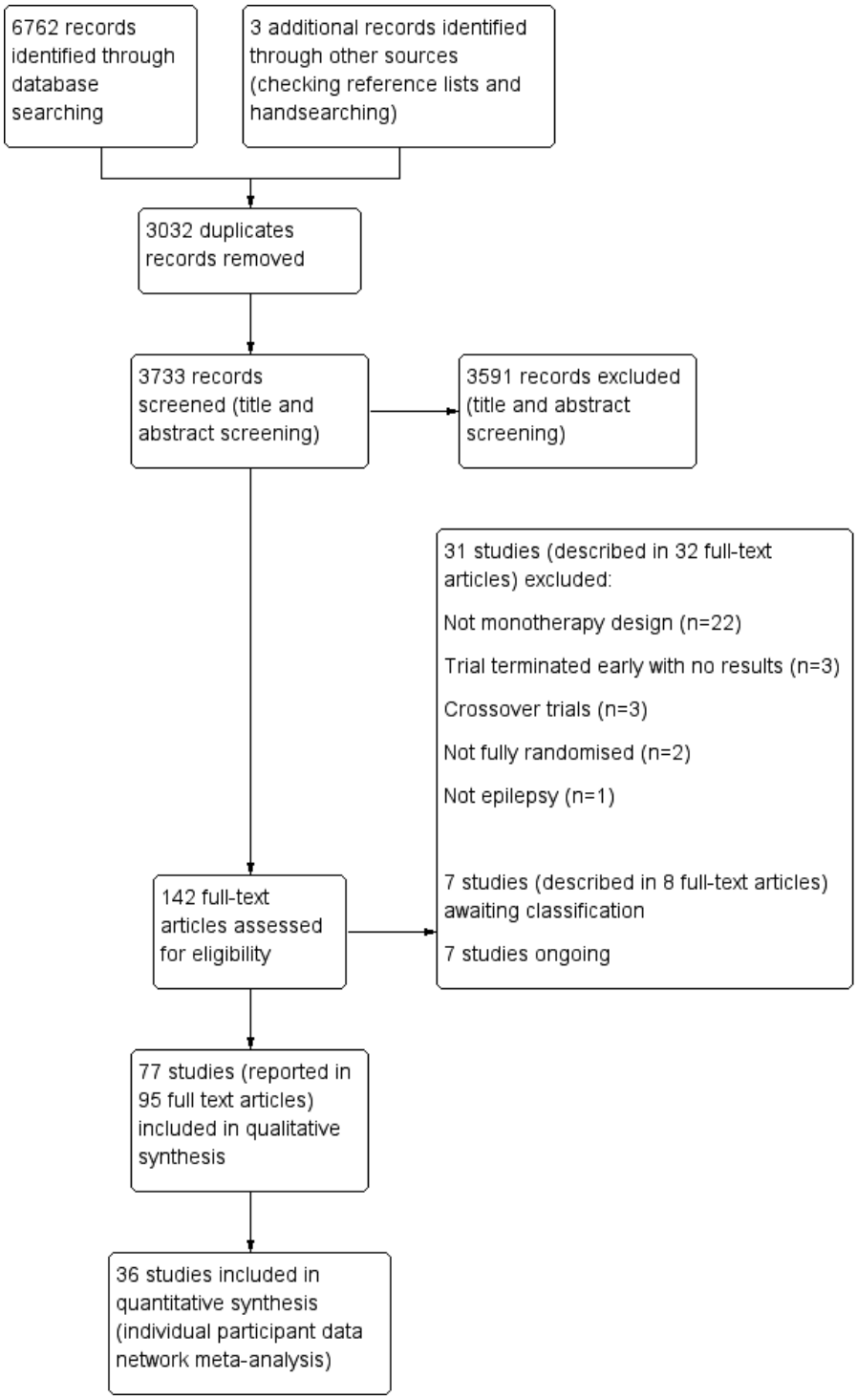
5.2.2.3 Results of the systematic search

Electronic searching identified 6762 records and a further three records were found by hand-searching and checking reference lists of included trials. Following removal of 3032 duplicate records, 3733 records were screened (title and abstract) and 3591 clearly irrelevant records were excluded. Full-text articles were accessed and screened for the remaining 142 records and 31 trials (described in 32 full-text articles) were excluded. In addition to the excluded trials, seven records were identified as ongoing trials and eight records were classified as 'awaiting assessment' as translation into English or further information was required to assess eligibility of the trials.

Figure 7 shows a study flow diagram of the screening process, including reasons for exclusion of full-text articles. In total, 77 trials (described in 95 full-text-articles) were included in the Cochrane review (see Appendix 10 for references of the primary publication of each trial).

One full-text article reported on a cohort of participants which were recruited into two separate trials [290]; these two trials are treated separately in analysis as 'Brodie 1995a' and 'Brodie 1995b.' It was unknown at the time of data request that the cohort of participants reported on were recruited in two separate trials, therefore Brodie 1995 is treated as a single data request in subsequent sections of this Chapter.

Figure 7: Study Flow Diagram: Selection of studies for Cochrane Epilepsy Individual Participant Data Network Meta-Analysis



5.2.3.4 Characteristics of included trials

Table 10 summarises the characteristics of the 77 trials eligible for inclusion in the NMA.

Table 10: Characteristics of trials included in Cochrane review and IPD-NMA

	Characteristics of included trials[6]						
Trial reference[1]	**Trial AEDs**	**Seizure type(s)[7]**	**Previous AED use permitted?**	**Ages**	**Single or Multi Centre**	**Type of study (sponsorship)[8]**	**IPD available**
Aikia 1992	PHT; OXC	Both	No	Adults	Not stated	Academic	No
Banu 2007	CBZ; PHB	Both	No	Children	Single	Academic	Yes
Baulac 2012	CBZ; ZNS	Partial only	No	Adults	Multi	Pharmaceutical	Yes
Bidabadi 2009[2]	CBZ; PHB	Partial only	Not stated	Children	Single	Academic	No
Bill 1997	PHT; OXC	Both	No	Adults	Multi	Pharmaceutical	Yes
Biton 2001	LTG; VPS	Both	Not stated	All ages	Multi	Pharmaceutical	Yes
Brodie 1995a[3]	CBZ; LTG	Both	No	All ages	Multi	Pharmaceutical	Yes
Brodie 1995b[3]	CBZ; LTG	Both	No	All ages	Multi	Pharmaceutical	Yes
Brodie 1999	CBZ; LTG	Both	No	Elderly	Multi	Pharmaceutical	Yes
Brodie 2002	GBP; LTG	Both	Not stated	Adults	Multi	Pharmaceutical	No
Brodie 2007	CBZ; LEV	Both	No	Adults	Multi	Pharmaceutical	Yes
Callaghan 1985	CBZ; PHT; VPS	Both	No	All ages	Single	Academic	No
Capone 2008[4]	CBZ; LEV	Not stated	New onset post-stroke seizures	Adults	Single	Academic	No
Castriota 2008[4]	CBZ; LEV	Partial only	No	Adults	Single	Academic	No
Chadwick 1998	CBZ; GBP	Partial only	Relapsed seizures permitted	All ages	Multi	Pharmaceutical	Yes
Chen 1996	CBZ; PHB; VPS	Both	No	Children	Single	Academic	No
Cho 2011	CBZ; LEV	Partial only	No	All ages	Single	Academic	No
Christe 1997	OXC; VPS	Both	No	Adults	Multi	Pharmaceutical	No
Consoli 2012	CBZ; LEV	Both	New onset post-stroke seizures	Adults	Multi	Academic	No

Cossu 1984[4]	CBZ; PHB	Partial only	No	Adults	Single	Academic	No
Craig 1994	PHT; VPS	Both	No	Elderly	Single	Pharmaceutical	Yes
Czapinski 1997[2]	CBZ; PHB; PHT; VPS	Partial only	No	Adults	Not stated	Academic	No
Dam 1989	CBZ; OXC	Not stated	No	Adults	Multi	Academic	No
de Silva 1996	CBZ; PHB; PHT; VPS	Both	No	Children	Multi	Academic	Yes
Dizdarer 2000	CBZ; OXC	Partial only	Not stated	Children	Single	Academic	Yes
Donati 2007	CBZ; OXC; VPS	Partial only	No	Children	Multi	Pharmaceutical	No
Eun 2012	CBZ; LTG	Partial only	No	Children	Multi	Academic	Yes
Feksi 1991	CBZ; PHB	Both	No	All ages	Single	Pharmaceutical	No
Forsythe 1991	CBZ; PHT; VPS	Not stated	No	Children	Single	Academic	No
Fritz 2006[2]	LTG; OXC	Not stated	Not stated	Adults	Not stated	Academic	No
Gilad 2007	CBZ; LTG	Partial only	New onset post-stroke seizures	Adults	Single	Academic	No
Guerreiro 1997	PHT; OXC	Both	No	Children	Multi	Pharmaceutical	Yes
Heller 1995	CBZ; PHB; PHT; VPS	Both	No	Adults	Multi	Academic	Yes
Jung 2015[9]	CBZ; LEV	Partial only	No	Children	Multi	Academic	No
Kalviainen 2002[2]	CBZ; LTG	Both	No	Not stated	Multi	Pharmaceutical	No
Kopp 2007[2]	CBZ; LEV; VPS	Both	No	Not stated	Single	Academic	No
Korean Lamotrigine Study Group 2008[2,9]	CBZ; LTG	Both	No	All ages	Multi	Pharmaceutical	No
Kwan 2009	LTG; VPS	Both	Relapsed seizures allowed	Adults	Multi	Academic	Yes
Lee 2011	CBZ; LTG	Partial only	No	Adults	Multi	Academic	Yes
Lukic 2005[2]	LTG; VPS	Both	No	Adults	Single	Academic	No
Mattson 1985	CBZ; PHT; PHB	Partial only	Under-treated seizures allowed	Adults	Multi	Government	Yes
Mattson 1992	CBZ; VPS	Partial only	Under-treated seizures allowed	Adults	Multi	Government	Yes

Mitchell 1987	CBZ; PHB	Partial only	No	Children	Single	Academic	No
Miura 1990	CBZ; PHT; VPS	Both	No	Not stated	Single	Academic	No
Motamedi 2013[9]	LEV; LTG	Both	No	Elderly	Single	Academic	No
NCT01498822[2,9]	OXC; LEV	Partial only	No	Adults	Multi	Pharmaceutical	No
NCT01954121[2,9]	CBZ; LEV	Partial only	No	Adults	Multi	Pharmaceutical	No
Nieto-Barrera 2001	CBZ; LTG	Partial only	No	All ages	Multi	Pharmaceutical	Yes
Ogunrin 2005	CBZ; PHT; PHB	Both	No	Adults	Single	Academic	Yes
Pal 1998	PHB; PHT	Both	No	Children	Single	Academic	Yes
Placencia 1993	CBZ; PHB	Both	No	All ages	Single	Academic	Yes
Privitera 2003[5]	CBZ; TPM; VPA	Both	No	All ages	Multi	Pharmaceutical	Yes
Pulliainen 1994	CBZ; PHT	Both	No	Adults	Single	Academic	No
Ramsey 1983	CBZ; PHT	Both	No	Adults	Multi	Government	No
Ramsey 1992	PHT; VPA	GTC only	No	All ages	Multi	Government	Yes
Ramsey 2007[2]	CBZ; LEV	Partial only	Under-treated seizures allowed	Elderly	Multi	Academic	No
Ramsey 2010	PHT; TPM	Both	Under-treated seizures allowed	All ages	Multi	Pharmaceutical	Yes
Rasgoti 1991	PHT; VPS	Both	Not stated	All ages	Single	Academic	No
Ravi Sudhir 1995	CBZ; PHT	Both	No	Adults	Single	Academic	No
Resendiz 2004[4]	CBZ; TPM	Partial only	No	Children	Multi	Academic	No
Reunanen 1996	CBZ; LTG	Both	Relapsed seizures allowed	All ages	Multi	Pharmaceutical	Yes
Richens 1994	CBZ; VPS	Both	Relapsed seizures allowed	Adults	Multi	Pharmaceutical	Yes
Rowan 2005	CBZ; GBP; LTG	Both	Under-treated seizures allowed	Elderly	Multi	Government	No
Saetre 2007	CBZ; LTG	Not stated	No	Elderly	Multi	Pharmaceutical	No
SANAD A 2007	CBZ; GBP; LTG; OXC; TPM	Partial only	Relapsed seizures allowed	All ages	Multi	Academic	Yes

SANAD B 2007	LTG; TPM; VPS	GTC only	Relapsed seizures allowed	All ages	Multi	Academic	Yes
Shakir 1981	PHT; VPS	Both	Relapsed seizures allowed	All ages	Multi	Academic	No
So 1992	CBZ; VPS	Partial only	Under-treated seizures allowed	Adults	Not stated	Academic	No
Steiner 1999	LTG; PHT	Both	No	Adults	Multi	Pharmaceutical	Yes
Steinhoff 2005[5]	CBZ; LTG; VPS	Both	No	All ages	Multi	Pharmaceutical	No
Stephen 2007	LTG; VPS	Both	No	All ages	Single	Academic	Yes
Suresh 2015[9]	CBZ; LEV	Partial only	No	Adults	Single	Academic	No
Thilothammal 1996	PHB; PHT; VPS	GTC only	No	Children	Single	Academic	No
Trinka 2013[5]	CBZ; LEV; VPS	Both	No	Adults	Multi	Pharmaceutical	Yes
Turnbull 1985	PHT; VPS	Both	No	Adults	Single	Academic	Yes
Verity 1995	CBZ; VPS	Both	Relapsed seizures permitted	Children	Multi	Pharmaceutical	Yes
Werhahn 2015	CBZ; LEV; LTG	Partial only	No	Elderly	Multi	Pharmaceutical	Yes

1. See Appendix 10 for reference of the primary publication of each trial and Chapter 5.2.1.3 for abbreviations of drugs.
2. Available only as abstract, online summary or clinical trial summary report.
3. Two trials reported in a single publication.
4. Translated from Italian or Spanish
5. Trials designed in two strata based on whether recommended treatment would be CBZ or VPS.
6. Further details of characteristics (e.g. proportions of each seizure type, specific age ranges recruited, geographical locations of centres etc.) are available in the published Cochrane IPD-NMA [69].
7. GTC: Generalised tonic clonic seizures with or without other generalised types. 'Both' – indicates individuals with partial seizures and individuals with GTCs recruited. 'Not stated' indicates that the proportion of each seizure type recruited was not stated.
8. Academic defined as study conducted within a university or hospital setting without clear government or pharmaceutical sponsorship or involvement.
9. Trial identified in an updated search in 2016, following closure of database for analysis. IPD request initiated and any IPD provided will be included in an update of the Cochrane IPD-NMA.

Most of the trials (63 out of 77 trials (82%)) had been published in at least one full-text article in English; seven trials were available in abstract form only, two trials were available only as an online summary on ClinicalTrials.gov and one trial was available in English only as a clinical trial summary report (full-text article published in Korean). Three trials published as a full-text article in Italian and one trial published as a full-text article in Spanish were translated.

One published full-text article (Brodie 1995) reported on two separate trials. Furthermore, three trials were designed in strata based on whether clinician recommended treatment would be CBZ or VPS. Within the two strata, participants were randomised to an experimental AED (TPM, LTG or LEV) compared to the clinician recommended treatment. To ensure that randomised comparisons are made, the strata in these trials were considered separately in this review (i.e. IPD would be analysed in a CBZ branch and a VPS branch).

Forty-five trials (58%) were multicentre, 28 trials (36%) were single centre and the number of centres was not stated for four trials (6%). Trials were conducted globally across North America, Europe, South America, Africa and Australasia. Twenty-nine studies (38%) were pharmaceutical sponsored (defined as pharmaceutical studies herein), five studies were government sponsored (6%, defined as government studies herein) and the remaining 43 studies (56%) conducted within a university or hospital setting without clear pharmaceutical or government sponsorship or involvement (defined as academic studies herein).

All trials recruited individuals of both genders. Twenty trials recruited individuals of all ages (26%), fifteen trials (19%) recruited children only; with the age limits ranging from under 12 years to under 18 years of age; 32 trials (41%) recruited adults only, with the age limits ranging from over 13 years to over 18 years of age; seven trials (9%) recruited elderly participants only, with age limits ranging from over 60 years to over 65 years of age; and three trials (5%) did not state age ranges of eligible participants.

Twenty-five trials (32%) were designed to recruit individuals with partial seizures only and three trials (4%) were designed to recruit individuals with generalised tonic clonic seizures with or without other generalised seizure types or unclassified seizure types only. The remaining 49 trials (64%) were designed to recruit individuals with partial or generalised tonic clonic seizures with or without other generalised seizure types. However, five of these trials did not state the proportion of individuals with each seizure type recruited.

All trials recruited individuals with new onset seizures; within three trials individuals with new onset seizures following stroke were recruited. Fifty-four trials (70%) recruited only individuals with no previous AED treatment and fourteen trials also permitted the

recruitment of individuals with relapsed or 'under-treated' seizures (18%). The remaining six trials did not state whether previous AED use was permitted for inclusion.

5.2.3.5 Methodological quality of included studies

Methodological quality of the included studies was assessed in all included studies using the Cochrane Collaboration's tool for assessing risk of bias [248]. The following methodological criteria are assessed according to this tool:

- Domain 1: Selection bias (sequence generation)
- Domain 2: Selection bias (allocation concealment)
- Domain 3: Performance bias (blinding of participants and personnel)
- Domain 4: Detection bias (blinding of outcome assessment)
- Domain 5: Attrition bias (incomplete outcome data)
- Domain 6: Reporting bias (selective outcome reporting)
- Domain 7: Other bias (any issues not covered by above domains)

Risk of bias assessments were made using information in all published reports of trials in addition to any unpublished information provided following IPD requests. Table 11 summarises the methodological quality of the 77 trials eligible for inclusion in the NMA and further discussion of the risk of bias assessment can be found in the Cochrane IPD-NMA [69].

All trials were described as randomised but 37 trials (48%) did not provide details of how the random sequence was generated so were judged to be at unclear risk of selection bias. The remaining trials provided details of randomisation methods; in 38 trials (49%) this was judged to be adequate and at low risk of selection bias and two trials using alternate randomisation (2%) were judged to be at high risk of selection bias. In 28 trials (36%), an adequate method of allocation concealment was described (low risk of selection bias), in two trials (3%) it was stated that allocation was not concealed for some or all participants (high risk of selection bias) and in 47 trials (61%), no details were provided regarding allocation concealment (unclear risk of selection bias).

Twenty-seven trials (35%) were double-blinded (low risk of performance bias), 32 trials (42%) were open label (high risk of performance bias) and it was not stated whether participants and personnel were blinded in the remaining 18 trials (23%, unclear risk of performance bias). Twelve trials (16%) also stated that outcome assessors were blinded (low risk of detection bias), 27 trials (35%) stated that outcome assessment was not blinded (high risk of detection

bias) and it was not stated whether outcome assessment was blinded in the remaining 38 trials (49%, unclear risk of detection bias).

Table 11: Methodological quality of studies included in the Cochrane Epilepsy Review and IPD-NMA

	Risk of bias domain[2,3]						
Trial Reference[1]	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Aikia 1992	Unclear	Unclear	Low	Unclear	High	Low	Low
Banu 2007	Unclear	Low	Low	Unclear	Low	Low	High
Baulac 2012	Low	Low	Low	Low	Low	Low	Low
Bidabadi 2009	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Bill 1997	Low	Low	Low	Unclear	Low	Low	Low
Biton 2001	Low	Unclear	Low	Low	Low	Low	Low
Brodie 1995a	Low	Low	Low	Unclear	Low	Low	Low
Brodie 1995b	Low	Low	Low	Unclear	Low	Low	Low
Brodie 1999	Low	Low	Low	Unclear	Low	Low	Low
Brodie 2002	Low	Unclear	Low	Unclear	Low	Low	Low
Brodie 2007	Low	Low	Low	Unclear	Low	Low	Low
Callaghan 1985	Unclear	Low	Unclear	Unclear	Low	Low	Low
Capone 2008	Unclear	Unclear	Unclear	Unclear	Low	Unclear	Low
Castriota 2008	Unclear	Unclear	High	High	Unclear	Unclear	Low
Chadwick 1998	Low	Low	High	Unclear	Low	Low	Low
Chen 1996	Low	Unclear	Unclear	Low	Unclear	Low	Low
Cho 2011	Unclear	Unclear	Unclear	Low	Unclear	Low	Low
Christe 1997	Unclear	Unclear	Low	Unclear	High	Low	Low
Consoli 2012	Low	Unclear	High	High	High	Low	High
Cossu 1984	Unclear	Unclear	Low	Unclear	Low	Unclear	High
Craig 1994	Low	Low	High	Low	Low	Low	Low
Czapinski 1997	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Dam 1989	Unclear	Unclear	Low	Unclear	High	Low	Low
de Silva 1996	Low	Low	High	High	Low	Low	Low
Dizdarer 2000	High	High	High	High	Low	Low	Low
Donati 2007	Low	Low	High	High	High	Low	Low
Eun 2012	Low	Unclear	High	High	Low	Low	Low
Feksi 1991	Low	Low	Unclear	Unclear	High	Low	High
Forsythe 1991	High	Unclear	High	Low	Low	Unclear	Low
Fritz 2006	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Gilad 2007	Unclear	Unclear	High	High	Low	Low	Unclear
Guerreiro 1997	Low	Low	Low	Unclear	High	Low	Low
Heller 1995	Low	Low	High	High	Low	Low	Low
Jung 2015	Low	Low	High	Low	High	Low	Low
Kalviainen 2002	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Kopp 2007	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Low

Korean Lamotrigine Study Group 2008	Unclear	Unclear	High	High	High	Low	Low
Kwan 2009	Unclear	Unclear	High	High	Low	Low	Low
Lee 2011	Low	Unclear	High	High	Low	Low	Low
Lukic 2005	Unclear	Unclear	High	High	Unclear	Unclear	Low
Mattson 1985	Unclear	Unclear	Low	Unclear	Low	Low	Low
Mattson 1992	Low	Low	Low	Unclear	Low	Low	Low
Mitchell 1987	Unclear	Unclear	High	High	Low	Low	High
Miura 1990	Unclear	Unclear	Unclear	Unclear	Low	Unclear	Low
Motamedi 2013	Low	Unclear	Low	Unclear	High	Low	Low
NCT01498822	Unclear	Unclear	High	High	High	Low	Low
NCT01954121	Unclear	Unclear	High	High	High	High	Low
Nieto-Barrera 2001	Low	Low	High	High	Low	Low	Low
Ogunrin 2005	Low	Low	Low	Low	Low	Low	Low
Pal 1998	Low	Unclear	High	Low	Low	Low	Low
Placencia 1993	Low	High	Unclear	Unclear	Low	Low	High
Privitera 2003	Low	Unclear	Low	Unclear	Low	Low	Low
Pulliainen 1994	Unclear	Unclear	Unclear	Low	High	Unclear	Low
Ramsey 1983	Unclear	Unclear	Low	Unclear	High	Low	Low
Ramsey 1992	Low	Unclear	High	High	Low	Low	Low
Ramsey 2007	Unclear	Unclear	Low	Unclear	Unclear	Unclear	Low
Ramsey 2010	Unclear	Unclear	Low	Low	Low	Low	Low
Rasgoti 1991	Unclear	Unclear	Unclear	Unclear	Unclear	Low	Low
Ravi Sudhir 1995	Unclear	Unclear	Unclear	Unclear	High	Unclear	Low
Resendiz 2004	Low	Unclear	High	High	High	Low	Low
Reunanen 1996	Low	Low	High	High	Low	Low	Low
Richens 1994	Low	Low	High	High	Low	Low	Low
Rowan 2005	Low	Low	Low	Unclear	Low	Low	Low
Saetre 2007	Unclear	Unclear	Low	Unclear	Low	Low	Low
SANAD A 2007	Low	Low	High	High	Low	Low	Low
SANAD B 2007	Low	Low	High	High	Low	Low	Low
Shakir 1981	Low	Low	Unclear	Unclear	Low	Low	Low
So 1992	Unclear	Unclear	Low	Unclear	High	Low	Low
Steiner 1999	Unclear	Unclear	Low	Low	Low	Low	Low
Steinhoff 2005	Unclear	Unclear	High	High	High	Low	Low
Stephen 2007	Unclear	Unclear	High	High	Low	Low	High
Suresh 2015	Unclear	Unclear	High	High	Unclear	High	Low
Thilothammal 1996	Low	Unclear	Unclear	Unclear	Low	Low	Low
Trinka 2013	Unclear	Low	High	High	Low	Low	Low
Turnbull 1985	Unclear	Unclear	Unclear	Unclear	Low	Low	Low
Verity 1995	Low	Low	High	High	Low	Low	Low
Werhahn 2015	Low	Low	Low	Unclear	Low	Low	Low

1. See Appendix 10 for reference of the primary publication of each trial
2. See Chapter 5.2.3.5 for definitions of the domains of risk of bias.
3. See Cochrane IPD-NMA for full details of risk of bias assessments [69].

In theory, a review using IPD should overcome issues of attrition bias and reporting bias as unpublished data can be provided, unpublished outcomes calculated and all randomised participants can be analysed by an ITT approach. Forty seven trials (61%, including all trials providing IPD) were at low risk of attrition bias, as attrition rates were reported and an ITT approach was used for analysis. Eighteen trials (23%) were judged to be at high risk of attrition bias as participants were excluded from analysis and/or an ITT approach was not used. For the remaining 12 trials, mostly those without full-text publications available, were judged to be at unclear risk of attrition bias as insufficient information was available to make a judgement. Sixty one trials (79%, including all trials providing IPD) were at low risk of reporting bias. Two trials (6%) were judged to be at high risk of reporting bias as results were not provided for all listed outcomes and for the remaining 14 trials, mostly without full-text publications available, were judged to be at unclear risk of reporting bias as insufficient information was available to make a judgement.

Another source of bias was detected in eight trials (10%). Inconsistencies between IPD provided and published results were found in four trials which could not be resolved by original trial authors. For one trial, too many inconsistencies were present for this data to be usable. Three trials were likely to be statistically underpowered and in one trial it was unclear if all participants were receiving AED monotherapy treatment.

5.3 IPD request methodology and preparation of IPD for network meta-analysis

5.3.1 IPD requests prior to 2012

As described in Chapter 5.2.2.3, 77 trials were identified as eligible for the current IPD-NMA. Two of the trials, recruiting 2437 participants, were conducted within the Clinical Trials Research Centre at the University of Liverpool, therefore IPD was available on site and did not need to be requested [264, 265]. Thirty of these trials reported in 29 publications (see Chapter 5.2.2.3 for details of two trials within Brodie 1995 publication) had been identified previously by the Cochrane Epilepsy Group for inclusion in IPD-MA or IPD-NMA and IPD requests had been initiated between the years of 1995 and 2005 approximately.

Many of the data requests were initiated at a time when IPD-MA was a relatively novel design and when e-mail was not commonly used. Exchanges were conducted by letter, fax, telephone and face-to-face meetings with trial investigators. Some datasets supplied had never been computerised. Due to the informal nature of many of these requests, no data sharing agreements were exchanged and very little documentation was retained regarding

the time-to-complete data requests (Professor Anthony Marson (Co-ordinating Editor – Cochrane Epilepsy Group), personal communication, January 2012). A summary of the results of these 29 IPD requests made is provided in Chapter 5.4.1.

5.3.2 IPD requests from 2012 to 2015

IPD requests for newly identified trials were initiated from January 2012 following the first systematic search in November 2011. Further requests were initiated in February 2013 and December 2013 following updated searches in 2012 and 2013 (see Figure 6). In total, requests for IPD from 39 eligible trials were initiated over this time frame.

For all trials meeting inclusion criteria, a data request letter and a data request form were sent to the first or corresponding author of the trial or to the trial sponsor(s) as appropriate. The recipients of IPD requests are referred to as data providers herein. Data request letters and data request forms were sent by as many methods as possible (e-mail, postal mail, fax). A copy of a template request letter and the data request form can be found in Appendix 11. The data request form asked data providers if the following information was available:

- **Trial methods:**
    - method of generation of random list
    - method of concealment of randomisation
    - stratification factors
    - blinding methods
- **Participant covariates:**
    - sex and age
    - seizure types
    - epilepsy status (newly diagnosed / relapsed seizures following drug withdrawal)
    - time between first seizure and randomisation
    - number of seizures prior to randomisation (with dates)
    - presence of neurological signs
    - electroencephalography (EEG) results, computed tomography (CT) and/or magnetic resonance imaging (MRI) results
    - aetiology of seizures (if known)
- **Follow-up data:**
    - treatment allocation
    - date of randomisation and dates of follow-up

- dates of seizures post randomisation or seizure frequency data between follow-up visits
- dates of treatment withdrawal and reason(s) for treatment withdrawal
- starting dose of treatment
- dates of dose changes
- adverse events reported.

The request also included any available, related documents such as case report forms, trial protocols, clinical summaries etc. Following the return of the data request form, data providers were asked to provide the data indicated to be available; data was accepted in any computerised format.

In the event of no response to the request, a follow-up letter, e-mail and/or fax (as previously sent) was sent to the same data provider first contacted. If no response was received following the second communication, an alternative trial author or sponsor was contacted if their contact details could be sourced. All data requests were considered ongoing until IPD was provided or a data provider confirmed that IPD could not be made available. Where IPD could not be made available, the quoted reason for non-availability was recorded and an additional request for any unpublished AD related to the outcomes of interest of the review was made if appropriate.

Any outstanding data requests were considered unsuccessful at the end of 2015; at this point, the database was closed to begin analysis (see Figure 6). Where IPD was not available for analysis (confirmation from data provider that IPD could not be provided or no responses received to any requests), an assessment was made by one reviewer (SJN) of whether any relevant and appropriate AD had been reported in the publication or could be indirectly estimated for inclusion in a combined IPD and AD analysis.

5.3.3 Preparation of IPD for analysis

All IPD provided (prior to 2012 or in requests made from 2012 to 2015) was stored on a secure, dedicated network drive which was accessible only to the statisticians performing analysis (SJN, MS, CTS) and the Computer Services Department of the University of Liverpool for maintenance purposes. All provided data was checked for consistency and prepared for analysis according to a pre-specified procedure which is detailed in Appendix 12.

The procedure was designed to ensure a standardised and consistent approach to preparation of all data for the IPD-NMA. However given differences in format and content of the datasets provided, the procedure serves as a guidance document rather than a direct algorithm for the preparation of IPD for analysis. The procedure was piloted firstly by SJN on a dataset provided prior to 2012 and secondly on the first dataset received from the 2012-2015 requests by SJN and MS independently and results were compared. Updates were made accordingly following piloting and the procedure outlined in Appendix 12 was applied to all newly provided datasets and the remaining datasets provided prior to 2012. Stages of the procedure were ignored if not applicable to the dataset in question and additional steps were added in on a case-by-case basis if deemed necessary.

Time to check and prepare a dataset for analysis following the procedure outlined in Appendix 12 (including time required for data providers to provide clarification of inconsistencies) was monitored approximately, but not recorded precisely. Due to different formats and content of IPD provided, resulting in varying numbers and extents of checks required, a comparison of total data checking and preparation time across all datasets was not deemed appropriate. Furthermore, datasets provided prior to 2012 had been checked and prepared for analysis by different statisticians working on the original Cochrane Reviews and IPD-MAs [67, 285, 291-296] and it is unlikely that a standardised procedure was used. Applicable sections of the procedure were applied to the original datasets provided to 2012 by SJN where possible (replication of published results, calculation TTE outcomes etc.). However, contact with original data providers had not been maintained, therefore clarification of any inconsistencies was not possible.

5.3.3.1 Consistency checking

Appendix 12 provides full details of relevant consistency checks. In summary, after basic checks on the content of the IPD provided has been performed, IPD was cross-checked against any published reports of the trial and published results reproduced where possible. Also where possible, a review was also conducted of the chronological randomisation sequence by checking the balance of prognostic factors, taking account of factors stratified for in randomisation procedure. Where any missing data, errors or inconsistencies were found, data providers were contacted for clarification. If large or major inconsistencies were present which could not be resolved by data providers, the data was not included in any analyses. If minor inconsistencies were present, data was included in analysis and sensitivity analyses were conducted to test the robustness of results [68].

5.3.3.2 Calculation of time-to-event outcomes

The following outcome measures were of interest in the IPD-NMA (see Chapter 3.1.2 for further discussion of the clinical relevance of these outcomes). Reporting of these outcomes in the original trial report was not an eligibility requirement for the Cochrane IPD-NMA.

- **Primary outcome**
  - Time-to-withdrawal of allocated treatment (retention time)
- **Secondary Outcomes**
  - Time-to-12-month remission after randomisation
  - Time-to-6-month remission after randomisation
  - Time-to-first seizure post randomisation

Outcomes were calculated from IPD provided following consistency checks and resolution of inconsistencies as far as possible (see Chapter 5.3.3 and Appendix 12 for further details). Outcomes for all datasets provided from 2012-2015 were calculated by one statistician (SJN or MS) and verified by the other. Outcome data which had been previously prepared for the original Cochrane IPD-MAs was verified by SJN (see Chapter 5.3.3 for further details).

For the analysis of 'time-to-withdrawal of allocated treatment,' an 'event' was defined as either the withdrawal of the allocated treatment due to poor seizure control or adverse events or both. Non-compliance with the treatment regimen or the addition of another antiepileptic drug were also be classed as 'events'. The outcome was censored if treatment was withdrawn because the individual achieved a period of remission, if a participant withdrew from allocated treatment for reasons not related to the treatment (such as loss to follow-up) or if the individual was still on allocated treatment at the end of follow-up. Two authors (SJN and AGM) independently reviewed reasons for treatment withdrawal for classification as events or censored observations, and disagreements were resolved by discussion or by involving a third author (CTS).

Calculation of secondary outcomes required seizure dates after randomisation. If seizure data were provided in terms of the number of seizures recorded between clinic visits rather than specific dates of seizures, to enable the calculation of TTE outcomes, linear interpolation was applied to estimate dates of seizures between follow-up visits. For example, if the trial recorded 4 seizures between 2 visits that occurred on 1 March 2010 and 1 May 2010 (interval of 61 days), then the date of first seizure would be approximately 13 March 2010.

Time-to-6-month and 12-month remission were calculated from the date of randomisation to the date (or estimated date) the individual had first been free of seizures for six or 12 months respectively. If the person had one or more seizures during the trial, a six-month or 12-month seizure-free period could also occur between the date of two seizures during the trial (or between the date of the last seizure and date of last follow-up). Time-to-first seizure was calculated from the date of randomisation to the date (or estimated date) that their first seizure occurred.

If seizure data were missing for a particular visit, these secondary outcomes were censored at the previous visit. These outcomes were also censored if the individual died or if follow-up ceased prior to the occurrence of the event of interest. Under an ITT approach, individuals who withdrew from allocated treatment but did not withdraw from follow-up (e.g. those who remained in the trial on an alternative treatment) were not censored at the date of withdrawal from treatment and remained 'at risk' for the secondary seizure outcomes.

5.4 Results of IPD Requests

At the end of 2015 the IPD database was closed to begin analysis and any outstanding requests at that point were considered to be unsuccessful. An additional systematic search was carried out in July 2016 in line with Methodological Expectations of Cochrane Intervention Reviews that published reviews must be as up-to-date as possible [297]. This search identified six eligible trials (recruiting 1460 participants). IPD requests have been initiated for each of these trials (via methods described in Chapter 5.3.2) and any IPD provided from these trials will be included in an update of the Cochrane IPD-NMA.

5.4.1 IPD requests prior to 2012

As outlined in Chapter 5.3.1, 30 unique trials reported in 29 publications were identified previously by the Cochrane Epilepsy Group for inclusion in IPD-MA or IPD-NMA and IPD requests had been initiated between the years of 1995 and 2005 approximately. According to the type of study sponsorship, fifteen of these studies are defined as academic studies, ten are defined as pharmaceutical studies and four are defined as government studies (see Chapter 5.2.3.4 for further details of definitions).

IPD was requested for a total of 5887 participants from these 30 trials (29 requests) and IPD was provided for 4703 (80%) participants from 18 (62%) of these 29 requests. IPD was provided from trials published between 1985 and 2001. Over 90% of IPD requested from pharmaceutical and government sponsored studies was successfully received (data provided

for 3695 out of 4084 participants from 12 out of 14 studies (86%)). However, only 56% of IPD (from 1008 out of 1803 participants) requested from 6 out of 15 academic sponsored studies (40% of studies) could be retrieved (see Table 12). IPD was not retrieved from a total of 11 eligible trials, published between 1981 and 1997, recruiting 1184 participants (38% of all eligible trials); for the majority of these trials, data had been lost or was no longer available due to the time elapsed since the trial (Table 12).

5.4.2 IPD requests from 2012-2015

As outlined in Chapter 5.3.2, 39 IPD requests were initiated between 2012 and 2015. According to the type of study sponsorship, 24 of these studies are defined as academic studies, 14 are defined as pharmaceutical studies and one is defined as a government study (see Chapter 5.2.3.4 for further details of definitions).

IPD was requested for a total of 8177 participants from these 39 trials. Four of the requests for pharmaceutical studies were made via data sharing portal ClinicalStudyDataRequest.com (CSDR) (or original platform 'GSK Share' between May 2013 and January 2014, see Chapter 1.3 for further details of data sharing portals). All other requests were made directly to the relevant sponsor.

At the close of the database at the end of 2015, IPD had been received for 5251 participants (64% of the total requested) from 15 (38%) trials requested from 2012 to 2015 (Table 12).

Figure 8 shows the duration and outcome of the 39 IPD requests. For the fifteen successful requests, the median time from initial request to receiving data was similar between 24 academic studies (343 days (range 154 to 861 days)) and 14 pharmaceutical studies (363 days (range 280 to 725 days)). The time taken to receive IPD for a single trial via CSDR was 364 days, but it must be noted that the request was first submitted in June 2013 when the platform was newly initiated and processes still under development so this may not reflect current timelines to providing data in CSDR.

IPD was not retrieved from a total of 24 eligible trials published between 1989 and 2012. For 11 trials recruiting 1537 participants, the data provider confirmed that IPD was not available (i.e. a negative response to the data request). The median time from initial request to negative response from these 11 trials was 287 days (range 1 to 784 days).

Figure 8: Duration and outcome of data requests for 39 randomised controlled trials of antiepileptic drugs

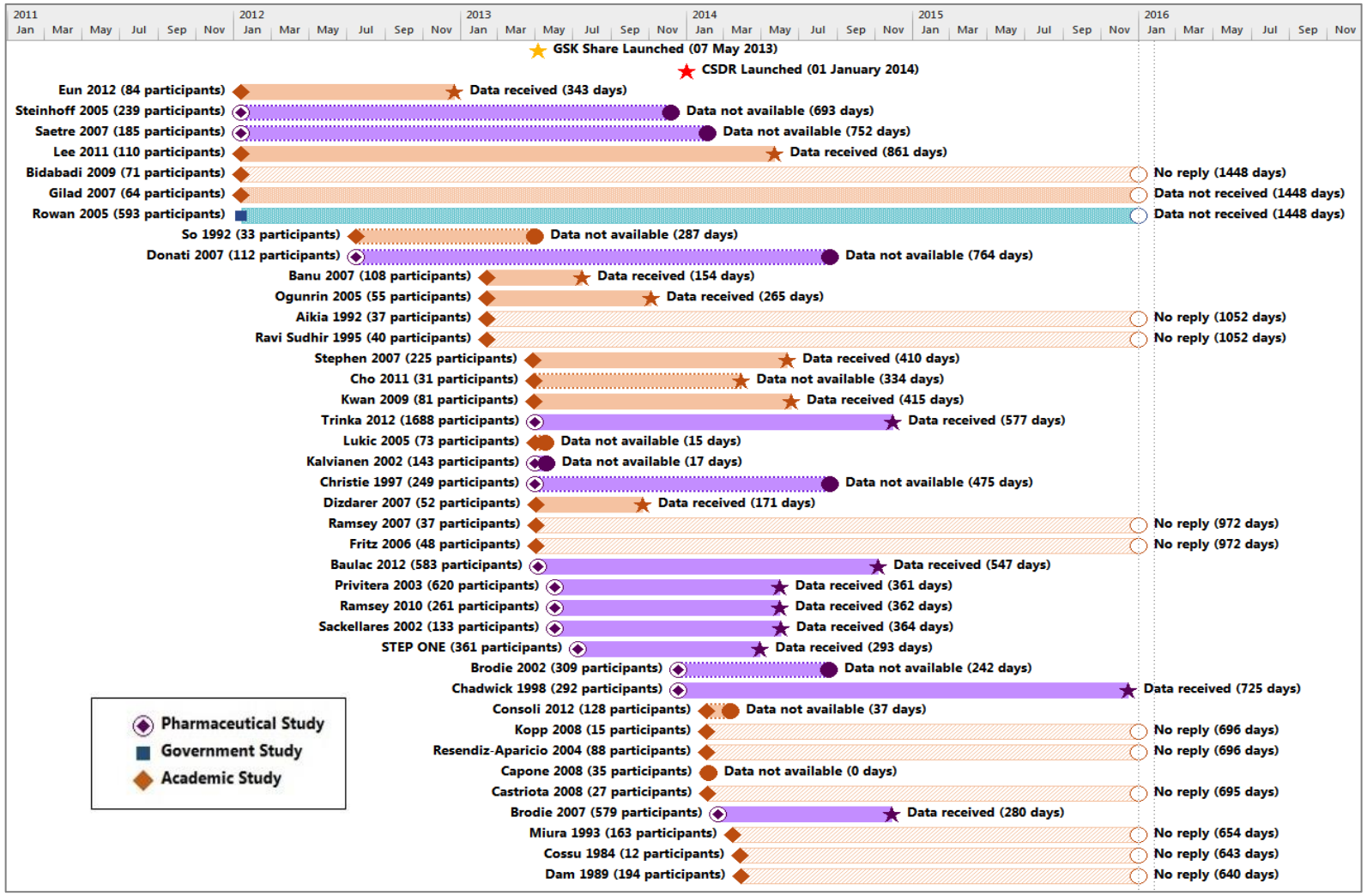


Table 12: Outcome of individual participant data requests conducted between approx. 1995 to 2005 and 2012 to 2015.

Number of studies and participants requested[1,2]	Original requests (approx. 1995 – 2005)				New requests (2012-2015)				All requests (approx. 1995-2015)			
	Ac	**Go**	**Ph**	**Total**	**Ac**	**Go**	**Ph**	**Total**	**Ac**	**Go**	**Ph**	**Total**
Eligible studies requested	15	4	10	**29[3]**	24	1	14	**39**	39	5	24	**68**
Studies providing IPD (n (%))	6 (40%)	3 (75%)	9 (90%)	**18 (62%)**	7 (29%)	0 (0%)	8 (57%)	**15 (38%)**	13 (33%)	3 (60%)	17 (71%)	**33 (49%)**
Eligible participants requested	1803	1178	2906	**5887**	1813	593	5771	**8177**	3616	1771	8677	**14084**
Participants IPD is provided for (n (%))	1008 (56%)	1091 (93%)	2604 (90%)	**4703 (80%)**	717 (40%)	0 (0%)	4534 (79%)	**5251 (64%)**	1725 (48%)	1091 (62%)	7138 (82%)	**9954 (71%)**
Reason data was not available: Number of studies (n (%))												
Data lost	5 (33%)	1 (25%)	0 (0%)	**6 (21%)**	3 (13%)	0 (0%)	0 (0%)	**3 (8%)**	8 (21%)	1 (20%)	0 (0%)	**9 (13%)**
Relevant data not recorded	2 (13%)	0 (0%)	0 (0%)	**2 (7%)**	1 (4%)	0 (0%)	0 (0%)	**1 (3%)**	3 (8%)	0 (0%)	0 (0%)	**3 (4%)**
Unable to make contact with an author / sponsor	1 (7%)	0 (0%)	0 (0%)	**1 (3%)**	11 (46%)	0 (0%)	0 (0%)	**11 (28%)**	12 (31%)	0 (0%)	0 (0%)	**12 (18%)**
Positive response but no data received	1 (7%)	0 (0%)	0 (0%)	**1 (3%)**	1 (4%)	1 (100%)	0 (0%)	**2 (5%)**	2 (5%)	1 (20%)	0 (0%)	**3 (4%)**
Incomplete dataset provided which could not be used	0 (0%)	0 (0%)	1 (10%)	**1 (3%)**	0 (0%)	0 (0%)	0 (0%)	**0 (0%)**	0 (0%)	0 (0%)	1 (4%)	**1 (1%)**
Local authority / ethical restrictions	0 (0%)	0 (0%)	0 (0%)	**0 (0%)**	1 (4%)	0 (0%)	0 (0%)	**1 (3%)**	1 (3%)	0 (0%)	0 (0%)	**1 (1%)**
"Data not available"[4]	0 (0%)	0 (0%)	0 (0%)	**0 (0%)**	0 (0%)	0 (0%)	3 (21%)	**3 (8%)**	0 (0%)	0 (0%)	3 (13%)	**3 (4%)**
Costs of providing data are prohibitive	0 (0%)	0 (0%)	0 (0%)	**0 (0%)**	0 (0%)	0 (0%)	2 (14%)	**2 (5%)**	0 (0%)	0 (0%)	2 (8%)	**2 (3%)**
Country specific restrictions	0 (0%)	0 (0%)	0 (0%)	**0 (0%)**	0 (0%)	0 (0%)	1 (7%)	**1 (3%)**	0 (0%)	0 (0%)	1 (4%)	**1 (1%)**
Total	**9 (60%)**	**1 (25%)**	**1 (10%)**	**11 (38%)**	**17 (61%)**	**1 (100%)**	**6 (43%)**	**24 (62%)**	**26 (67%)**	**2 (40%)**	**7 (29%)**	**35 (51%)**

Abbreviations: Ac: Academic studies, IPD: Individual Participant Data, Go: Government Studies, NMA: Network Meta-Analysis, Ph: Pharmaceutical Studies (see next page for footnotes)

1. In addition, we had IPD available from our own 'SANAD' trial [264, 265], the largest ever in epilepsy at the time, which randomised 2437 participants
2. An additional search was conducted in July 2016 and six eligible trials (recruiting 1460 participants) were identified. IPD requests for these trials have been initiated and any IPD made available will be included in an update of the Cochrane IPD-NMA (see Chapter 5.4.2 for further details).
3. In total, 29 data requests were made, one request resulted in the provision of data from two trials (reported in a single publication). These two trials are treated as a single request in this table.
4. Refers to a non-specific reason (data not available for secondary analysis with no further reason provided).

Reasons for negative response were:

(i) country specific restrictions over anonymisation of data (one request submitted to CSDR for a pharmaceutical study conducted in 2005)

(ii) cost of retrieving and preparing data prohibitive due to age of study (two requests submitted to CSDR for pharmaceutical studies conducted in 2002 and 2007)

(iii) data cannot be made available, no more specific details provided (three requests directly to pharmaceutical sponsors for studies conducted between 1997 and 2007)

(iv) concerns regarding ethical approval for sharing data (one academic author, study conducted 2011)

(v) the data requested were not recorded (one academic author, study conducted 2005)

(vi) data were lost (three academic authors of studies conducted between 1992 and 2012; one of which provided additional unpublished summary data)

For the remaining 13 trials, two (one government and one academic) had indicated an initial positive response to data requests but data was not provided by the close of database, whilst for 11 trials (nine academic and two pharmaceutical) no response was received to any communications. These 13 data requests were closed at a median of 972 (range 640 to 1448 days) after initial request (Figure 8). If response is received to any of these 13 data requests following the close of database and IPD is subsequently provided, it will be included in an update of the Cochrane IPD-NMA.

It must be emphasised when interpreting the timelines of the requests between 2012 and 2015, that data sharing policies and platforms were under development, and that all of the pharmaceutical sponsors contacted directly at the time of request have since committed to CSDR or an equivalent data sharing platform such as YODA [298].

5.4.3 Total IPD available for Cochrane IPD-NMA

At the close of database at the end of 2015, the total number of participants data provided from IPD requests was 9954 out of 14084 participants (71% of total eligible participants requested) from 33 out of 68 trials (49% of eligible trials requested, Table 12).

As outlined in Chapter 5.3.1, IPD for two trials recruiting 2437 participants were available on site at the University of Liverpool (data requests not required) and a single data request resulted in provision of data from two separate trials. Further, IPD requests for six trials, recruiting 1460 participants, identified and initiated following the close of database did not provide any IPD for the current Cochrane IPD-NMA (but any IPD provided will be included in an update of the IPD-NMA, see Chapter 5.4 for further details).

Therefore, IPD for 12,391 out of a total of 17,961 eligible participants (69% of total data) from 36 out of the 77 eligible trials (47%) was provided for the current Cochrane IPD-NMA. The statistical analyses and results of the Cochrane IPD-NMA are presented in Chapter 6.

Data were available for the following participant characteristics (percentage of 12,391 participants with data available): sex (99.5%, data missing for 75 participants), seizure type (96%, data missing for 555 participants), drug randomised (99.9%, data missing for 11 participants), age at randomisation (99%, data missing for 98 participants), number of seizures in six months prior to randomisation (83%, data missing for 2135 participants), and time since first seizure to randomisation (37%, data missing for 7820 participants).Thirteen trials provided the results of neurological examinations for 5367 participants (43%). Seventeen trials provided electroencephalographic (EEG) results for 2990 participants (24%). Fifteen trials provided computerised tomography/magnetic resonance imaging (CT/MRI) results for 2083 participants (16%).

Sufficient IPD was provided to calculate all four outcomes (see Chapter 5.3.3.2) for 20 of the 36 trials. Time-to-12-month remission could not be calculated for nine trials as the duration of the trial was less than 12 months and for a further four trials, neither time-to-12-month remission or time-to-6-month remission could be calculated as the duration of the trial was less than 6 months. For one additional trial, only the date of first seizure recurrence after randomisation was provided (dates of subsequent seizures not available), therefore only time-to-first seizure could be calculated and remission outcomes could not. For four trials, time-to-withdrawal of allocated treatment could not be calculated; for three trials, insufficient information was available regarding dates or reasons for withdrawal and for one trial, all participants completed the twelve week follow-up without treatment withdrawal.

Table 13 shows the number of participants randomised to each of the 10 drugs, split according to the trials for which IPD were available and not available.

Table 13: Number of participants randomised to each drug for trials with IPD provided or not provided

Trials providing IPD											
Trial \ Drug[1]	**CBZ**	**PHB**	**PHT**	**VPS**	**LTG**	**OXC**	**LEV**	**TPM**	**GBP**	**ZNS**	**Total[5]**
Banu 2007	54	54	0	0	0	0	0	0	0	0	108
Baulac 2012	301	0	0	0	0	0	0	0	0	282	583
Bill 1997	0	0	144	0	0	143	0	0	0	0	287
Biton 2001	0	0	0	69	66	0	0	0	0	0	136
Brodie 1995a	66	0	0	0	70	0	0	0	0	0	136
Brodie 1995b	63	0	0	0	61	0	0	0	0	0	124
Brodie 1999	48	0	0	0	102	0	0	0	0	0	150
Brodie 2007	291	0	0	0	0	0	288	0	0	0	579
Chadwick 1998	74	0	0	0	0	0	0	0	218	0	292
Craig 1994	0	0	81	85	0	0	0	0	0	0	166
de Silva 1996	54	10	54	49	0	0	0	0	0	0	173
Dizdarer 2000	26	0	0	0	0	26	0	0	0	0	52
Eun 2012	41	0	0	0	43	0	0	0	0	0	84
Guerreiro 1997	0	0	94	0	0	99	0	0	0	0	193
Heller 1995	61	58	63	61	0	0	0	0	0	0	243
Kwan 2009	0	0	0	44	37	0	0	0	0	0	81
Lee 2011	53	0	0	0	57	0	0	0	0	0	110
Mattson 1985	155	155	165	0	0	0	0	0	0	0	475
Mattson 1992	236	0	0	244	0	0	0	0	0	0	480
Nieto-Barrera 2001	202	0	0	0	420	0	0	0	0	0	622
Ogunrin 2005	19	18	18	0	0	0	0	0	0	0	55
Pal 1998	0	47	47	0	0	0	0	0	0	0	94
Placencia 1993	95	97	0	0	0	0	0	0	0	0	192
Privitera 2003 (CBZ)[2]	129	0	0	0	0	0	0	266	0	0	395
Privitera 2003 (VPS)[2]	0	0	0	78	0	0	0	147	0	0	225
Ramsey 1992	0	0	50	86	0	0	0	0	0	0	136
Ramsey 2010	0	0	128	0	0	0	0	133	0	0	261
Reunanen 1996	121	0	0	0	230	0	0	0	0	0	351
Richens 1994	151	0	0	149	0	0	0	0	0	0	300

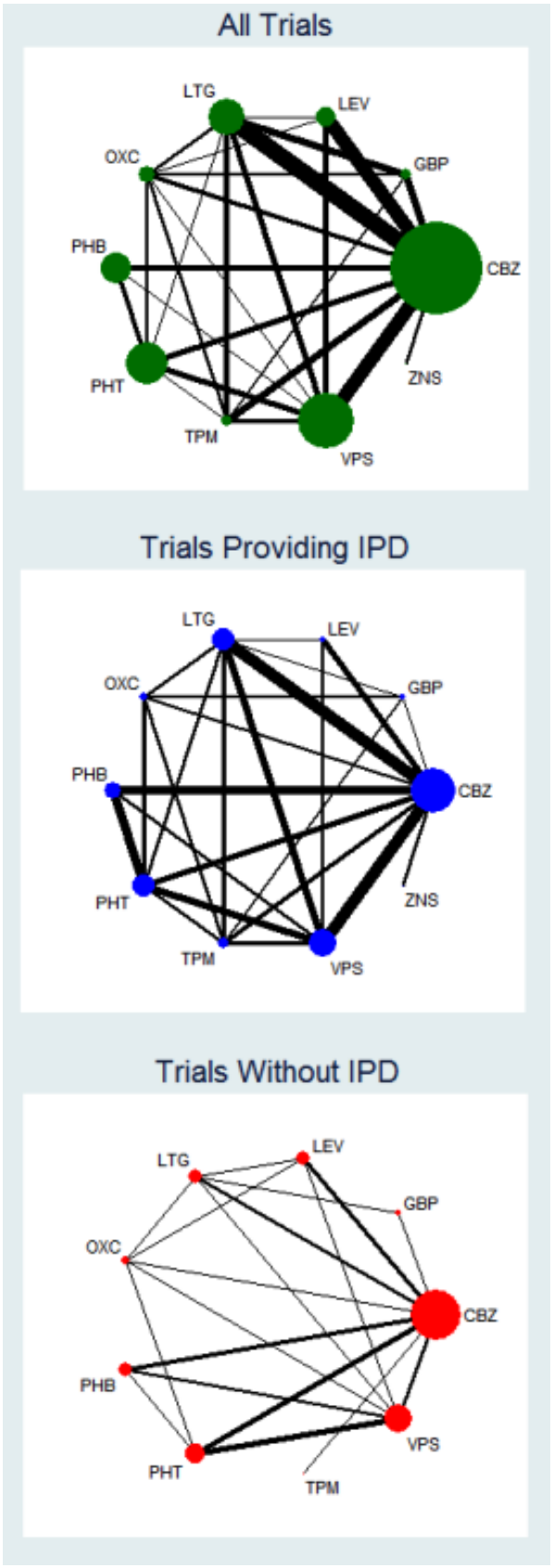
SANAD A 2007	378	0	0	0	378	210	0	378	377	0	1721
SANAD B 2007	0	0	0	238	239	0	0	239	0	0	716
Steiner 1999	0	0	95	0	86	0	0	0	0	0	181
Stephen 2007	0	0	0	109	117	0	0	0	0	0	227
Trinka 2013 (CBZ)[2]	503	0	0	0	0	0	493	0	0	0	999
Trinka 2013 (VPS)[2]	0	0	0	353	0	0	350	0	0	0	703
Turnbull 1985	0	0	70	70	0	0	0	0	0	0	140
Verity 1995	130	0	0	130	0	0	0	0	0	0	260
Werhahn 2015	121	0	0	0	118	0	122	0	0	0	361
Total	**3372**	**439**	**1009**	**1765**	**2067**	**478**	**1253**	**1163**	**595**	**282**	**12391**
Trials not providing IPD											
Trial \ Drug	**CBZ**	**PHB**	**PHT**	**VPS**	**LTG**	**OXC**	**LEV**	**TPM**	**GBP**	**ZNS**	**Total[1]**
Aikia 1992	0	0	18	0	0	19	0	0	0	0	37
Bidabadi 2009	36	35	0	0	0	0	0	0	0	0	71
Brodie 2002	0	0	0	0	151	0	0	0	158	0	309
Callaghan 1985	59	0	58	64	0	0	0	0	0	0	181
Capone 2008	17	0	0	0	0	0	18	0	0	0	35
Castriota 2008	14	0	0	0	0	0	13	0	0	0	27
Chen 1996	26	25	0	25	0	0	0	0	0	0	76
Cho 2011	15	0	0	0	0	0	16	0	0	0	31
Christe 1997	0	0	0	121	0	128	0	0	0	0	249
Consoli 2012	66	0	0	0	0	0	62	0	0	0	128
Cossu 1984	6	6	0	0	0	0	0	0	0	0	12
Czapinski 1997	30	30	30	30	0	0	0	0	0	0	120
Dam 1989	100	0	0	0	0	94	0	0	0	0	194
Donati 2007	28	0	0	29	0	55	0	0	0	0	112
Feksi 1991	152	150	0	0	0	0	0	0	0	0	302
Forsythe 1991	23	0	20	21	0	0	0	0	0	0	64
Fritz 2006	0	0	0	0	21	27	0	0	0	0	48
Gilad 2007	32	0	0	0	32	0	0	0	0	0	64
Jung 2015[3]	64	0	0	0	0	0	57	0	0	0	121
Kalviainen 2002	70	0	0	0	73	0	0	0	0	0	143
Kopp 2007	6	0	0	3	0	0	6	0	0	0	15
Korean LTG Study Group 2008[3]	129	0	0	0	264	0	0	0	0	0	393
Lukic 2005	0	0	0	38	35	0	0	0	0	0	73
Mitchell 1987	15	18	0	0	0	0	0	0	0	0	33
Miura 1990	66	0	51	46	0	0	0	0	0	0	163

Motamedi 2013[3]	0	0	0	0	50	0	50	0	0	0	100
NCT01498822[3]	0	0	0	0	0	178	175	0	0	0	353
NCT01954121[3]	215	0	0	0	0	0	218	0	0	0	433
Pulliainen 1994	23	0	20	0	0	0	0	0	0	0	43
Ramsey 1983	42	0	45	0	0	0	0	0	0	0	87
Ramsey 2007[4]	?	0	0	0	0	0	?	0	0	0	37
Rasgoti 1991	0	0	45	49	0	0	0	0	0	0	94
Ravi Sudhir 1995	20	0	20	0	0	0	0	0	0	0	40
Resendiz 2004	42	0	0	0	0	0	0	46	0	0	88
Rowan 2005	198	0	0	0	200	0	0	0	195	0	593
Saetre 2007	92	0	0	0	93	0	0	0	0	0	185
Shakir 1981	0	0	15	18	0	0	0	0	0	0	33
So 1992	17	0	0	16	0	0	0	0	0	0	33
Suresh 2015[3]	30	0	0	0	0	0	30	0	0	0	60
Steinhoff 2005 (CBZ)[2]	88	0	0	0	88	0	0	0	0	0	176
Steinhoff 2005 (VPA)[2]	0	0	0	30	33	0	0	0	0	0	63
Thilothammal 1996	0	51	52	48	0	0	0	0	0	0	151
Total	**1721**	**315**	**374**	**538**	**1040**	**501**	**645**	**46**	**353**	**0**	**5570**
Grand total	**5093**	**754**	**1383**	**2303**	**3064**	**979**	**1898**	**1209**	**948**	**282**	**17950**
IPD provided	**66%**	**58%**	**73%**	**77%**	**66%**	**49%**	**66%**	**96%**	**63%**	**100%**	**69%**

1. See Appendix 10 for reference of the primary publication of each trial and Chapter 5.2.1.3 for abbreviations of drugs.
2. Trials designed in strata based on clinician recommended treatment. Within the two strata, participants were randomised to TPM (Privitera 2003), LEV (Trinka 2013) or LTG (Steinhoff 2005) or CBZ / VPS depending on the strata. Data analysed according to the separate strata.
3. Trial identified in an updated search in 2016, following closure of database for analysis. IPD request initiated, any IPD provided will be included in an update of the Cochrane IPD-NMA.
4. One trial provided the total number randomised but not the numbers randomised to each group. The 37 participants randomised are counted in the overall totals.
5. Drug allocated missing for 11 participants in the IPD provided.

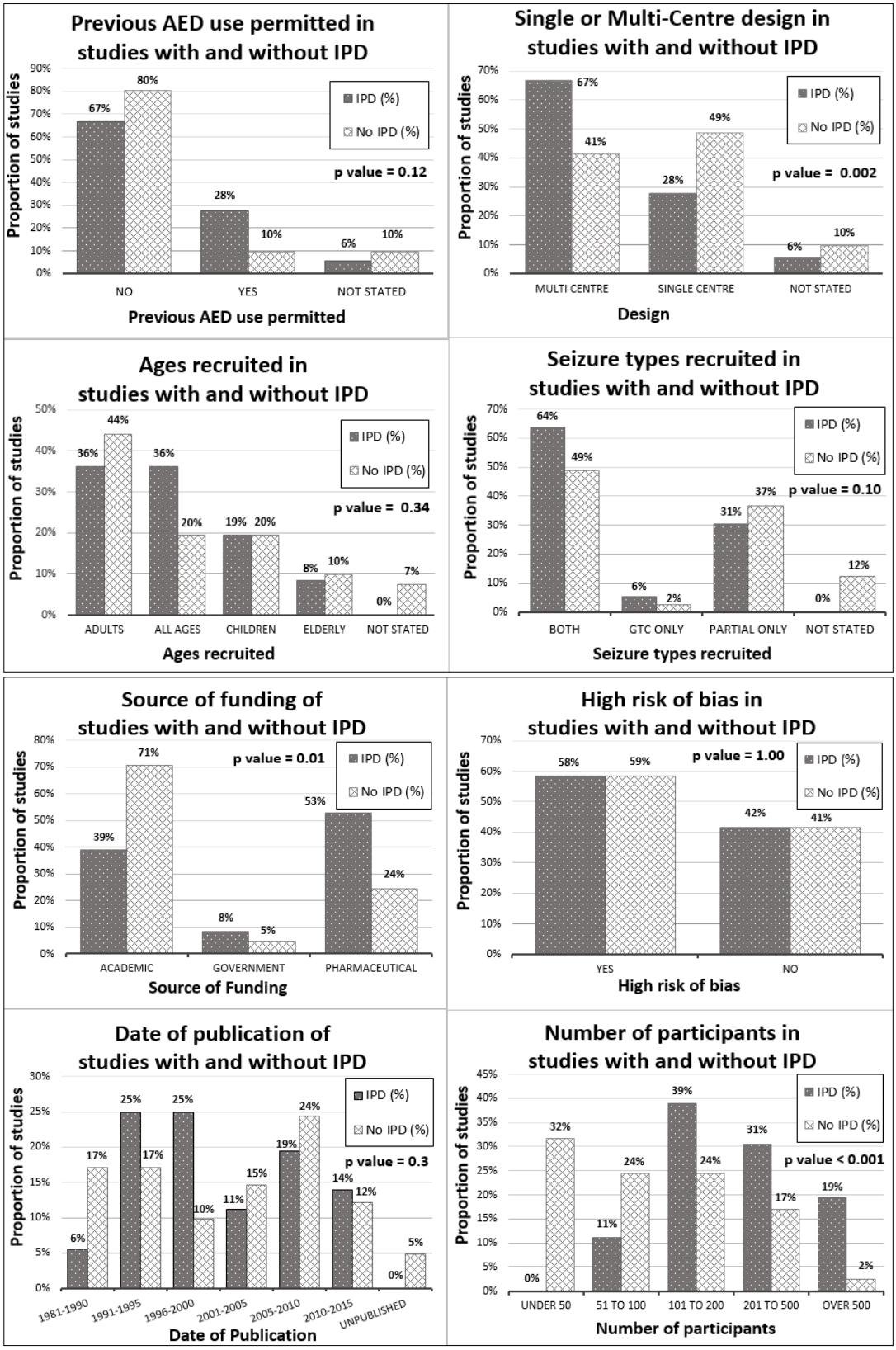
Figure 9 shows the networks of trials with and without IPD provided. Specifically, IPD was provided for all direct pairwise comparisons in the overall network except for OXC compared to VPS and OXC compared to LEV. In fact, out of all drugs included in the network, the lowest proportion of data was received for OXC (49%, Table 13). Aside from the lack of IPD for the OXC / VPS and OXC / LEV comparisons, the networks of the trials with and without IPD appear visually similar (Figure 9).

Figure 9: Network plots of pairwise comparisons for all trials, for trials providing IPD and for trials not providing IPD for a network meta-analysis of 10 antiepileptic drugs



See Table 12 and Table 13 for numbers of trials and participants providing and not providing IPD and Chapter 5.2.1.3 for abbreviations of drugs.

Figure 10: Characteristics of studies providing or not providing IPD



1. See Chapter 5.2.3.4 for further details of the definitions of characteristics
2. High risk of bias defined as at least one domain of the Cochrane Risk of Bias tool deemed to be at high risk of bias, see Chapter 5.2.3.5 for further details
3. P-values calculated from Fishers Exact test due to small numbers in some categories of characteristics. χ^2 p-values were also calculated for completeness, conclusions were unchanged.

Further exploratory examination of the characteristics of the studies with and without IPD was carried out via graphical plots and using Fisher's Exact Test (due to some small numbers in categories of characteristics); see Table 10 and Figure 10. Principally, due to additional communications with data providers regarding trial design for trials providing IPD compared to replying on published information only for trials without IPD; there were much fewer cases of a characteristic 'not stated' in trials providing IPD across all characteristics.

There were no significant differences between trials with and without IPD in terms of ages and epilepsy type recruited and designs permitting participants with previous AED use. There was also no significant difference in terms of methodological quality of the trials with and without IPD (in terms of a least one domain of the Cochrane Risk of bias tool at high risk of bias, Table 11) or in terms of the date of publication of trials with and without IPD.

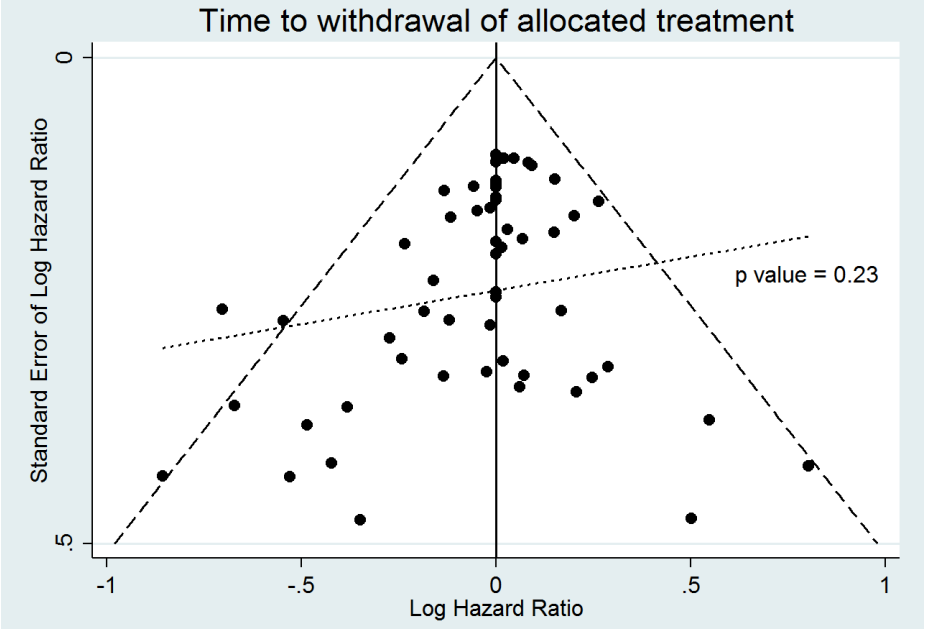
There were significant differences between trials with and without IPD in terms of the design and source of funding of trials; with more multi-centre trials and more pharmaceutical trials providing IPD and more single centre trials and more academic trials not providing IPD. There were also significant differences in the sizes of trials with and without IPD; trials providing IPD tended to be larger than trials not providing IPD. This is also reflected in the proportion of participant data (69%) provided compared to the proportion of trials providing data (49%).

These three significant characteristics are all related to the resources of the trial; larger multi-centre pharmaceutical trials are likely to have more resources to prepare IPD than small, single centre academic trials (also see Table 12 for reasons IPD was not provided).

The majority of studies not providing IPD did not report any of the outcomes of interest to the Cochrane review; this was not a criteria for inclusion in the review as the intention was to calculate these outcomes in a standardised manner using IPD. Therefore, it was difficult to examine whether 'significant' or 'positive' trial results were related to provision of IPD or conversely whether trials with 'negative' results were less likely to provide IPD. Figure 11 shows a comparison-adjusted funnel plot of the trial-specific effect sizes for primary outcome 'Time-to-withdrawal of allocated treatment' for the network of interventions, produced via the 'netfunnel' command in Stata version 14 [299].

Visual inspection of this network funnel plot and a test for asymmetry does not suggest that 'publication bias' (i.e. bias from missing results from studies without IPD) is present for this outcome. However, an association between trial resources and provision of IPD may have implications for updates of this Cochrane IPD-NMA (and for systematic IPD-MA generally).

Figure 11: Comparison-adjusted network funnel plot of trial-specific ln HRs for time-to-withdrawal of allocated treatment



1. Produced via 'netfunnel' command in Stata 14 [299].
2. Dotted line corresponds to the study effect sizes centred at the comparison-specific summary plotted against the standard error of the log hazard ratio. P-value corresponds to a t-test of the gradient of the line (difference from zero which would indicate asymmetry).

5.5 Discussion

5.5.1 Reflection of experiences Group and implications

The first in the series of Cochrane Epilepsy Group IPD-MA was published in 2000 when such an approach was relatively novel and methodology limited [67]. This meta-analysis included IPD from 63% of total trials and 83% of total participants, a good retrieval rate in the wider context of all IPD-MAs (see Chapter 4). Success rate of the group of retrieving IPD has declined from 80% up to around 2005 to 65% between 2012 and 2015 (Table 12). It should be noted that this difference in IPD retrieval rate could be due to chance, nonetheless, the apparent decline in the proportion of IPD provided is of concern.

Results of the systematic review outlined in Chapter 4 of this thesis showed that Cochrane reviews were less likely to retrieve all or a high proportion of IPD than non-Cochrane reviews. This may be explained by the inclusion of thorough search methods within Cochrane reviews, as well as advances in systematic searching of larger electronic databases generally, leading to the identification of larger numbers of studies including more grey literature studies where IPD may be difficult to retrieve with the resources available to review authors, such as Cochrane review authors who usually undertake systematic reviews on a voluntary basis.

Also of concern are changes in the reported reasons for lack of data availability. Table 12 demonstrates that loss of datasets is an issue for academic trials and has been for many years, highlighting a need for better methods of data curation and solutions for long-term storage and access. As highlighted in Chapter 4, lack of specific reasons for unavailability of IPD remains an issue in the reporting of IPD-MA and Cochrane Epilepsy Group experiences of data requesting show that this issue is not restricted to the reporting of IPD-MA and also exists at the study request level; IPD from three out of 35 studies was 'not available' with no further reason stated (Table 12). During more recent requests, 'prohibitive costs' have prevented the sharing of pharmaceutical data. Additional costs and resources associated with IPD-MA are generally considered to be incurred by the meta-analysts [22-24]; however in this new era of commercial data sharing platforms [91] and requirement of high level data de-identification, costs to data providers have certainly increased and should be taken under consideration when planning an IPD-MA [300]. Collaboration, financial or otherwise, between meta-analysts and data providers may assist in sharing costs and resources, potentially maximising retrieval rates of IPD.

Despite our highlighted concerns, recent changes in methods of data sharing have resulted in several benefits to our analyses. Our most common reason for not retrieving data, an issue only for academic trials, was due to failing to make contact with data providers. In our experience, facilities within pharmaceutical data sharing platforms allowed a clear and transparent pathway of communication between data requestors and providers; but the continued benefit of such facilities will require increasing uptake of such platforms from both data users and data providers, from both a pharmaceutical and non-pharmaceutical setting.

In addition to improvements in Good Clinical Practice over time, resulting from regulations such as the European Union Clinical Trials Directive [298], a greater focus on data privacy and additional preparation required to share a dataset has resulted in 'cleaner' datasets provided in the most recent requests compared to earlier and previous requests. As described in Chapter 5.3.3, the exact time required to check and prepare datasets was not recorded but the most recent datasets provided (see Figure 8) required minimal to no clarification of inconsistencies, therefore reducing the relative time to prepare the dataset for analysis.

While under the new framework of data sharing platforms, additional time and resources must allow for constructing a research proposal, independent scientific review, signing of data sharing agreements and de-identification of data; recent datasets provided have required much less data cleaning prior to analysis than in previous years, implying a shift in

the time required to perform an IPD-MA, rather than an increase. However, the research community must ensure that procedures to access IPD do not become over-burdensome, over-costly and prohibitive and that common sense and responsible risk proportionate approaches should be used [79, 84].

5.5.2 Relation to the wider context of clinical trial data sharing

A summary of the development and processes of data sharing platforms such as CSDR and YODA is provided in Chapter 1.3. Both CSDR and the YODA project provide metrics related to submitted research proposals [91, 93]. Metrics provided by YODA relate to requests for de-identified individual participant datasets and to de-identified clinical study reports (CSRs) combined whereas requests to CSDR relate only to individual participant datasets.

To 31st May 2017, out of 291 research proposals submitted to CSDR and processed, 235 (81%) met initial requirements and 198 (68%) were approved (or approved with conditions) by the independent review panel. Out of 296 proposals submitted since January 2014, 50 (17%) were multi-sponsor proposals. Data sharing agreements have been signed for 163 proposals and de-identified datasets provided for 153 proposals. Results for eight data requests (provided with access to between one and fourteen datasets from a single sponsor) have been published to date [69, 100-103, 106-108] and one publication is listed to be in press according to CSDR [91] with other requests remaining 'in process.'

According to metrics provided on the YODA project website, to 1st July 2017, all 65 fully reviewed research proposals were approved by YODA for Johnson & Johnson datasets or CSRs with a median of eight days for YODA project review. Data access (to IPD or CSRs) has been granted for 59 proposals and results of two proposals have been published to date [104, 105]. Further partners of the YODA project are Medtronic Inc. and SI-Bone Inc. The YODA data sharing model was first put into practice via independently performed systematic reviews of the safety and efficacy of Medtronic's recombinant human bone morphogenetic protein-2 product [301, 302]. Up to July 2017, no requests had been made via the YODA project to use SI-Bone data.

Metrics are also provided for the number of 'enquiries' for non-listed studies on both the CSDR and YODA project websites. To 31st March 2017, out of 784 unique studies requested via enquiries on CSDR, following feasibility checks, a positive response was provided for 321 studies (41%) (I.e. a researcher would be able to submit a research proposal for that study and the study would subsequently be listed on CSDR). A negative response was provided for the remaining 389 studies (59%) and access to data from the study would not be provided.

To 1st July 2017; out of 164 unique studies within answered enquiries to the YODA project, a positive response was provided for 100 studies (61%) and a negative response was provided for the remaining 64 studies (38%).

Table 14: Reasons why access could not be provided to de-identified study data following enquiries to CSDR and YODA

Reason why access cannot be provided to data[1]	Number of studies		
	CSDR	YODA[2]	Total
Sponsor did not agree to share data from Phase I or Phase IV trials	143 (31%)	25 (39%)	168 (32%)
Interventional product not approved / unapproved indication or terminated	77 (16%)	11 (17%)	88 (17%)
Considered out of scope as per the Sponsor Specific Information[1]	75 (16%)	0 (0%)	75 (14%)
Cannot be anonymised or likelihood of re-identification	53 (11%)	0 (0%)	53 (10%)
Study ongoing or unpublished	33 (7%)	16 (25%)	49 (9%)
Interventional product not or no longer owned by that sponsor	43 (9%)	0 (0%)	43 (8%)
Sponsor does not have legal authority to share the data	15 (3%)	0 (0%)	15 (3%)
Data sharing commitment with a development partner, partner does not agree to share	1 (<1%)	10 (17%)	11 (2%)
Foreign language studies or documentation not available in English	7 (2%)	4 (7%)	11 (2%)
Data unavailable	7 (2%)	0 (0%)	7 (1%)
Costs / resources prohibitive to providing data due to age of studies	5 (1%)	0 (0%)	5 (<1%)
Beyond the period that sponsor retains study data and documentation	2 (<1%)	0 (0%)	2 (<1%)
Requested data is already publicly available	2 (<1%)	0 (0%)	2 (<1%)
Data cannot be converted to electronic format	0 (0%)	1 (2%)	1 (<1%)
Total	**463**	**67**	**530**

Abbreviations: CSDR: ClinicalStudyDataRequest; YODA: Yale University Data Access

1. Classified into 'general' reasons and merged for all sponsors, exact wording for each sponsor (including Sponsor Specific Information) can be found on CSDR and YODA websites [91, 93]
2. Number of reasons listed (67) is greater than number of studies listed (64); assumed that more than one reason for a negative response can be given per study

A summary of the list of sponsor reported reasons for a negative response to enquiries to CSDR and the YODA project is provided in Table 14. The majority of reported reasons (around 70%) were related to the legal authority of the sponsor to share the requested data; e.g. data from early phase studies (a third of negative responses were requests for Phase I study data), from studies of unapproved or terminated intervention products, from ongoing studies and from studies with a development partner who does not agree to share data. Such reasons,

other than restrictions around ongoing or unpublished studies, would generally not be applicable in the context of IPD syntheses with the objective of establishing clinical effectiveness or safety of approved interventions. Research proposals or further objectives of projects for which early phase or off-licence product use were requested via enquiries to CSDR or the YODA project are not available; but would make an interesting comparison to the objective of the list of approved research proposals provided on the respective websites.

As reflected in the Cochrane Epilepsy Group requests, a proportion (around 15%) of sponsor reported reasons listed on the CSDR website were non-specific such as 'Data not available' or 'Considered out of scope as per the Sponsor Specific Information' without further detail Furthermore, 'Sponsor Specific Information' for all sponsors states (in various wording) that studies with a risk of re-identification such as single centre studies, studies in rare diseases or studies with very small sample sizes will not be provided [91]. Documentation of the anonymisation standards employed by each sponsor are provided but details are not provided on any methodology employed by sponsors to quantify the re-identification risk [303] or criteria for judging when the risk of re-identification is too high to share data, even though this reason is listed for 9% of studies requested via enquires to CSDR.

As outlined in Chapter 5.4.2, three out of the four requests made by the Cochrane Epilepsy Group via CSDR resulted in a negative response related to resources to fulfil the requests (i.e. prohibitive costs or facilities to translate documentation of foreign language studies). Such resource related reasons seem to have been specific to the studies under request for Cochrane Epilepsy reviews and are in the minority (around 2%) of sponsor reported reasons for all negative responses to enquires made to CSDR and the YODA project. Informal comparison of the IPD retrieval rate of the Cochrane Epilepsy Group for the conduct of systematic IPD-MAs and NMAs and reasons provided for unavailability of data to the researcher reported 'approved research proposals' and sponsor reported reasons for negative responses to data requests may provide some insight to the type of data requests and research projects suited to data request platforms.

Considering the 161 approved research proposals published on the CSDR website up to July 2017 (restricted to IPD requests, whereas requests to YODA may be for IPD or CSRs), 30 (19%) clearly describe a proposal for meta-analysis but only 12 (7% of total proposals) clearly describe a systematic approach to meta-analysis. Information provided on the CSDR website regarding research proposals is limited so a larger proportion may in fact intend to perform a systematic IPD-MA, however IPD-MA (systematic or not), are currently in minority of approved research proposals on CSDR. At the time of writing, two systematic IPD syntheses including data provided from a data sharing platform have been published [69, 108].

Results from Chapter 4 suggest as many as 105 new IPD-MAs per year are being published up to 2015, yet only 30 IPD-MA projects seem to have been submitted as research proposals on CSDR between October 2013 and July 2017. This could be for the simple reason that only a small proportion of the 3461 studies listed on CSDR have been identified as eligible for an IPD-MA, or this could indicate that authors conducting IPD-MA over recent years are not using data request platforms; perhaps due to lack of awareness of the existence of such platforms or perhaps because IPD-MA projects are not suited to data request platforms. In fact, up to July 2017, three projects, two identified as IPD-MAs within their research proposals, could not be completed due to the restrictions of the remote platforms preventing merging of individual participant datasets [93]

The suitability and practicality of data request platforms and remote data access for the conduct of IPD-MA should become clear in future years as further research proposals are submitted and further research projects making use of the data available via data request platforms are published.

5.5.3 Concluding remarks

Twenty years of Cochrane Epilepsy IPD-MA and NMA have shown a decline in IPD retrieval rate from has declined from 80% in 2005 to 65% between 2012 and 2015. Recent years, in line with data transparency initiatives in the pharmaceutical industry and across the research community as a whole, have shown that provision of IPD seems to be related to the resources of the trial, with larger multi-centre pharmaceutical trials more likely to have more resources to prepare IPD than small, single centre academic trials. However, resources alone do not guarantee provision of IPD with 'prohibitive costs' preventing sharing of some pharmaceutical data in recent years. On the other hand, loss of datasets continues to be an issue for academic trials and has been for many years, highlighting a need for better methods of data curation and solutions for long-term storage and access.

Consideration of approved research proposals for access to IPD via CSDR show that less than 20% of proposals have been for an IPD-MA, whether systematic or 'opportunistic.' IPD-MAs which are systematic may identify older studies, small studies and foreign language studies, many of which are outside of the scope of CSDR sponsors to provide data due to resources, restrictions and risks from re-identification. While data sharing platforms such as CSDR and YODA may be suitable for some objectives of secondary research, the implications of limited resources, increased costs and increased awareness of data privacy and risks from re-identification may result in a decline in the amount of IPD made available from both a study and patient-level for systematic IPD-MA.

Chapter 6: Approaches to network meta-analysis of antiepileptic drug monotherapy

6.1 Introduction

The clinical setting and rationale of IPD-MAs and IPD-NMAs of AED monotherapy trials performed by the Cochrane Epilepsy Group is described in Chapter 1.2. There are strong clinical beliefs that some AEDs are more effective in certain seizure types than others and current NICE guidelines recommend different first-line treatments for individuals with partial seizures (carbamazepine or lamotrigine) and for individuals with generalised seizures (sodium valproate)[74]. Furthermore, some RCTs of AED monotherapy have built this clinical preference for certain drugs for different seizure types into their design by recruiting participants with different seizure types separately [264, 265] or by stratifying randomisation by 'clinician choice' of first-line treatment according to seizure type [304, 305].

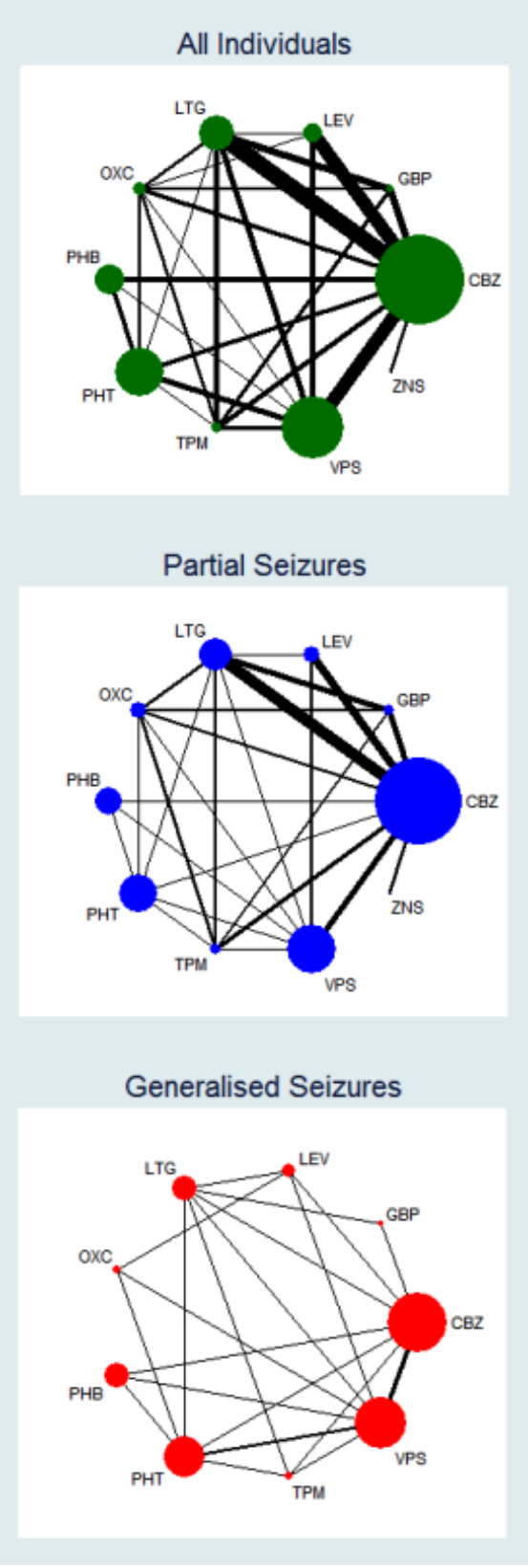
A previous IPD-NMA of AED monotherapy published in 2007 demonstrated results which are in line with current NICE guidelines for recommended first-line AED treatment for individuals with new onset partial or generalised seizures [74, 285] and the objective of the illustrative example of the IPD-NMA described throughout this thesis was to update the previous IPD-NMA, including evidence from all trials published since 2007 and for two additional AEDs, licenced for epilepsy monotherapy treatment after 2007 (levetiracetam and zonisamide).

The objective of this Chapter is to describe the approaches for the statistical methodology of an IPD-NMA of AED monotherapy, taking account of current clinical practice for treating individuals with partial onset and generalised onset seizures. IPD-NMA results for the different approaches described as also presented.

6.2 Statistical methodology

The outcomes of the IPD-NMA are outlined in Chapter 5.3.3.2 and details of data preparation is provided in Chapter 5.3.3. All preparation of data for analysis was performed in SAS statistical software (version 9.3) [306] and all syntheses were performed in Stata statistical software (version 14) [307].

Figure 12: Network plots of pairwise comparisons for all individuals, participants with partial seizures and participants generalised seizures.



11978 participants classified as experiencing partial seizures (66.7% of total), 4407 participants classified as experiencing generalised seizures (24.5% of total) and 1576 had an unclassified or missing seizure type (8.8% of total). Generalised tonic-clonic seizures with or without other seizure types was shortened to 'generalised seizures' for brevity. See Chapter 5.2.1.3 for abbreviations of drugs.

Epilepsy type is classified according to the main seizure type an individual has experienced at baseline (partial-onset or generalised-onset). Partial seizures (simple or complex) and partial secondarily generalised seizures are classified as partial epilepsy. Primarily generalised tonic-clonic seizures (with or without other seizure types) are classified as generalised epilepsy. Figure 12 presents the networks of evidence for participants with partial seizures and with generalised seizures (for all individuals eligible for inclusion in the review, not just those for whom IPD was provided for analyses). Overall and for individuals with partial seizures there are a total of 45 possible pairwise comparisons between the ten AEDs of interest. At the time of analysis, zonisamide (ZNS) had been randomised in one trial recruiting individuals with partial-onset seizures only [267], therefore ZNS does not feature in the network of evidence for generalised seizures. Hence, there are a total of 36 possible pairwise comparisons between the nine AEDs of interest in network of evidence for generalised seizures.

As outlined in Chapter 6.1, current clinical practice treats different seizure types with different drugs, suggesting the existence of a 'clinical' treatment-by-epilepsy type (partial or generalised) interaction. If such an interaction is also present statistically, then the key assumption made in NMA of an exchangeable treatment effect across all included trials would be violated. Furthermore, it would be of little relevance to current and future medical decision making to perform an NMA ignoring the differences in the treatment of partial and generalised seizure types in 'real-world' clinical practice.

Within this Chapter, several approaches are outlined with regard to the association between epilepsy type and treatment effect in IPD-NMA. The NMA was performed in two stages. Firstly, trial-specific estimates of treatment effects relating to a reference AED were obtained by fitting a stratified Cox PH regression model (based on Equation 23, outlined in Chapter 2.3.3) to the entire individual participant dataset. Secondly, the trial-specific estimates obtained from the 'first-stage' were synthesised in an NMA under a multivariate framework. The following sections describe the how epilepsy type was modelled in the 'first-stage' of each approach and the 'second-stage' multivariate approach to NMA.

6.2.2 First-stage: IPD-NMA with separate models by epilepsy type

In this approach, NMA is performed separately for individuals with partial seizures and for individuals with generalised seizures (i.e. no treatment-by-epilepsy type interaction term).

In the first-stage, for the i^{th} participant in the j^{th} trial, the following fixed-effects models, stratified by trial to preserved within-trial randomisation to the IPD datasets for individuals with partial and generalised seizures respectively:

$$h_{ij}(t) = h_{0j}(t)exp((\alpha_{2j}x_{2ij} + \ \ldots + \alpha_{10j}x_{10ij}) \qquad \text{(Equation 29)}$$

$$h_{ij}(t) = h_{0j}(t)exp((\beta_{2j}x_{2ij} + \ \ldots + \beta_{9j}x_{9ij}) \qquad \text{(Equation 30)}$$

Where $x_{2ij}\ \ldots x_{10ij}$ are indicator variables for each of comparator AED of interest in relation to reference treatment CBZ (first-line treatment for partial seizures, constrained to be 0) in Equation 29 and VPS (first-line treatment for generalised seizures, constrained to be 0) in Equation 30. Note that treatment with ZNS (i.e. x_{10ij}) is not included in Equation 30 for individuals with generalised seizures as this treatment is not present in the network for generalised seizures (see Figure 12).

Coefficients $\alpha_{2j}\ \ldots\alpha_{10j}$ for individuals with partial seizures and $\beta_{2j}\ldots\beta_{9j}$ for individuals with generalised seizures correspond to trial-specific fixed effect estimates (i.e. the $\ln(HR)$) of each AED compared to the reference treatment which will be synthesised in NMA.

To facilitate NMA, the models described in Equation 29 and Equation 30 were applied by running the 'stcox' command in Stata via 'mvmeta_make' command. In other words, Equation 29 is fit using the following command for $x_{2ij}\ \ldots x_{10ij}$:

xi: mvmeta_make stcox _ x2 _x3 _x4 _x5 _x6 _x7 _x8 _x9 _x10, strata(trial) ppfix(none) nohr by(trial) names(y S) esave(N) keepmat saving(Partial1) replace

This command outputs and saves a dataset '*Partial1*' with trial-specific estimates (coefficients $\alpha_{2j}\ \ldots\alpha_{10j}$ or equivalently $\beta_{2j}\ldots\beta_{9j}$ for generalised seizures), their associated variances and covariances if applicable (i.e. correlation between treatment effects for trials with more than two treatment arms)[308]. The subsequent dataset is within the correct format for the 'mvmeta' command to be run (see Chapter 6.2.5 for further details).

6.2.3 First stage: IPD-NMA with treatment-by-epilepsy type interaction

In this approach, in the first-stage a single Cox PH model is fitted to the IPD with a treatment-covariate interaction. In other words, for the i^{th} participant in the j^{th} trial, the following model, stratified by trial with treatment-by-epilepsy type interaction, was fitted to the entire individual participant dataset:

$$h_{ij}(t) = h_{0j}(t)exp((\alpha_{2j}x_{2ij} + \ \ldots + \alpha_{10j}x_{10ij}) + z_{ij}\,(\pi_j + \gamma_{2j}x_{2ij} + \ \ldots \ + \ \gamma_{9j}x_{9ij})) \qquad \text{(Equation 31)}$$

Where $x_{2ij} \ \ldots x_{10ij}$ are indicator variables for each of comparator AED of interest in relation to reference treatment CBZ constrained to be 0 and z_{ij} is an indicator variable for whether an individual has partial ($z_{ij} = 0$) or generalised seizures ($z_{ij} = 1$).

Coefficients $\alpha_{2j} \ \ldots \alpha_{10j}$ correspond to the trial-specific effect sizes (i.e. the $\ln(HR)$) of each AED compared to CBZ for individuals with partial seizures, coefficient π_j is the trial-specific effect size of generalised seizures compared to partial seizures (reference) and coefficients $\gamma_{2j} \ldots \gamma_{9j}$ correspond to trial-specific interaction effects of each AED compared to the reference treatment CBZ (i.e. the additional effect of the drug in individuals with generalised seizures compared to individuals with partial seizures).

It should be noted that under this approach, the treatment effect sizes for individuals with generalised seizures are calculated as follows. For example in the j^{th} trial, let $\hat{\alpha}_{2j}$ and $\mathrm{var}(\hat{\alpha}_{2j})$ be the estimates of $log\ (HR)$ and variance for drug 2 compared to CBZ for individuals with partial seizures and let $\hat{\gamma}_{2j}$ and $\mathrm{var}(\hat{\gamma}_{2j})$ be the $log\ (HR)$ and variance for the treatment-epilepsy type interaction term for drug 2 compared to CBZ. To obtain the $log\ (HR)$ and variance for individuals with generalised seizures, $\hat{\beta}_{2j}$ and $\mathrm{var}(\hat{\beta}_{2j})$:

$$\hat{\beta}_{2j} \ = \hat{\alpha}_{2j} + \hat{\gamma}_{2j} \qquad \text{(Equation 32)}$$

$$\mathrm{var}\big(\hat{\beta}_{2j}\big) = \mathrm{var}\big(\hat{\alpha}_{2j}\big) + \mathrm{var}\big(\hat{\gamma}_{2j}\big) + 2\ covar(\hat{\alpha}_{2j}, \hat{\gamma}_{2j}) \qquad \text{(Equation 33)}$$

Where $covar(\hat{\alpha}_{2j}, \hat{\gamma}_{2j})$ is the covariance between treatment effect $\hat{\alpha}_{2j}$ and interaction effect $\hat{\gamma}_{2j}$. As described above in Chapter 6.2.2, the model described in Equation 31 is fit by running the 'stcox' command in Stata via 'mvmeta_make' command to produce a dataset in correct format for 'mvmeta' (see Chapter 6.2.5 for further details).

This approach was pre-specified in the protocol of Cochrane IPD-NMA, which was written in 2014 [68] and detailed results of the Cochrane IPD-NMA [69], including investigation of heterogeneity and inconsistency, sensitivity analyses and clinical interpretation of the IPD-NMA results are presented in Chapter 7 of this thesis.

Recent work has highlighted the importance of the appropriate specification of one-stage IPD models with treatment-covariate interactions by separating within-study and across-study interactions to avoid inadvertent ecological bias [25, 38, 41, 185], particularly within a TTE setting [25]. The most recent work in this area has illustrated this within the context of

epilepsy, showing that the magnitude and statistical significance of age as a treatment effect modifier is reduced when within-trial and across-trial interactions are separated in IPD-MA compared to the original IPD-MA where they are amalgamated [25]. Furthermore, the importance of separating within-study and across-study associations increases as the variability of the covariate value across the included studies increases [25, 41, 46]. In the present example, some studies recruit only individuals with partial seizures, some recruit only individuals with generalised seizures and the proportion of individuals with partial seizures within the studies designed to recruit both seizure types ranges from 18% to 86%, therefore, there is a risk of ecological bias in the approach presented above.

An additional analysis was performed following the completion of main analysis for the Cochrane Review [69]. The additional analysis separates the within and across-study interactions in the Cox PH model (specified in Equation 31) by centering the treatment-by-epilepsy type interaction:

$$h_{ij}(t) = h_{0j}(t)exp((\alpha_{2j}x_{2ij} + \ \ldots + \alpha_{10j}x_{10ij}) + (z_{ij} - \bar{z}_j)(\pi_j + \gamma_{2j}x_{2ij} + \ \ldots \ + \ \gamma_{9j}x_{9ij}))$$

(Equation 34)

Where $\bar{z}_j$ is the proportion of individuals with generalised seizures within each trial and other parameters are defined as in Equation 31.

6.2.4 First stage: IPD-NMA via the 'meta-analysis of interactions' approach

This approach is based on the 'meta-analysis of interactions' approach by Simmonds and Higgins [309]. In this approach, rather than a 'complete' IPD analysis, IPD is reduced to summary statistics with a treatment-by-epilepsy type interaction.

In other words, a model of the structure outlined in Equation 31 (with a treatment-by-epilepsy type interaction) is fitted separately to the IPD of each trial, producing separate summary statistics of treatment effect for individuals with partial seizures $\alpha_{2j} \ \ldots \alpha_{10j}$ and individuals with generalised seizures $\beta_{2j} \ldots \beta_{9j}$ (calculated as described above in Equation 32 and Equation 33).

The summary statistics for each epilepsy type from each trial are then combined in separate NMAs by epilepsy type as if they were AD. This was achieved by producing a dataset of the summary statistics structured as a list of pairwise comparisons and converted from 'pairs' to 'augmented' format via the 'network' command within Stata version 14 [310] (see Appendix 13) and NMA is performed via 'mvmeta' as described in Chapter 6.2.5.

6.2.5 Second stage: IPD-NMA methods

NMA was performed under a multivariate meta-analysis framework, where the pairwise treatment comparisons are treated as different outcomes and NMA is performed via multivariate meta-regression techniques [220].

Assume that each trial provides p treatment effects of interest in relation to a reference treatment. This framework will be demonstrated assuming a simple example of a three-armed trial which randomises participants of both seizure types to the reference treatment CBZ and two other AEDs (say drug 2 and drug 3). Using the notation outlined in Chapter 6.2.2 and Chapter 6.2.3, in the j^{th} trial with i participants, with reference CBZ (constrained to be 0) and two other AEDs defined by indicator variables x_{2ij} and x_{3ij}, then the j^{th} trial will provide four effect estimates of interest (i.e. here $p = 4$): α_{2j} and α_{3j} for individuals with partial seizures and β_{2j} and β_{3j} for individuals with generalised seizures. For trial j, the p effects of interest can be written as a (1 x p) vector $\boldsymbol{y_j}$ and the within-trial variance-covariance (p x p) matrix can be written as $\boldsymbol{S_j}$. So in this example:

$$\boldsymbol{y_j} = \begin{pmatrix} \alpha_{2j} \\ \alpha_{3j} \\ \beta_{2j} \\ \beta_{3j} \end{pmatrix} \quad \text{(Equation 35)}$$

$$\boldsymbol{S_j} = \begin{pmatrix} var(\alpha_{2j}) & cov(\alpha_{2j},\alpha_{3j}) & cov(\alpha_{2j},\beta_{2j}) & cov(\alpha_{2j},\beta_{3j}) \\ cov(\alpha_{3j},\alpha_{2j}) & var(\alpha_{3j}) & cov(\alpha_{3j},\beta_{2j}) & cov(\alpha_{3j},\beta_{3j}) \\ cov(\beta_{2j},\alpha_{2j}) & cov(\beta_{2j},\alpha_{3j}) & var(\beta_{2j}) & cov(\beta_{2j},\beta_{3j}) \\ cov(\beta_{3j},\alpha_{2j}) & cov(\beta_{3j},\alpha_{3j}) & cov(\beta_{3j},\beta_{2j}) & var(\beta_{3j)} \end{pmatrix} \quad \text{(Equation 36)}$$

The aim of the multivariate analysis is to estimate the 'network parameters' which are unknown i.e. in this case the treatment effect of drug 2 in relation to CBZ and the drug 3 in relation to CBZ, for individuals with partial seizures and for individuals with generalised seizures across all trials and across the whole network of AEDs. These 'network parameters' can be written as (p x 1) matrix $\boldsymbol{\mu}$**.** So in this example:

$$\boldsymbol{\mu} = (\alpha_2, \alpha_3, \beta_2, \beta_3) \quad \text{(Equation 37)}$$

The general form of multivariate random-effects meta-regression for j trials is [220]:

$$\boldsymbol{y_j} \sim N(\boldsymbol{\mu}\, \boldsymbol{X_j}, \Sigma + S_j) \quad \text{(Equation 38)}$$

Where $\boldsymbol{y_j}$, $\boldsymbol{\mu}$ and $\boldsymbol{S_j}$ are defined as above for the p treatment effects of interest in the j^{th} trial, $\boldsymbol{X_j}$ is a (q x p) matrix with all elements equal to 0 or 1 to indicate which treatment effects are present in each trial (up to a maximum of q 'network parameters' to be estimated) and Σ is a ($q \times q$) matrix of between-trials variance-covariance which is assumed to be the same for all trials. Various assumptions can be made about the structure of this variance-covariance matrix; see White *et al* [220] for further discussion. Multivariate meta-regression under this framework can also be performed using fixed-effects, with Σ assumed to be a zero matrix (i.e. no heterogeneity).

Therefore for the two separate models for individuals of each seizure type described in Chapter 6.2.2, for individuals with partial seizures, the network parameters to be estimated are $\boldsymbol{\mu} = (\alpha_2 \ \ldots, \alpha_{10})$ (i.e. q =9) and for individuals with generalised seizures the network parameters to be estimated are $\boldsymbol{\mu} = (\beta_2 \ldots \beta_9)$ (i.e. q =8). Similarly, for the approaches described in Chapter 6.2.3 and Chapter 6.2.4 where seizure types are considered within the same model via a treatment-by-epilepsy type interaction, the network parameters to be estimated are $\boldsymbol{\mu} = (\alpha_2 \ldots \alpha_{10}, \beta_2 \ldots \beta_9)$, (i.e. $q = 17$).

Both fixed-effects and random-effects models were fitted via 'mvmeta' command in Stata using the default estimation method of restricted maximum likelihood estimation (REML) [220, 308]. Random-effects models were fitted assuming equal heterogeneity for all pairwise comparisons (i.e. between-trial-covariance structure (variance-covariance matrix Σ) proportional to single unknown heterogeneity parameter τ^2 rather than allowing for heterogeneity to vary across the comparisons). It was necessary to make an assumption regarding the between-trial covariance structure as not all possible pairwise comparisons between the treatments of interest were present in the network (see Figure 9). This multivariate meta-regression approach requires that all studies report on a common reference treatment (in this example, CBZ or VPS depending on the 'first-stage' model approach). For studies that do not include this reference treatment, to allow this approach to be used a minimally informative reference arm (i.e. a very small treatment effect and variance) can be imputed via a data augmentation technique describe by White *et al* [308].

Due to the assumption made of 'proportional' heterogeneity, an I^2 statistic cannot be directly calculated for the NMA. Alternatively, an R statistic can be estimated directly via the 'randfix' option of 'mvmeta' which is a ratio measure of the standard errors in the random-effects NMA model compared to the fixed-effects NMA model [311], and it has been shown that R can be used to calculate I^2 as follows [312]:

$$I^2 = (R^2 - 1)/R^2 \quad \text{(Equation 39)}$$

An I^2 statistic was estimated for the whole treatment network for each NMA and the estimate of τ^2 from the random-effects NMA model was also used in interpreting the presence of any heterogeneity in the treatment network (with higher I^2 and τ^2 values indicating more heterogeneity present). It should also be noted that the R statistic and therefore this I^2 statistic are dependent on the number of parameters in the NMA model. Therefore when seizure types are considered in separate NMA models (Chapter 6.2.2), these models have fewer network parameters to estimate than the models which consider the two seizure types via an interaction term (Chapter 6.2.3 and Chapter 6.2.4) so the models with the interaction terms are likely to have higher R and I^2 values which may be due to the number of parameters rather than necessarily due to increased variability in the model. For this reason, direct numerical comparisons of the τ^2, R and I^2 statistics across the NMA models was not made and any relative comparisons were made in the context of the number of parameters included in each model.

The estimated network parameters (matrix $\boldsymbol{\mu}$ for each model) represent the relative treatment effects for each AED compared to reference treatment CBZ. To obtain a complete set of pairwise comparisons for the treatments, some algebraic manipulation was required in a similar manner to the calculations described above in Equation 32 and Equation 33.

For example, let $\hat{\alpha}_2$ and $\mathrm{var}(\hat{\alpha}_2)$ be the 'network parameter' estimates of $log\ (HR)$ and variance for drug 2 compared to CBZ (for individuals with partial seizures) and let $\hat{\alpha}_3$ and $\mathrm{var}(\hat{\alpha}_3)$ be the 'network parameter' estimates of $log\ (HR)$ and variance for drug 3 compared to CBZ. To obtain the 'network parameter' estimates of $log\ (HR)$ and variance for drug 3 compared to drug 2, $\hat{\alpha}_{3vs.2}$ and $\mathrm{var}(\hat{\alpha}_{3vs.2})$:

$$\hat{\alpha}_{3vs.2} = \hat{\alpha}_3 - \hat{\alpha}_2 \quad \text{(Equation 40)}$$

$$\mathrm{var}(\hat{\alpha}_{3vs.2}) = \mathrm{var}(\hat{\alpha}_2) + \mathrm{var}(\hat{\alpha}_3) + 2\ covar(\hat{\alpha}_2, \hat{\alpha}_3) \quad \text{(Equation 41)}$$

Where $covar(\hat{\alpha}_2, \hat{\alpha}_3)$ is the covariance between treatment effects $\hat{\alpha}_{2j}$ and $\hat{\alpha}_3$. Other pairwise comparisons for each epilepsy type were calculated in a similar way.

6.3 Results

The main focus of this Chapter is the consideration of different methodological approaches to modelling epilepsy type in this IPD-NMA of AED monotherapy. Therefore, for brevity, results for primary outcome 'Time-to-withdrawal of allocated treatment' only are reported in this Chapter. Further details of secondary outcomes, additional consideration of inconsistency and heterogeneity and clinical implications are presented in Chapter 7.

6.3.1. IPD-NMA results according to the approach for modelling epilepsy type

Table 15 and Table 16 show IPD-NMA results for individuals with partial seizures and with generalised seizures respectively according to the approach for modelling epilepsy type; see Chapter 6.2 for further details of all methods.

Table 15: Fixed-effects IPD-NMA results for time-to-withdrawal of allocated treatment according to approach of modelling epilepsy type (individuals with partial seizures)

Comparison[1,2]	Model for partial seizures only[3]	Model with amalgamated interaction term[4]	Model with separated interaction term[4]	Meta-analysis of interactions model[5]
CBZ vs PHB	**1.57 (1.20 to 2.05)**	**1.55 (1.18 to 2.04)**	**1.56 (1.19 to 2.04)**	**1.58 (1.14 to 2.18)**
CBZ vs PHT	1.16 (0.93 to 1.45)	1.13 (0.92 to 1.38)	1.12 (0.92 to 1.36)	1.20 (0.89 to 1.62)
CBZ vs VPS	1.10 (0.90 to 1.35)	1.04 (0.86 to 1.25)	0.90 (0.76 to 1.07)	0.90 (0.70 to 1.17)
CBZ vs LTG	**0.72 (0.63 to 0.83)**	**0.75 (0.65 to 0.86)**	**0.73 (0.64 to 0.85)**	**0.78 (0.61 to 0.98)**
CBZ vs OXC	1.07 (0.84 to 1.37)	1.09 (0.84 to 1.42)	1.09 (0.80 to 1.47)	1.34 (0.48 to 3.72)
CBZ vs TPM	1.17 (0.99 to 1.38)	1.18 (0.98 to 1.43)	1.17 (0.94 to 1.47)	1.01 (0.50 to 2.05)
CBZ vs GBP	**1.18 (1.01 to 1.39)**	**1.20 (1.00 to 1.43)**	1.18 (0.95 to 1.47)	1.12 (0.83 to 1.53)
CBZ vs LEV	**0.83 (0.70 to 0.99)**	**0.82 (0.69 to 0.97)**	0.85 (0.70 to 1.02)	**0.69 (0.51 to 0.95)**
CBZ vs ZNS	1.08 (0.81 to 1.44)	1.08 (0.79 to 1.48)	1.08 (0.75 to 1.55)	1.08 (0.81 to 1.44)
PHB vs PHT	0.74 (0.49 to 1.11)	**0.73 (0.55 to 0.96)**	**0.72 (0.54 to 0.95)**	0.76 (0.44 to 1.31)
PHB vs VPS	**0.70 (0.50 to 0.99)**	**0.67 (0.48 to 0.92)**	**0.58 (0.43 to 0.78)**	**0.57 (0.37 to 0.88)**
PHB vs LTG	**0.46 (0.34 to 0.62)**	**0.48 (0.35 to 0.66)**	**0.47 (0.35 to 0.64)**	**0.49 (0.33 to 0.73)**
PHB vs OXC	**0.68 (0.48 to 0.98)**	0.70 (0.48 to 1.03)	0.70 (0.47 to 1.04)	0.85 (0.28 to 2.54)
PHB vs TPM	0.75 (0.54 to 1.02)	0.76 (0.55 to 1.06)	0.75 (0.53 to 1.07)	0.64 (0.29 to 1.41)
PHB vs GBP	0.76 (0.55 to 1.03)	0.77 (0.55 to 1.07)	0.76 (0.54 to 1.07)	0.71 (0.46 to 1.11)
PHB vs LEV	**0.53 (0.39 to 0.73)**	**0.53 (0.38 to 0.73)**	**0.54 (0.39 to 0.75)**	**0.44 (0.28 to 0.69)**
PHB vs ZNS	0.69 (0.47 to 1.02)	0.70 (0.46 to 1.06)	0.69 (0.44 to 1.09)	0.69 (0.45 to 1.05)
PHT vs VPS	0.95 (0.70 to 1.29)	0.92 (0.70 to 1.21)	0.81 (0.62 to 1.04)	0.76 (0.49 to 1.16)
PHT vs LTG	**0.62 (0.48 to 0.80)**	**0.66 (0.52 to 0.85)**	**0.66 (0.51 to 0.83)**	**0.65 (0.44 to 0.95)**
PHT vs OXC	0.92 (0.67 to 1.28)	0.97 (0.69 to 1.35)	0.97 (0.68 to 1.39)	1.12 (0.37 to 3.40)
PHT vs TPM	1.01 (0.77 to 1.32)	1.05 (0.80 to 1.39)	1.05 (0.78 to 1.41)	0.85 (0.38 to 1.87)
PHT vs GBP	1.02 (0.78 to 1.34)	1.06 (0.81 to 1.40)	1.05 (0.79 to 1.41)	0.94 (0.61 to 1.45)
PHT vs LEV	**0.72 (0.54 to 0.95)**	**0.73 (0.56 to 0.95)**	**0.76 (0.58 to 0.99)**	**0.58 (0.37 to 0.90)**

PHT vs ZNS	0.93 (0.65 to 1.34)	0.96 (0.66 to 1.39)	0.97 (0.64 to 1.46)	0.90 (0.60 to 1.37)
VPS vs LTG	**0.65 (0.51 to 0.84)**	**0.72 (0.58 to 0.90)**	0.81 (0.66 to 1.01)	0.86 (0.60 to 1.22)
VPS vs OXC	0.97 (0.71 to 1.34)	1.05 (0.76 to 1.44)	1.21 (0.86 to 1.70)	1.48 (0.51 to 4.28)
VPS vs TPM	1.06 (0.81 to 1.38)	1.14 (0.88 to 1.48)	1.30 (0.98 to 1.72)	1.12 (0.51 to 2.45)
VPS vs GBP	1.07 (0.83 to 1.39)	1.15 (0.89 to 1.49)	1.31 (0.99 to 1.73)	1.24 (0.83 to 1.85)
VPS vs LEV	0.76 (0.58 to 0.98)	0.79 (0.61 to 1.03)	0.94 (0.73 to 1.21)	0.77 (0.51 to 1.16)
VPS vs ZNS	0.98 (0.69 to 1.39)	1.04 (0.73 to 1.50)	1.20 (0.80 to 1.79)	1.20 (0.82 to 1.76)
LTG vs OXC	**1.49 (1.10 to 2.02)**	**1.46 (1.11 to 1.92)**	**1.48 (1.08 to 2.03)**	1.73 (0.60 to 4.95)
LTG vs TPM	**1.62 (1.27 to 2.06)**	**1.59 (1.29 to 1.95)**	**1.60 (1.26 to 2.03)**	1.31 (0.61 to 2.78)
LTG vs GBP	**1.64 (1.29 to 2.09)**	**1.60 (1.31 to 1.96)**	**1.61 (1.27 to 2.04)**	1.45 (0.99 to 2.13)
LTG vs LEV	1.16 (0.92 to 1.45)	1.10 (0.89 to 1.35)	1.15 (0.92 to 1.44)	0.89 (0.58 to 1.38)
LTG vs ZNS	**1.50 (1.09 to 2.07)**	**1.45 (1.03 to 2.04)**	**1.47 (1.00 to 2.18)**	1.40 (0.97 to 2.02)
OXC vs TPM	1.09 (0.78 to 1.51)	1.09 (0.82 to 1.44)	1.08 (0.78 to 1.49)	0.76 (0.21 to 2.74)
OXC vs GBP	1.10 (0.80 to 1.53)	1.10 (0.83 to 1.45)	1.09 (0.78 to 1.50)	0.84 (0.29 to 2.44)
OXC vs LEV	0.78 (0.58 to 1.05)	0.75 (0.55 to 1.03)	0.78 (0.55 to 1.11)	0.52 (0.18 to 1.51)
OXC vs ZNS	1.01 (0.69 to 1.47)	0.99 (0.66 to 1.49)	0.99 (0.62 to 1.59)	0.81 (0.28 to 2.33)
TPM vs GBP	1.01 (0.78 to 1.32)	1.01 (0.82 to 1.25)	1.01 (0.77 to 1.31)	1.11 (0.51 to 2.41)
TPM vs LEV	**0.71 (0.56 to 0.90)**	**0.69 (0.54 to 0.89)**	**0.72 (0.54 to 0.96)**	0.69 (0.31 to 1.49)
TPM vs ZNS	0.93 (0.67 to 1.29)	0.91 (0.64 to 1.31)	0.92 (0.60 to 1.41)	1.07 (0.50 to 2.29)
GBP vs LEV	**0.70 (0.56 to 0.89)**	**0.69 (0.54 to 0.88)**	**0.72 (0.54 to 0.95)**	**0.62 (0.40 to 0.96)**
GBP vs ZNS	0.91 (0.66 to 1.27)	0.90 (0.63 to 1.30)	0.92 (0.60 to 1.40)	0.96 (0.63 to 1.47)
LEV vs ZNS	1.30 (0.93 to 1.81)	1.32 (0.93 to 1.88)	1.28 (0.85 to 1.92)	**1.56 (1.02 to 2.39)**
τ^2 statistic	7×10^{-21}	0.0037	7×10^{-13}	2×10^{-20}
R statistic	1.000	1.064	1.131	1.000
I^2 statistic	0%	11.7%	21.8%	0%

1. Order of drugs in the table: most commonly used drug first (CBZ), then drugs are ordered approximately by the date they were licenced as a monotherapy treatment (oldest first).
2. HRs and 95% CIs are calculated from fixed-effects analyses. HR < 1 indicates an advantage to the second drug in the comparison; results in highlighted in bold are statistically significant.
3. Results taken from the model outlined in Chapter 6.2.2
4. Results taken from the models outlined in Chapter 6.2.3
5. Results taken from the model outlined in Chapter 6.2.4

For individuals with partial seizures for the majority of comparisons, the conclusions that would be drawn in terms of statistical significance are consistent across the four modelling approaches (i.e. either all statistically significant or non-significant). For eleven comparisons, there is variability across the modelling approaches in terms of the conclusions that would be drawn based on statistical significance, most notably for the important comparisons of CBZ vs LEV and VPS vs LTG (see Chapter 7 for further discussion of important AEDs); although the general direction of effect is consistent for all of the eleven comparisons.

Notably, when comparing the results of the two models described in Chapter 6.2.3 with the within-trial and across-trial interaction terms amalgamated or separated respectively, numerical results of these two models are mostly very similar (to one or two decimal places) and conclusions mostly the same. Also notably, for the comparisons of CBZ vs LEV and VPS vs LTG, the conclusions drawn in terms of statistical significance would be different, with the amalgamated interaction model showing a significant effect and the separated interactions model showing no statistically significant difference; although the direction of effect is the same for both comparisons across the two approaches and numerical results for the CBZ vs LEV comparison are the same to one decimal place.

In comparison to the other modelling approaches, the 'meta-analysis of interactions' approach seems to produce the most different results to the other methods, particularly in terms of the precision of the results but also in terms of the numerical results (e.g. see comparisons of LTG vs OXC, TPM, GBP and ZNS in Table 15). The differences in results from this approach may have occurred for two reasons. Firstly, other approaches analyse IPD only in the 'first-stage' and this is the only method which reduces the IPD to aggregate data for each trial in the 'first-stage' and then performs the NMA as if the data was aggregate data. This aspect of the approach could have resulted in the loss of precision observed in the results in Table 15.

Table 16: Fixed-effects IPD-NMA results for time-to-withdrawal of allocated treatment according to approach of modelling epilepsy type (individuals with generalised seizures)

Comparison[1,2]	Model for generalised seizures only[3]	Model with amalgamated interaction term[4]	Model with separated interaction term[4]	Meta-analysis of interactions model[5]
CBZ vs PHB	1.63 (0.86 to 3.08)	1.47 (0.83 to 2.61)	1.27 (0.55 to 2.92)	0.76 (0.25 to 2.28)
CBZ vs PHT	1.05 (0.73 to 1.52)	0.92 (0.59 to 1.42)	0.97 (0.54 to 1.75)	0.73 (0.32 to 1.63)
CBZ vs VPS	0.83 (0.62 to 1.12)	**0.70 (0.54 to 0.92)**	0.70 (0.45 to 1.08)	0.63 (0.36 to 1.09)
CBZ vs LTG	0.69 (0.48 to 1.01)	**0.63 (0.45 to 0.89)**	0.75 (0.47 to 1.19)	0.89 (0.49 to 1.61)
CBZ vs OXC	0.83 (0.16 to 4.43)	1.00 (0.21 to 4.81)	0.90 (0.17 to 4.66)	0.73 (0.20 to 2.59)
CBZ vs TPM	0.90 (0.42 to 1.93)	1.24 (0.90 to 1.71)	0.83 (0.42 to 1.63)	0.82 (0.40 to 1.68)
CBZ vs GBP	0.92 (0.11 to 8.07)	0.90 (0.11 to 7.29)	0.76 (0.09 to 6.55)	0.65 (0.03 to 12.6)
CBZ vs LEV	0.78 (0.43 to 1.42)	0.74 (0.44 to 1.23)	0.80 (0.48 to 1.32)	0.84 (0.45 to 1.58)
PHB vs PHT	0.65 (0.30 to 1.39)	0.62 (0.32 to 1.24)	0.77 (0.30 to 1.97)	0.95 (0.22 to 4.20)
PHB vs VPS	0.51 (0.24 to 1.09)	**0.48 (0.27 to 0.86)**	0.55 (0.23 to 1.33)	0.83 (0.22 to 3.11)
PHB vs LTG	**0.43 (0.20 to 0.89)**	**0.43 (0.22 to 0.83)**	0.59 (0.23 to 1.52)	1.16 (0.32 to 4.28)
PHB vs OXC	0.51 (0.09 to 3.06)	0.68 (0.13 to 3.60)	0.71 (0.11 to 4.48)	0.95 (0.16 to 5.57)
PHB vs TPM	0.55 (0.20 to 1.49)	0.84 (0.44 to 1.60)	0.66 (0.22 to 1.92)	1.08 (0.27 to 4.25)
PHB vs GBP	0.57 (0.06 to 5.43)	0.61 (0.07 to 5.34)	0.60 (0.06 to 6.03)	0.86 (0.04 to 20.7)
PHB vs LEV	0.48 (0.20 to 1.15)	0.50 (0.23 to 1.09)	0.63 (0.24 to 1.67)	1.11 (0.31 to 4.03)

PHT vs VPS	0.79 (0.49 to 1.28)	0.77 (0.46 to 1.27)	0.72 (0.36 to 1.45)	0.87 (0.28 to 2.69)
PHT vs LTG	0.66 (0.39 to 1.12)	0.69 (0.39 to 1.20)	0.77 (0.36 to 1.62)	1.22 (0.38 to 3.90)
PHT vs OXC	0.79 (0.14 to 4.38)	1.09 (0.21 to 5.56)	0.92 (0.16 to 5.31)	1.00 (0.16 to 6.23)
PHT vs TPM	0.85 (0.36 to 1.99)	1.35 (0.79 to 2.30)	0.85 (0.35 to 2.10)	1.13 (0.33 to 3.92)
PHT vs GBP	0.88 (0.10 to 7.91)	0.98 (0.12 to 8.30)	0.78 (0.08 to 7.29)	0.90 (0.04 to 20.9)
PHT vs LEV	0.74 (0.37 to 1.49)	0.80 (0.42 to 1.55)	0.82 (0.38 to 1.77)	1.16 (0.40 to 3.37)
VPS vs LTG	0.83 (0.51 to 1.36)	0.90 (0.60 to 1.35)	1.07 (0.59 to 1.93)	1.41 (0.55 to 3.62)
VPS vs OXC	1.00 (0.18 to 5.46)	1.42 (0.29 to 6.92)	1.28 (0.24 to 7.01)	1.15 (0.26 to 5.13)
VPS vs TPM	1.08 (0.47 to 2.44)	**1.76 (1.22 to 2.53)**	1.19 (0.53 to 2.65)	1.30 (0.45 to 3.76)
VPS vs GBP	1.11 (0.12 to 9.89)	1.28 (0.16 to 10.5)	1.09 (0.12 to 9.74)	1.04 (0.05 to 22.0)
VPS vs LEV	0.93 (0.48 to 1.82)	1.05 (0.58 to 1.90)	1.14 (0.58 to 2.23)	1.34 (0.53 to 3.39)
LTG vs OXC	1.20 (0.21 to 6.91)	1.58 (0.33 to 7.67)	1.20 (0.23 to 6.27)	0.82 (0.18 to 3.77)
LTG vs TPM	1.29 (0.53 to 3.13)	**1.96 (1.25 to 3.08)**	1.11 (0.51 to 2.41)	0.93 (0.32 to 2.65)
LTG vs GBP	1.33 (0.14 to 12.4)	1.42 (0.17 to 11.6)	1.02 (0.12 to 8.81)	0.74 (0.03 to 16.1)
LTG vs LEV	1.12 (0.55 to 2.28)	1.17 (0.63 to 2.19)	1.07 (0.54 to 2.11)	0.95 (0.39 to 2.35)
OXC vs TPM	1.07 (0.15 to 7.86)	1.24 (0.26 to 5.94)	0.93 (0.18 to 4.80)	1.13 (0.23 to 5.61)
OXC vs GBP	1.11 (0.06 to 21.8)	0.90 (0.08 to 9.96)	0.85 (0.07 to 9.91)	0.90 (0.03 to 26.0)
OXC vs LEV	0.93 (0.16 to 5.49)	0.74 (0.14 to 3.86)	0.89 (0.16 to 4.99)	1.16 (0.27 to 4.95)
TPM vs GBP	1.03 (0.09 to 11.6)	0.73 (0.09 to 5.89)	0.92 (0.11 to 7.90)	0.8 (0.03 to 18.36)
TPM vs LEV	0.87 (0.33 to 2.29)	0.60 (0.33 to 1.09)	0.96 (0.41 to 2.23)	1.03 (0.38 to 2.78)
GBP vs LEV	0.84 (0.09 to 7.97)	0.82 (0.10 to 7.10)	1.05 (0.11 to 9.58)	1.29 (0.06 to 27.0)
τ^2 statistic	0.0701	0.0037	7×10^{-13}	2×10^{-20}
R statistic	1.225	1.064	1.131	1.000
I^2 statistic	33.3%	11.7%	21.8%	0%

1. See Table 15 for abbreviations and other footnote labels

Secondly, the largest differences in the results of this approach compared to the other modelling approaches seem to occur within the pairwise comparisons where no direct evidence exists or only very limited direct evidence from a single study contributes to the network (see Figure 12 and Chapter 7.2 for further discussion of direct evidence contributed for each comparison). The 'meta-analysis of interactions' approach fits only within-trial interactions and does not take account of across-trial interactions. The differences in the results across the modelling approaches suggest that any association between treatment effect and epilepsy type may be different within the trials compared to across the trials; in other words that ecological bias may be present within the model with an amalgamated interaction. If this is the case, then an approach to NMA which does not allow for across-trial relationships is likely to produce different results, particularly in this example where within-trial information is quite limited or non-existent for some comparisons.

Table 16 shows that statistically significant differences between pairs of AEDs are rarely found across any of the methodological approaches for individuals with generalised seizures; in fact no differences between the drugs in any of the pairwise comparisons is found from the model with separated within-trial and across-trial interactions, or from the 'meta-analysis of interactions' model. There is also more variability in the numerical results across methods for individuals with generalised seizures and more changes in the statistical significance and direct of effect of numerical results. As for individuals with partial seizures, the 'meta-analysis of interactions' approach seems to produce the most numerically different results to the other methods. Overall, it is not surprising that a different approach to modelling the relationship between treatment and epilepsy type has had less impact on the numerical results and conclusions for individuals with partial seizures (the majority epilepsy type, around 70%) but a larger impact on the numerical results and conclusions for individuals with generalised seizures (the minority epilepsy type, around 25%).

It is well documented that VPS is the most effective drug for controlling generalised seizures [69, 74]. However this AED is not suitable for all individuals, particularly females of childbearing age due to the potential teratogenic effects of the drug [74, 75]. Therefore interest lies in identifying AEDs which are suitable alternatives for VPS, rather than AEDs which are significantly better than VPS. This is starting to be reflected within clinical practice, with recent trials aiming to demonstrate non-inferiority of new AEDs compared to 'standard' treatments, rather than superiority [261, 262]. In terms of the important comparisons shown in Table 16 (see Chapter 7 for further discussion), consistently across all of the methods, despite some numerical differences in results and general direction of effect, no statistically significant differences are shown for the comparisons VPS vs LTG and VPS vs LEV, suggesting that LTG and LEV do not seem to be any worse than VPS in terms of retention which may make these two treatments potential alternatives to VPS.

Another notable difference is that the conclusion that may be drawn for the comparison of VPS over CBZ for individuals with generalised seizures in terms of statistical significance differs according to the methodological approach. Despite this, the direction of effect for the comparison of VPS and CBZ is consistent across all approaches and the statistical significance (or lack thereof) of this comparison is unlikely to impact on clinical practice due to documented evidence that CBZ may exacerbate some generalised seizure types [287, 288].

6.3.2 Additional consideration of treatment-covariate interaction

Hua *et al* [25] note the risk of an interaction being missed or incorrectly interpreted where within-study and across-study interactions are amalgamated. While this is an important and valid concern when performing analysis to establish the existence of treatment-covariate interactions, within this context, the aim was not to test for a statistical interaction. Strong clinical evidence and current guidelines [74] assume that a clinical interaction between treatment and epilepsy type exists, hence within current clinical practice clinicians do tend to have a preference for specific drugs for different seizure types (see Chapter 1.2 and Chapter 6.1 for further details). Therefore for this IPD-NMA to be useful to future clinical practice, it was essential to provide results by seizure type and missing an 'interaction' was not a specifically a concern to this NMA.

However, previous Cochrane IPD-MAs (e.g. [59, 67]) as well as the analysis of Hua *et al* [25] have shown mixed results regarding whether a statistical interaction between treatment and epilepsy type actually exists. It is possible that misclassification of seizure type has confounded the results of previous Cochrane IPD-MAs, with one of the Cochrane IPD-MAs showing a significant interaction between treatment and epilepsy type following reclassification [59]; see Chapter 7.1.3 for further details of additional analyses to account for misclassification of seizure type in this IPD-NMA.

Within this example, due to the 'two-stage' nature of the methodology (see Chapter 6.2), statistical interactions could only be tested at the 'first-stage' of the modelling approach (see Chapter 6.2.3). A further analysis was performed for 'Time-to-withdrawal of allocated treatment' comparing results for the interaction terms of one-stage IPD-MA models (with amalgamated or separated interactions, of the structure of the models outlined in Equation 31 and Equation 34) fitted to subsets of the IPD for the most important pairwise comparisons in this context; CBZ vs LTG, CBZ vs LEV, CBZ vs VPS and VPS vs LTG (see Chapter 7).

Results of these analyses are presented in Table 17. It should be noted that these analyses are considered exploratory and that across-study interactions (the $\hat{\beta}_A$ defined by Hua *et al* [25]) are not presented due to the computational complexity of estimating this parameter within the current framework.

For the comparisons of CBZ vs LTG, CBZ vs LEV, CBZ vs VPS, results presented in Table 17 show that the interaction term is not statistically significant (i.e. no evidence of a statistical interaction between treatment effect and epilepsy type for these particular AEDs).

Furthermore, the same conclusions regarding the interaction between treatment and epilepsy type would be drawn (based on the statistical significance), although the magnitude of the interaction appears numerically larger for CBZ vs LTG where interactions are amalgamated rather than separated (-0.433 compared to -0.313).

Table 17: One stage meta-analysis results for time-to-withdrawal of allocated treatment according to approach of modelling treatment-by-epilepsy type interaction

Comparison[1]	Parameter	Amalgamated interaction			Separated interactions[2]		
		β	SE	P-value	β	SE	P-value
CBZ vs LTG N=1,889 (8 trials)	Drug	-0.328	0.100	0.001	-0.402	0.091	<0.001
	Epilepsy type	0.248	0.174	0.154	0.195	0.191	0.308
	Interaction	-0.433	0.241	0.072	-0.313	0.274	0.253
CBZ vs VPS N=1,219 (5 trials)	Drug	0.020	0.120	0.869	-0.024	0.104	0.816
	Epilepsy type	-0.178	0.179	0.32	-0.127	0.195	0.515
	Interaction	-0.136	0.235	0.562	-0.238	0.279	0.395
CBZ vs LEV N=1,818 (3 trials)	Drug	-0.165	0.087	0.059	-0.181	0.082	0.028
	Epilepsy type	-0.127	0.179	0.478	-0.079	0.177	0.655
	Interaction	-0.105	0.252	0.676	-0.204	0.257	0.427
VPS vs LTG N=774 (3 trials)	Drug	-0.782	0.249	0.002	-0.028	0.138	0.84
	Epilepsy type	-0.392	0.234	0.094	-0.237	0.268	0.376
	Interaction	1.108	0.300	<0.001	0.719	0.393	0.067

Abbreviations: β = parameter estimate; CBZ = carbamazepine; LEV= Levetiracetam; LTG = lamotrigine; SE=standard error; VPS = sodium valproate

1. The first drug in the comparison is the reference treatment
2. For the separated interaction model, only the within-study interaction presented ($\hat{\beta}_W$ of Hua *et al* [25]) due to computational time of estimating across study interaction ($\hat{\beta}_A$ of Hua *et al* [25]) within Stata version 14.

For the comparison of VPS vs LTG, the interaction is highly significant from the amalgamated model (p<0.001) but the within-study interaction is not significant from the separated model (p=0.067). The magnitude of the interaction is also numerically reduced in the separated model compared to the amalgamated model (0.719 compared to 1.108). The differences across the modelling approaches shown for this comparison are in line with the results presented above for the NMA in Table 16.

6.3 Discussion

6.3.1 Summary of main results and implications

This Chapter presents several methodological approaches for an IPD-NMA of AED monotherapy, taking account of current clinical practice for treating individuals with partial onset and generalised onset seizures. Results from the different approaches presented in this Chapter demonstrate that for individuals with partial seizures (the majority epilepsy type, around 70%), numerical results and conclusions that could be drawn from them are fairly robust to the approach of modelling epilepsy type (within a separate model or via interaction terms). However, across methodological approaches for individuals with generalised seizures (the minority epilepsy type, around 25%) some numerical results change quite substantially, as well as some changes in the statistical conclusions.

Also of note, the 'meta-analysis of interactions approach' which reduces IPD to aggregate data and models the treatment-by-epilepsy type interaction separately within each trial (rather than also the interaction across-trials as the methods analysing the entire IPD dataset across all trials do) seems to produce the most numerically different results to the other methods, with these differences even more pronounced for individuals with generalised seizures where less data is available. This may be due to potential ecological bias originating from this approach.

It should be noted that definition of the primary outcomes of the IPD-NMA considered in this Chapter 'time-to-withdrawal of allocated treatment' is complex as multiple reasons for treatment withdrawal that are possible for each individual and how these reasons may be classified as events or censored observations compared to the definition of planned secondary outcomes of the IPD-NMA such as 'time-to-first seizure,' (i.e. whether an individual experiences a seizure during the study or not). Where different modelling approaches were applied to an IPD-NMA of 'time-to-first seizure,' numerical results again were similar and conclusions mostly unchanged for both seizure types (results not shown for brevity). For such complex outcomes, modelling of the association between treatment effect and epilepsy type on the individual patient level is even more important where outcome definitions may vary across trials (i.e. more withdrawal events in some trials may occur due to adverse events and due to lack of efficacy in other trials), whereas the definition of seizure recurrence is the same across all trials so the analysis of 'time-to-first seizure' is more robust to varying assumptions regarding the treatment-covariate relationships.

Further consideration of outcome definitions, particularly 'time-to-withdrawal of allocated treatment' which can be modelled under a competing risks framework and the impact on treatment-covariate relationships (including modelled as interactions) for both meta-analysis and NMA would be of interest for future research.

Result of additional exploratory analyses within this Chapter have also shown no evidence of a statistical interaction between treatment effect and epilepsy type for several commonly used AEDs which are assumed within current clinical practice to be more effective for particular seizure types [74]. This may indicate that such an interaction does not actually exist, despite the clinical perception of this interaction. Alternatively, it is possible that the existence of a statistical interaction has been not been identified due to lack of power or due to confounding from misclassification of epilepsy type, as was the case within a Cochrane IPD-MA of phenytoin (perceived to be a better treatment for partial seizures) compared to sodium valproate (perceived to be a better treatment for generalised seizures) [59].

Therefore to inform future clinical practice which assumes a 'clinical' interaction between AEDs and seizure type, it is essential that future trials reflect clinical practice and this 'interaction' within their designs; by recruiting participants with different seizure types separately (e.g. [264, 265]) or performing stratified randomisation and presenting results for each seizure type separately (e.g. [304, 305]). Accurate seizure classification (as far as possible) of individuals recruited into future trials is also of great importance to avoid confounding of any associations between specific AEDs and seizure types; further discussion of this implication can be found in Chapter 7.3.3.

6.3.2 Strengths and weaknesses

The use of IPD in these analyses allowed several approaches to consider the relationship between treatment effect and epilepsy type and allowed for results to be presented separately by epilepsy type in the context of the recommended first line treatment of the epilepsy type, such an approach which would not have been possible without the use of IPD.

At the time of planning this analysis during 2014 [68], there were few methodological publications for IPD-NMA, and no work has been published which had considered IPD-NMA of TTE data with treatment-covariate interactions (amalgamated interactions or separated into within and across-trial interactions). To my knowledge at the time of writing, there are still no published methods which allow such a 'one-stage' model to be fitted.

Hence the proposed methodology was adapted from a previous IPD-NMA in epilepsy conducted by the group [285] and a 'two-stage' approach was taken for all of the analyses presented within this Chapter. Recent work has outlined scenarios where one-stage and two-stage approaches to meta-analysis may produce different results depending on modelling assumptions and estimation methods, choice of fixed or random-effects, clustering within studies, correlation between model parameters and handling of treatment-covariate interactions [169]. It should be noted that two-stage approaches are methodologically accessible and generally extend to any type of synthesis or type of data [40].

Further investigation of whether any existing NMA methodology (e.g. [44, 158, 159]) can be extended or whether new methodology could be developed within the context of this IPD-NMA, within either the present multivariate framework or alternative frameworks (such as hierarchical Bayesian frameworks), would be of interest for further work.

A common approach for presenting NMA results is to present the 'ranks' of the treatments (i.e. the probability that each treatment in the network is the best)[313] which may be of interest to readers, particularly clinicians or participants facing a treatment choice. Within the present analysis of a complex and chronic condition (epilepsy) with multiple outcomes relating to efficacy and tolerability, including a complex composite primary outcome of 'time-to-withdrawal of allocated treatment' to which the individual participant can make a contribution, it was not deemed appropriate to present 'best' treatment for each outcome. For some individuals with epilepsy, complete remission of seizures may be a priority whereas for others intolerable side effects may be deemed more unacceptable than an occasional seizure so what would be deemed the 'best' treatment for one individual would not be the 'best' treatment for another. Further discussion of the clinical implications of the results and an informal 'ranking' by ordering according treatment effect sizes from the IPD-NMA compared to the reference treatment (e.g. better or worse than carbamazepine) and are presented graphically in Chapter 7 and in the Cochrane IPD-NMA [69].

It should be noted that for some of the present analysis approaches, if it had been deemed appropriate to present treatment 'ranks,' due to the treatment-by-epilepsy type interaction in this model, rankings separated by epilepsy type cannot be calculated directly. Not being able to present a clearly 'best' treatment from NMA results according to different covariate values (where appropriate) is a general limitation and future research which allows ranking treatments from models with treatment-covariate interactions would be of value.

6.3.3 Concluding remarks

Results across the statistical approaches outlined in this Chapter demonstrate that for individuals with partial seizures are fairly robust to the approach of modelling epilepsy type (within a separate model or via interaction terms) but results are more variable across the methodological approaches for individuals with generalised seizures.

Current clinical practice assumes that a clinical interaction between treatment effect and epilepsy type exists, with clinician preference for certain drugs and current NICE guidelines recommend different first-line treatments for individuals with different seizure types [74]. Despite this perception, the statistical approaches investigated within this Chapter do not provide consistent evidence that such an interaction exists statistically.

Nonetheless, it is essential for this IPD-NMA to be informative to future clinical practice of epilepsy monotherapy that separate inferences by seizure type can be made and that future trials of epilepsy monotherapy reflect clinical practice and this perceived 'interaction' within their designs. Such future designs would better allow for more accurate investigation of any true statistical interaction between treatment effect and epilepsy type as well as further consideration of any differences of interactions across-trials and within-trials.

Chapter 7: A Cochrane IPD-NMA of antiepileptic drug monotherapy: additional considerations

Chapter 6 of this thesis presents detailed statistical methodology for several approaches to modelling the association between epilepsy type and treatment effect in IPD-NMA. Methodology outlined in Chapter 6.2.3 was pre-specified in the protocol of Cochrane IPD-NMA, which was written in 2014 [68]. This Chapter presents the clinical results and implications of the Cochrane IPD-NMA [69] and additional methodological considerations including comparison of NMA results to direct evidence, investigation of heterogeneity and inconsistency, additional and sensitivity analyses.

7.1 Additional methodological considerations

7.1.1 Direct (pairwise) evidence

Where pairwise evidence was available (see Figure 9 in Chapter 5 for network diagram of comparisons made between the ten AEDs of interest), the j^{th} trial, a trial-specific $ln\ (HR)_j$ and $se(ln\ (HR))_j$ was estimated using methods outlined in Chapter 6.2.5 of this thesis. These trial specific estimates were then combined via the 'metan' command in Stata (inverse-variance meta-analysis, see Equation 13 and Equation 14 in Chapter 2.3.2), resulting in a pooled $ln\ (HR)$ and $var\ (ln\ (HR))$ for each available pairwise comparison.

For each pairwise comparison where data could be synthesised for at least two trials, the presence of statistical heterogeneity was assessed using the Q test (P-value < 0.10 for significance) and the I^2 statistic [314]. The presence of clinical heterogeneity was assessed by comparing trial design characteristics and participant demographics and by visual inspection of forest plots, particularly in terms of the magnitude and direction of effects. In the first instance, a fixed-effects model was to be used for all pairwise meta-analyses. If an important amount of heterogeneity (defined here as $I^2 > 50\%$) was found to be present which could not be explained by differences in trial or participant characteristics, pairwise meta-analyses would be repeated using a Der Simonanian-Laird random-effects model [12].

It should be noted that this is a 'two-stage' method of IPD-MA which was performed to allow visual comparison forest plots to assess for clinical heterogeneity. Such an approach may be associated with some limitations compared to one stage IPD-MA but is unlikely to produce different results to a one-stage method as both approaches have otherwise made the same assumptions [169]. For completeness, one-stage IPD-MA was also performed fitting a one-

stage stratified Cox PH model with fixed-treatment effects, treatment-by-epilepsy type interaction and stratified by trial for each pairwise comparison. Numerical results were very similar and conclusions unchanged for all pairwise comparisons (results not shown).

7.1.2 Investigation of consistency

As outlined in Chapter 1.1.2, the key assumption made in NMA is that average treatment effect is "exchangeable" across all included trials. Due to the clinical perception of a clinical (and potentially statistical) interaction between treatment effect and epilepsy type within this network, judgements of exchangeability were made separately by epilepsy type.

In the context of the present network, transitivity requires that all treatments are "jointly randomisable;" given that all of the ten drugs within this network are licenced and commonly used as monotherapy treatments for individuals with newly diagnosed partial-onset seizures or generalised-onset tonic-clonic seizures (with or without other generalised seizure types) and have all been used within trials of similar designs, there were no concerns over transitivity in this network.

The consistency assumption can be evaluated statistically comparing the difference between the direct treatment effect estimate and the indirect estimate for each loop of evidence. In general, methods for evaluating inconsistency take either a ‘local’ or ‘global’ approach; the former approach is loop specific, focusing on inconsistency within each specific treatment comparison while the latter approach evaluates the presence of inconsistency across the entire network of evidence [18]. A recent literature review of NMA methodology recommends the use of both ‘local’ and ‘global’ methods, where possible, to gain a better understanding of both the plausibility of the consistency assumption in the network as a whole and also the sources of possible discrepancies within the network [18].

Given the complexity of the network model fitted (with treatment-by-epilepsy type interaction) and the number of multi-arm trials included in analysis, two methods were used to evaluate inconsistency within this network. Firstly, node-splitting was performed via the ‘network sidesplit’ command in Stata [310, 315]; this ‘local’ method separates evidence from a particular comparison (node) into direct and indirect, allowing for formal comparison of difference between direct (pairwise), indirect and NMA (direct and indirect evidence combined) estimates. Secondly, a ‘design-by-treatment’ inconsistency model was fitted via the ‘network meta inconsistency’ command in Stata [310, 316]; a ‘global’ method which evaluates both loop and design inconsistencies, particularly within multi-arm trials. For

example, the A versus B treatment effect in a study comparing only treatments A and B (i.e. an AB design) may vary from the A versus B effect in a three arm study of an ABC design [18].

Here a 'design' was classified to reflect both the treatment comparisons made in the study and the epilepsy types recruited within the trial. For example, a trial of CBZ compared to LTG in individuals with partial-onset seizures only was considered to be a different design to a trial of CBZ compared to LTG recruiting individuals of both epilepsy types.

For investigation of inconsistency via node-splitting, treatment effect estimates are presented graphically for direct evidence (pairwise meta-analysis from the trials which make direct comparisons of the pair of drugs), indirect evidence (from the node-splitting model which makes indirect comparisons from the trials which do not make a direct comparison of the pair of drugs) and direct plus indirect evidence (from NMA model of the whole network) for each pairwise comparison. Numerical results for direct evidence, indirect evidence and NMA results for each pairwise comparison were examined, particularly the overlap of CIs of the estimate. It would be expected that numerical results for the NMA would be the most precise as the largest amount of data contributes to these analyses and due to the 'borrowing of strength' across the network. P-values and heterogeneity statistics are also presented from 'design-by-treatment' inconsistency models for each outcome.

Potentially important inconsistency was noted to be present if the global test for inconsistency from the 'design-by treatment' model was statistically significant and/or where CIs of results from direct evidence and NMA results do not overlap (i.e. there is a statistically significant difference between pairwise and NMA results). If deemed present, potential origins of inconsistency were investigated. Also, while of less concern than the potential inconsistency described above, it was also noted where CIs of results from indirect evidence did not overlap with the CIs of the direct and NMA estimates and reasons for any numerical differences in results were considered.

7.1.3 Additional analyses and sensitivity analyses

As outlined in Chapter 5.3.2, a range of participant covariates (including age, sex, seizure history, EEG and scan results) were requested in addition to outcome data in all IPD requests in order to explore these covariates as potential modifiers of treatment effect and as potential sources of heterogeneity and/or inconsistency. However, due to large proportions of missing data and variability in the definitions of data provided for most of these covariates (see Chapter 5.4.3 for further details), adjusted analyses with all covariates of interest was deemed not appropriate.

Age at randomisation was provided for the majority of participants (99% of total eligible participants) and previous Cochrane IPD-MAs have shown an association between age and treatment effect for commonly used drugs CBZ, VPS and LTG [61, 67]. Therefore, an additional analysis was performed adjusting for age at randomisation where an additional interaction term of treatment-by-age (centred by mean age of participants in each trial) was added to the initial Cox PH model (described in Chapter 6.2.3). NMA was repeated and results compared to those from the primary analysis (without age).

A range of sensitivity analyses were conducted to examine the robustness of results to assumptions made in pairwise and NMA:

- Trial-specific treatment effects were estimated via a Cox PH model (see Chapter 6.2.3). To assess the validity of the PH assumption, the statistical significance of time-varying covariates for all covariates in the model were tested. If there was indication that the PH assumption has been violated, in sensitivity analysis, for the i^{th} participant in the j^{th} trial, a parametric accelerated failure time (AFT) model, stratified by trial, was fitted to the entire individual participant dataset:

  $$h_{ij}(t) = h_{0j}(t/\exp(\varphi_{ij})exp(-\varphi_{ij}) \quad \text{(Equation 42)}$$

  Where $\varphi_{ij} = (\alpha_{2j}x_{2ij} + \ \ldots + \alpha_{10j}x_{10ij}) + z_{ij}\,(\pi_j + \gamma_{2j}x_{2ij} + \ \ldots + \ \gamma_{9j}x_{9ij})$

  Other parameters are defined as in Equation 31, outlined in Chapter 6.2.3. NMA was repeated according to the methods outlined in Chapter 6.2.5 and results were compared to those from the primary analysis with the Cox PH model.
- As outlined in Chapter 5.3.3, where minor inconsistencies remained in IPD provided following clarification from data providers, sensitivity analyses were conducted. Details of inconsistencies identified in IPD and sensitivity analyses conducted to account for these inconsistencies are presented in Appendix 14.
- Misclassification of epilepsy type is a recognised problem in epilepsy; whereby some individuals with partial seizures have been mistakenly classed as having generalised-onset seizures and vice versa. Such a misclassification has impacted on results of previous Cochrane IPD-MAs [66, 67]. Clinical evidence suggests that individuals with generalised-onset seizures are unlikely to have an 'age of onset' greater than 25 to 30 years [317]; out of the 12,371 participants with IPD provided for analysis, 1,164 (9%) were classified as experiencing generalised seizures and had an estimated age of onset as greater than 30 years. Two sensitivity analyses were performed to investigate the impact of potential misclassification:

- Re-classification of 1,164 individuals with generalised seizures and age of onset over 30 years as having partial-onset seizures. NMA was then repeated with the interaction term of treatment-by-epilepsy type with the reclassified epilepsy type.
- Re-classification of 1,164 individuals with generalised seizure and age at onset over 30 years and 574 participants with missing epilepsy type into an 'unclassified epilepsy type' group. NMA was then repeated with the interaction term of treatment-by-epilepsy type where epilepsy type is partial epilepsy compared to generalised or unclassified epilepsy.

It was not possible to achieve convergence of NMA with a 'three-way' interaction (i.e. partial epilepsy compared to generalised epilepsy compared to unclassified epilepsy). This is likely due to small numbers of participants with unclassified epilepsy and with generalised epilepsy following reclassification) receiving some AEDs.

7.2 Results

For brevity, results for 'Time-to-withdrawal of allocated treatment' and 'Time-to-first seizure' only are reported within this Chapter. Numerical results for the two remission outcomes can be found in Appendix 15 of this thesis and further discussion of the clinical implications of these results within the published Cochrane IPD-NMA [69].

All tables and figures of NMA results indicate the proportion of the treatment effect estimate which is contributed by direct evidence (ranging from 0% where no direct comparison exists to 100% for the CBZ vs ZNS comparison which is disconnected from the rest of the network, see Figure 12). These proportions were determined from a 'contribution plot' (derived via the 'netweight' command in Stata version 14), described by Chaimani *et al* [299]. It should be noted that due to a limited amount of evidence for individuals with generalised seizures for some comparisons in the network, some CIs of treatment effect sizes are very wide.

Table 18: Reasons for withdrawal from allocated treatment

Reason for withdrawal	Classification for analysis	Randomised Drug[3]										
		CBZ	**PHB**	**PHT**	**VPS**	**LTG**	**OXC**	**TPM**	**GBP**	**LEV**	**ZNS**	**Total**
Adverse events	Event	505 (45%)	24 (20%)	93 (35%)	132 (28%)	235 (41%)	56 (41%)	259 (48%)	73 (20%)	134 (39%)	31 (32%)	1542 (38%)
Inadequate response	Event	232 (20%)	20 (16%)	46 (17%)	140 (29%)	144 (26%)	36 (26%)	148 (27%)	223 (62%)	89 (26%)	23 (24%)	1101 (27%)
Adverse events and inadequate response	Event	148 (13%)	51 (41%)	54 (20%)	107 (22%)	32 (6%)	11 (8%)	46 (8%)	32 (9%)	0 (0%)	0 (0%)	481 (12%)
Protocol violation / non-compliance	Event	102 (9%)	15 (12%)	41 (15%)	11 (2%)	68 (12%)	27 (20%)	0 (0%)	21 (6%)	21 (6%)	3 (3%)	309 (8%)
Withdrew consent	Event	121 (11%)	13 (11%)	25 (9%)	64 (13%)	65 (11%)	2 (1%)	55 (10%)	4 (1%)	68 (20%)	35 (36%)	452 (11%)
Other[1]	Event	29 (3%)	0 (0%)	7 (3%)	24 (5%)	26 (5%)	5 (4%)	37 (7%)	9 (2%)	32 (9%)	4 (4%)	173 (4%)
Total events[3]		**1137 (35%)**	**123 (38%)**	**266 (31%)**	**478 (28%)**	**570 (29%)**	**137 (29%)**	**545 (47%)**	**362 (61%)**	**344 (27%)**	**96 (34%)**	**4058 (34%)**
Illness or death	Censored	34 (2%)	10 (5%)	17 (3%)	7 (1%)	20 (1%)	1 (0%)	10 (2%)	9 (4%)	0 (0%)	0 (0%)	108 (1%)
Remission of seizures	Censored	49 (2%)	4 (2%)	38 (6%)	75 (6%)	40 (3%)	12 (4%)	44 (7%)	21 (9%)	0 (0%)	0 (0%)	283 (4%)
Lost to follow-up	Censored	81 (4%)	31 (16%)	51 (9%)	63 (5%)	33 (3%)	24 (7%)	18 (3%)	0 (0%)	41 (5%)	0 (0%)	342 (4%)

Other[2]	Censored	104 (5%)	6 (3%)	22 (4%)	82 (7%)	31 (2%)	5 (2%)	26 (4%)	26 (12%)	0 (0%)	25 (13%)	327 (4%)
Completed Study	Censored	1829 (87%)	139 (73%)	468 (79%)	949 (81%)	1272 (91%)	291 (87%)	501 (84%)	166 (75%)	868 (95%)	161 (87%)	6644 (86%)
Total censored[3]		**2097 (65%)**	**190 (62%)**	**596 (69%)**	**1176 (72%)**	**1396 (71%)**	**333 (71%)**	**599 (53%)**	**222 (39%)**	**909 (73%)**	**186 (66%)**	**7704 (66%)**
Missing[4]		24	7	1	26	12	8	14	11	0	0	103
Total[5]		**3258**	**320**	**863**	**1680**	**1978**	**478**	**1158**	**595**	**1253**	**282**	**11,865**

1. Other treatment related reasons included: Physician's decision, drug-related death, pregnancy or perceived remission or non-specific (drug related) reason.

2. Other non-treatment related reasons included: epilepsy diagnosis changed, participants developed other medical disorders including neurological and psychiatric disorders or non-specific (non-drug related) reason.

3. Proportions for specific reasons indicate proportion of total events or total censored. Proportion for total events and total censored indicate the proportion of total participants.

4. Those with missing reason for withdrawal were primarily classified as censored in analysis and performed a sensitivity analysis treating these individuals as having withdrawal 'events.' Results of sensitivity analysis were practically identical and conclusions unchanged.

5. Four studies did not contribute to analysis of time-to-withdrawal of allocated treatment.

7.2.1 Primary outcome: Time-to-withdrawal of allocated treatment

11,865 out of 12,391 participants (96%) contributed to analysis of 'Time-to-withdrawal of allocated treatment'. Withdrawal information was not available for three trials and all participants completed follow-up in one trial so 'Time-to-withdrawal of allocated treatment' could not be calculated for four trials (4% of participants with IPD provided).

Table 18 shows the reported reasons for premature withdrawal from the trial and these reasons were classified in analysis. In some instances, participants many have withdrawn from treatment for a combination of reasons. For the purpose of analysis a judgement was made regarding the primary reason for withdrawal (see Chapter 5.3.3.2 for further details of classification). It should be noted that the information reported in Table 18 does not take account of randomisation within trials and should be interpreted as exploratory.

Out of 11865 individuals, 4058 (34%) prematurely withdrew and 7704 were censored in analysis (65%). For 103 participants, reason for withdrawal was missing (ranging by drug from 0 participants (LEV and ZNS) to 26 participants (VPS)). We treated those with missing reason for withdrawal as censored in analysis and performed a sensitivity analysis treating these individuals as having withdrawal 'events'. Results of sensitivity analysis were practically identical and conclusions unchanged (results not presented for brevity), therefore these individuals are censored in results presented.

7.2.1.1 Direct evidence

Table 19 (individuals with partial seizures) and Table 20 (individuals with generalised seizures) show the number of trials and participants contributing direct evidence for each of the pairwise comparisons in the network. Results highlighted in bold indicate statistically significant results and HR <1 indicates an advantage to the second drug in the comparison. All results presented are calculated with fixed-effects.

Twenty out of 45 comparisons had no direct evidence for individuals with partial seizures. Thirteen out of 36 comparisons had no direct evidence for individuals with generalised seizures and eight comparisons for individuals with generalised seizures had less than 20 individuals contributing direct evidence resulting in wide CIs around the treatment effect estimate. Comparisons with the most participants contributing to analysis were CBZ vs LTG and CBZ vs LEV (partial seizures) and VPS vs LEV and VPS vs TPM (generalised seizures).

Table 19: Pairwise and network meta-analysis results - Time-to-withdrawal of allocated treatment for individuals with partial seizures

Comparision[1]	Direct Evidence (Pairwise meta-analysis)			Direct plus Indirect Evidence (Network Meta-Analysis)	
	Number of studies (participants)	HR (95% CI)[2,3]	I^2 (%)	Proportion of direct evidence[5]	HR (95% CI)[2,3]
CBZ vs PHB	4 (520)	**1.57 (1.16 to 2.13)**	0%	52.50%	**1.55 (1.18 to 2.04)**
CBZ vs PHT	3 (428)	1.03 (0.74 to 1.42)	63.6%	12.80%	1.13 (0.92 to 1.38)
CBZ vs VPS	5 (814)	0.94 (0.73 to 1.19)	0%	40.10%	1.04 (0.86 to 1.25)
CBZ vs LTG	9 (2268)	**0.76 (0.61 to 0.95)**	39.3%	28.90%	**0.75 (0.65 to 0.86)**
CBZ vs OXC	2 (562)	4.62 (0.95 to 22.4)	0%	5.70%	1.09 (0.84 to 1.42)
CBZ vs TPM	2 (937)	1.04 (0.52 to 2.07)	0%	7.40%	1.18 (0.98 to 1.43)
CBZ vs GBP	2 (954)	1.14 (0.84 to 1.55)	0%	87.10%	**1.20 (1.00 to 1.43)**
CBZ vs LEV	3 (1567)	**0.70 (0.52 to 0.94)**	0%	37.90%	**0.82 (0.69 to 0.97)**
CBZ vs ZNS	1 (583)	1.08 (0.81 to 1.44)	NA[4]	100%	1.08 (0.79 to 1.48)
PHB vs PHT	3 (404)	**0.67 (0.50 to 0.91)**	65%	15.20%	**0.73 (0.55 to 0.96)**
PHB vs VPS	2 (75)	**0.68 (0.34 to 1.36)**	23%	8.80%	**0.67 (0.48 to 0.92)**
PHB vs LTG	No direct evidence			0%	**0.48 (0.35 to 0.66)**
PHB vs OXC	No direct evidence			0%	0.70 (0.48 to 1.03)
PHB vs TPM	No direct evidence			0%	0.76 (0.55 to 1.06)
PHB vs GBP	No direct evidence			0%	0.77 (0.55 to 1.07)
PHB vs LEV	No direct evidence			0%	**0.53 (0.38 to 0.73)**
PHB vs ZNS	No direct evidence			0%	0.70 (0.46 to 1.06)
PHT vs VPS	4 (168)	1.00 (0.60 to 1.64)	58.5%	9%	0.92 (0.70 to 1.21)
PHT vs LTG	1 (90)	1.10 (0.57 to 2.14)	NA	11.60%	**0.66 (0.52 to 0.85)**
PHT vs OXC	2 (325)	0.65 (0.32 to 1.32)	0%	40.40%	0.97 (0.69 to 1.35)
PHT vs TPM	1 (53)	0.77 (0.38 to 1.57)	NA	10.90%	1.05 (0.80 to 1.39)
PHT vs GBP	No direct evidence			0%	1.06 (0.81 to 1.40)
PHT vs LEV	No direct evidence			0%	**0.73 (0.56 to 0.95)**
PHT vs ZNS	No direct evidence			0%	0.96 (0.66 to 1.39)
VPS vs LTG*	3 (221)	**1.40 (1.00 to 1.96)**	45.1%	5.10%	**0.72 (0.58 to 0.90)**
VPS vs OXC	No direct evidence			0%	1.05 (0.76 to 1.44)
VPS vs TPM	2 (111)	**1.66 (1.24 to 2.23)**	48.1%	33.70%	1.14 (0.88 to 1.48)
VPS vs GBP	No direct evidence			0%	1.15 (0.89 to 1.49)
VPS vs LEV	1 (190)	1.14 (0.73 to 1.75)	NA	17.20%	0.79 (0.61 to 1.03)
VPS vs ZNS	No direct evidence			0%	1.04 (0.73 to 1.50)
LTG vs OXC	1 (506)	0.69 (0.12 to 4.14)	NA	4.40%	**1.46 (1.11 to 1.92)**
LTG vs TPM	1 (648)	1.18 (0.86 to 1.62)	NA	20.90%	**1.59 (1.29 to 1.95)**
LTG vs GBP	1 (659)	0.62 (0.06 to 6.01)	NA	1%	**1.60 (1.31 to 1.96)**
LTG vs LEV	1 (240)	0.86 (0.58 to 1.28)	NA	23.70%	1.10 (0.89 to 1.35)
LTG vs ZNS	No direct evidence			0%	**1.45 (1.03 to 2.04)**
OXC vs TPM	1 (496)	0.87 (0.16 to 4.73)	NA	4.90%	1.09 (0.82 to 1.44)
OXC vs GBP	1 (507)	0.90 (0.08 to 9.96)	NA	2.30%	1.10 (0.83 to 1.45)
OXC vs LEV	No direct evidence			0%	0.75 (0.55 to 1.03)

OXC vs ZNS	No direct evidence			0%	0.99 (0.66 to 1.49)
TPM vs GBP	1 (649)	1.04 (0.12 to 9.33)	NA	1.10%	1.01 (0.82 to 1.25)
TPM vs LEV	No direct evidence			0%	**0.69 (0.54 to 0.89)**
TPM vs ZNS	No direct evidence			0%	0.91 (0.64 to 1.31)
GBP vs LEV	No direct evidence			0%	**0.69 (0.54 to 0.88)**
GBP vs ZNS	No direct evidence			0%	0.90 (0.63 to 1.30)
LEV vs ZNS	No direct evidence			0%	1.32 (0.93 to 1.88)

1. Order of drugs in the table: most commonly used drug first (CBZ), then drugs are ordered approximately by the date they were licenced as a monotherapy treatment (oldest first).

2. HRs and 95% CIs are calculated from fixed-effects analyses.

3. HR<1 indicates an advantage to the second drug in the comparison; results in highlighted in bold are statistically significant.

4. NA - heterogeneity is not applicable as only one study contributes direct evidence.

5. Proportion of the estimate contributed by direct evidence (see Chaimani *et al* [299]) .

Comparisons marked with a *, confidence intervals of direct evidence and network meta-analysis do not overlap indicating the inconsistency may be present in the results.

Table 19 and Table 20 also show heterogeneity in the direct treatment effects. No substantial heterogeneity was present (I^2 greater than 50%) for any comparison for individuals with generalised seizures. For three comparisons for individuals with partial seizures, substantial heterogeneity was present (I^2 greater than 50%). The heterogeneity in these comparisons seemed to originate from difference in trial designs contributing to the pooled result; i.e. pooling of trials recruiting children only, adults only or elderly participants only and pooling of double-blind and open label trials (previous pairwise Cochrane Epilepsy IPD reviews in this series have discussed the importance of blinding to the outcome of time-to-withdrawal of allocated treatment) [61, 318]. Repeating analysis with random-effects did not change conclusions for two of the comparisons (CBZ vs PHT and PHT vs VPS) but for one comparison (PHB vs PHT), when repeating analysis with random-effects there was no longer a statistically significant advantage to PHT; pooled HR 0.42 (95% CI 0.16 to 1.06).

Table 20: Pairwise and network meta-analysis results - Time-to-withdrawal of allocated treatment for individuals with generalised seizures

Comparision[1]	Direct Evidence (Pairwise meta-analysis)			Direct plus Indirect Evidence (Network Meta-Analysis)	
	Number of studies (participants)	HR (95% CI)[2,3]	I^2 (%)	Proportion of direct evidence[5]	HR (95% CI)[2,3]
CBZ vs PHB	3 (156)	1.21 (0.51 to 2.86)	11.8%	27.30%	1.47 (0.83 to 2.61)
CBZ vs PHT	2 (118)	2.68 (0.95 to 7.57)	0%	11.30%	0.92 (0.59 to 1.42)
CBZ vs VPS	4 (405)	1.26 (0.73 to 2.20)	6.6%	27.30%	**0.70 (0.54 to 0.92)**
CBZ vs LTG	7 (302)	1.23 (0.72 to 2.10)	0%	39.20%	**0.63 (0.45 to 0.89)**
CBZ vs OXC	1 (9)	0.39 (0.03 to 4.35)	NA[4]	3.90%	1.00 (0.21 to 4.81)
CBZ vs TPM	2 (101)	1.10 (0.51 to 2.36)	0%	23.20%	1.24 (0.90 to 1.71)
CBZ vs GBP	1 (6)	0.49 (0.03 to 7.90)	NA	8.50%	0.90 (0.11 to 7.29)
CBZ vs LEV	2 (251)	1.22 (0.74 to 2.02)	0%	57%	0.74 (0.44 to 1.23)
PHB vs PHT	2 (95)	1.56 (0.49 to 4.99)	0%	16.10%	0.62 (0.32 to 1.24)
PHB vs VPS	2 (94)	0.56 (0.20 to 1.54)	0%	19.40%	**0.48 (0.27 to 0.86)**
PHB vs LTG	No direct evidence			0%	**0.43 (0.22 to 0.83)**
PHB vs OXC	No direct evidence			0%	0.68 (0.13 to 3.60)
PHB vs TPM	No direct evidence			0%	0.84 (0.44 to 1.60)
PHB vs GBP	No direct evidence			0%	0.61 (0.07 to 5.34)
PHB vs LEV	No direct evidence			0%	0.50 (0.23 to 1.09)
PHT vs VPS	3 (326)	0.66 (0.30 to 1.45)	22.6%	19.30%	0.77 (0.46 to 1.27)
PHT vs LTG	1 (91)	1.11 (0.42 to 2.94)	NA	14.90%	0.69 (0.39 to 1.20)
PHT vs OXC	2 (155)	1.05 (0.44 to 2.52)	0%	37.90%	1.09 (0.21 to 5.56)
PHT vs TPM	1 (150)	1.68 (0.49 to 5.69)	NA	11.20%	1.35 (0.79 to 2.30)
PHT vs GBP	No direct evidence			0%	0.98 (0.12 to 8.30)
PHT vs LEV	No direct evidence			0%	0.80 (0.42 to 1.55)
VPS vs LTG	3 (387)	0.46 (0.22 to 0.97)	0%	14.80%	0.90 (0.60 to 1.35)
VPS vs OXC	No direct evidence			0%	1.42 (0.29 to 6.92)
VPS vs TPM*	2 (443)	0.53 (0.27 to 1.07)	48.5%	22.40%	**1.76 (1.22 to 2.53)**
VPS vs GBP	No direct evidence			0%	1.28 (0.16 to 10.5)
VPS vs LEV	1 (512)	0.68 (0.30 to 1.59)	NA	18.60%	1.05 (0.58 to 1.90)
LTG vs OXC	1 (10)	2.09 (0.34 to 12.8)	NA	7.60%	1.58 (0.33 to 7.67)
LTG vs TPM	1 (14)	1.10 (0.42 to 2.89)	NA	7.30%	**1.96 (1.25 to 3.08)**
LTG vs GBP	1 (7)	2.63 (0.27 to 25.7)	NA	13.80%	1.42 (0.17 to 11.6)
LTG vs LEV	No direct evidence			0%	1.17 (0.63 to 2.19)
OXC vs TPM	1 (14)	1.31 (0.24 to 7.32)	NA	9%	1.24 (0.26 to 5.94)
OXC vs GBP	1 (7)	1.26 (0.11 to 14.1)	NA	12.70%	0.90 (0.08 to 9.96)
OXC vs LEV	No direct evidence			0%	0.74 (0.14 to 3.86)
TPM vs GBP	1 (11)	0.96 (0.11 to 8.67)	NA	14.60%	0.73 (0.09 to 5.89)
TPM vs LEV	No direct evidence			0%	0.60 (0.33 to 1.09)
GBP vs LEV	No direct evidence			0%	0.82 (0.10 to 7.10)

1. See Table 19 for details of footnotes

7.2.1.2 NMA results (direct plus indirect evidence)

Figure 13 shows how each AED performs compared to first-line treatment CBZ for individuals with partial seizures (ordered by treatment effect estimate); LTG and LEV are significantly better than CBZ and CBZ is significantly better than GBP and PHB. Figure 14 shows how each AED performs compared to first-line treatment LTG for individuals with partial seizures (ordered by treatment effect estimate); LTG is significantly better than all AEDs except for LEV. Figure 15 shows how each AED performs compared to first-line treatment VPS for individuals with generalised seizures (ordered by treatment effect estimate); VPS is significantly better than CBZ, TPM and PHB.

Table 19 and Table 20 (above) show treatment effect estimates for all pairwise comparisons in the network combining direct with indirect evidence (NMA). In addition to the performance of the AEDs compared to the first-line treatments (as described above), for individuals with partial seizures, LEV seems to perform better than most other AEDs and for individuals with generalised seizures, LTG seems to perform better than most other AEDs. For both individuals with partial seizures and individuals with generalised seizures, PHB seems to perform worse than most other AEDs.

Figure 13: All AEDs compared to carbamazepine (CBZ) for time to treatment withdrawal, individuals with partial seizures

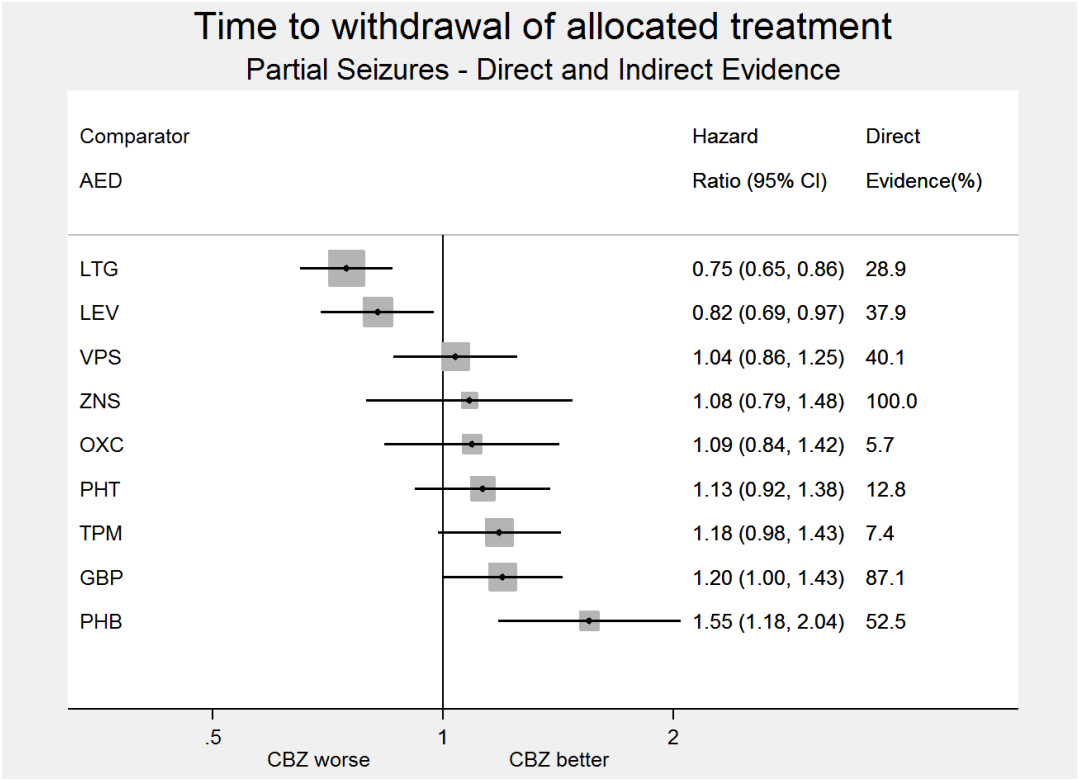


Figure 14: All AEDs compared to lamotrigine (LTG) for time to treatment withdrawal, individuals with partial seizures

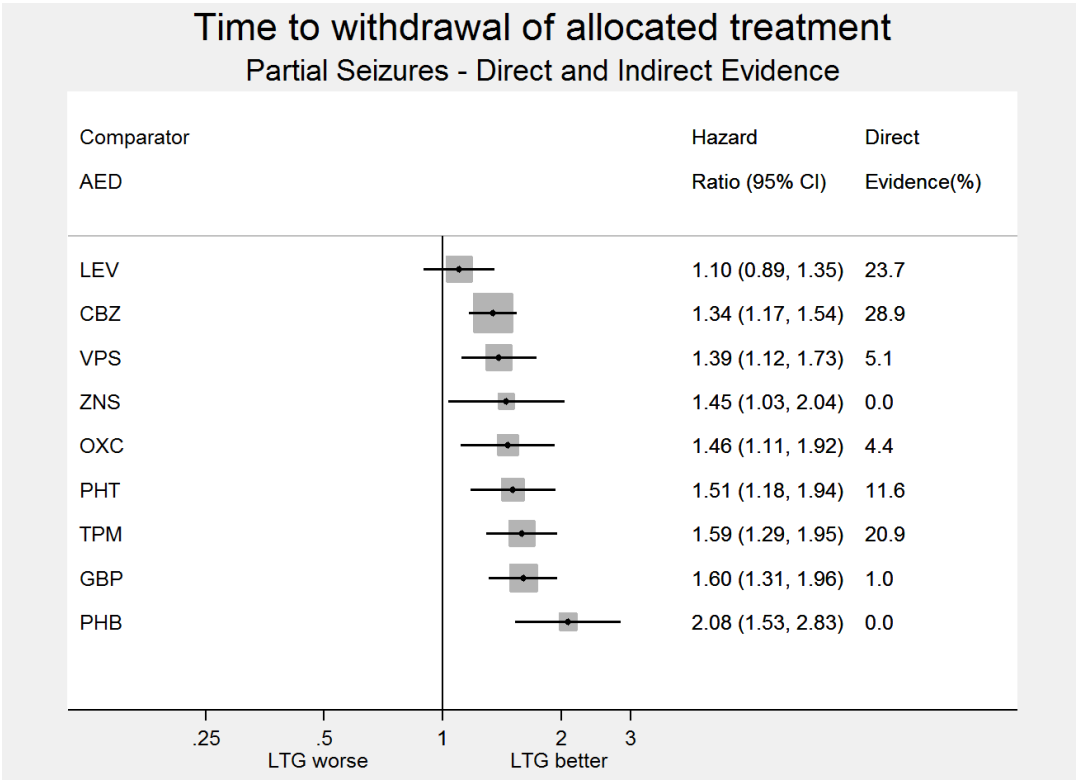
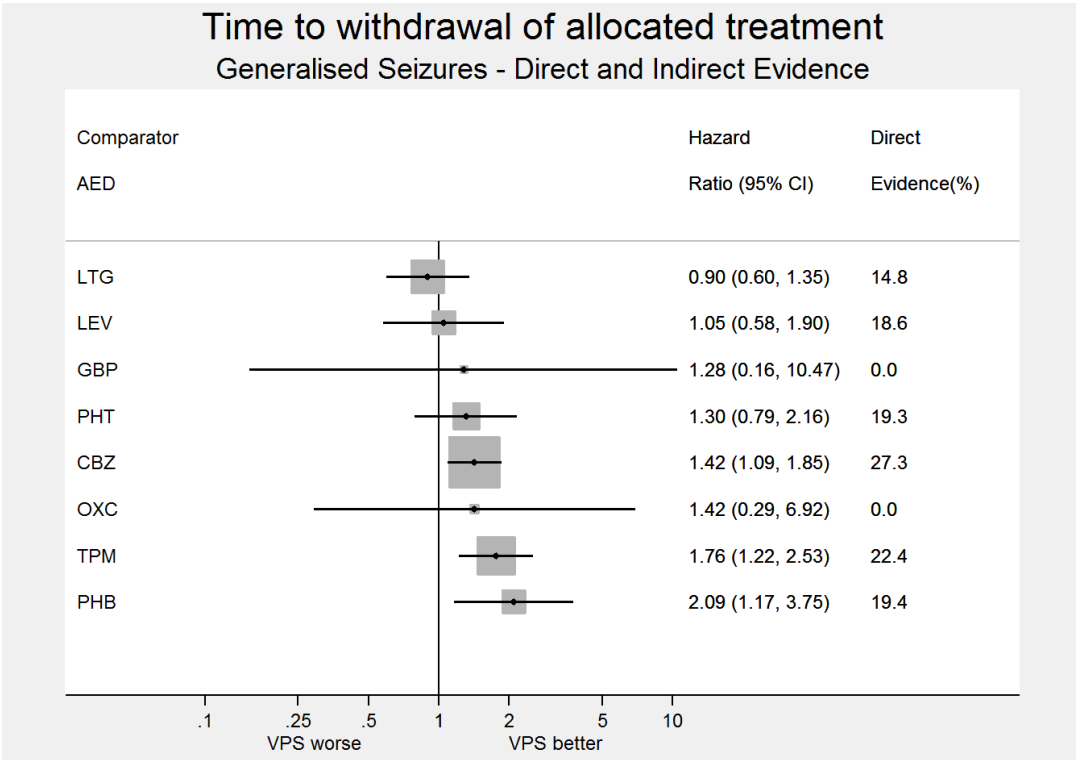


Figure 15: All AEDs compared to sodium valproate (VPS) for time to treatment withdrawal, individuals with generalised seizures



As described in Chapter 6.2.5., an I^2 statistic could not be directly calculated for the NMA but could be estimated. The estimated I^2 statistic was 11.7% and when repeating NMA with random-effects, calculated the τ^2 statistic was 0.0037. Numerical results for treatment effects were very similar (the same to one decimal place, results not presented) and conclusions remained unchanged. Therefore, if any heterogeneity is present within this NMA, the impact upon results is negligible.

7.2.1.3 Investigation of inconsistency

The 'design-by-treatment' inconsistency model was fitted to 17 variables and regressed on 23 designs, five of which were multi-arm trials (up to five treatment arms). Accounting for the multi-arm trials, this resulted in a χ^2 test for inconsistency with 36 degrees of freedom which was not significant: (χ^2 (36) = 45.6, p-value = 0.131, heterogeneity (τ) = 5.65 x 10^{-10}).

Notably, for most pairwise comparisons, numerical results of direct evidence and NMA are similar, mostly in the same direction and CIs of estimates overlap. For all pairwise comparisons, results from NMA are more precise than results from direct evidence (in some cases, much more precise where limited direct evidence exists, for example, see CBZ compared to OXC, Appendix 16, Figure 36). For the following comparisons, conclusions drawn from direct evidence and from NMA are different (see Table 19 and Table 20):

- Direct evidence shows a significant advantage to one of the AEDs and the NMA results show no significant difference between the AEDs: VPS vs TPM (partial seizures)
- Direct evidence shows no significant difference between the AEDs and NMA shows a significant advantage for one of the AEDs: CBZ vs GBP, LTG vs OXC, LTG vs TPM, LTG vs GBP (all partial seizures); CBZ vs VPS, CBZ vs LTG, PHB vs VPS, VPS vs TPM, LTG vs TPM (all generalised seizures)
- No direct evidence exists between the AEDs while NMA shows a significant advantage for one of the AEDs: PHB vs LTG, PHB vs LEV, LTG vs ZNS, TPM vs LEV, GBP vs LEV (all partial seizures); PHB vs LTG (generalised seizures)

For the following comparisons; CIs for the results from indirect evidence do not overlap with (see Appendix 16):

- Direct evidence: CBZ vs PHT (generalised seizures), PHB vs PHT (generalised seizures).
- NMA results: LTG vs PHT (partial seizures), CBZ vs PHT (generalised seizures), LTG vs PHT (generalised seizures).

For the following comparisons; CIs for the results from direct evidence and from NMA do not overlap which indicates potential inconsistency is present (see Table 19 and Table 20, results marked with *): VPS vs LTG (partial seizures), VPS vs TPM (generalised seizures).

For the comparison of VPS vs LTG for individuals with partial seizures, from direct evidence only, there is a statistically significant advantage to VPS (HR 1.40 (1.00 to 1.96), however from the NMA results, the direction of effect changes to a statistically significant advantage to LTG (HR 0.72 (0.58 to 0.90)). However, for this comparison, only 5.1% of the network estimate is contributed from direct evidence and a moderate amount of heterogeneity is present in this estimate (I^2=45%), likely due to variability in the trial design of the three trials contributing to this estimate (for example, one trial was designed to only recruit individuals with generalised or unclassified seizures but did recruit a small number of individuals with partial seizures who contribute to this outcome) [264].

For the comparison of VPS vs TPM for individuals with generalised seizures, from direct evidence, there is no significant difference between the drugs (HR 0.53 (0.27 to 1.07)), however from the NMA results, a statistically significant advantage is shown for VPS (HR 1.76 (1.22 to 2.53)). As above, for this comparison, only 22.4% of the network estimate is contributed from direct evidence and a moderate amount of heterogeneity is present in this estimate (I^2=48.5%). Again, this heterogeneity is likely due to difference in trial design of the two trials contributing direct evidence (see characteristics of Privitera et al [304] for details of stratification).

Furthermore, the 'design-by treatment' inconsistency model does not show any significant evidence of inconsistency within the network. Therefore, we are not concerned about any impact of this observed inconsistency of numerical results on the conclusions of the review.

7.2.2 Secondary outcome: Time-to-first seizure after randomisation

12,152 out of 12,391 participants (98%) contributed to analysis of 'Time-to-first seizure post-randomisation.' For 239 participants (2%) seizure dates after randomisation were missing so these individuals could not contribute to analysis.

7.2.2.1 Direct evidence

Table 21 (individuals with partial seizures) and Table 22 (individuals with generalised seizures) show the number of trials and participants contributing direct evidence for each of the pairwise comparisons in the network. Results highlighted in bold indicate statistically significant results and HR <1 indicates an advantage to the second drug in the comparison. All results presented are calculated with fixed-effects.

Twenty out of 45 comparisons had no direct evidence for individuals with partial seizures. Thirteen out of 36 comparisons had no direct evidence for individuals with generalised seizures and eight comparisons for individuals with generalised seizures had less than 20 individuals contributing direct evidence resulting in wide CIs around the treatment effect estimate. Comparisons with the most participants contributing to analysis were CBZ vs LTG and CBZ vs LEV (partial seizures) and VPS vs LEV and VPS vs TPM (generalised seizures).

Table 21: **Pairwise and network meta-analysis results - Time-to-first seizure for individuals with partial seizures**

Comparision[1]	Direct Evidence (Pairwise meta-analysis)			Direct plus Indirect Evidence (Network Meta-Analysis)	
	Number of studies (participants)	HR (95% CI)[2,3]	I^2 (%)	Proportion of direct evidence[5]	HR (95% CI)[2,3]
CBZ vs PHB	6 (581)	0.99 (0.78 to 1.26)	54.3%	21%	**0.79 (0.64 to 0.97)**
CBZ vs PHT	4 (432)	0.91 (0.72 to 1.16)	16.1%	27.10%	0.98 (0.85 to 1.13)
CBZ vs VPS	5 (813)	1.01 (0.86 to 1.19)	32%	34.60%	**1.20 (1.06 to 1.37)**
CBZ vs LTG	9 (2252)	0.98 (0.75 to 1.27)	0%	40.70%	**1.29 (1.17 to 1.42)**
CBZ vs OXC	2 (555)	1.47 (0.57 to 3.81)	57.3%	4.80%	1.09 (0.89 to 1.32)
CBZ vs TPM	2 (925)	1.03 (0.51 to 2.08)	69.3%	1.50%	1.12 (0.97 to 1.29)
CBZ vs GBP	2 (943)	**1.64 (1.14 to 2.36)**	17.7%	49%	**1.44 (1.25 to 1.66)**
CBZ vs LEV	3 (1552)	1.18 (0.85 to 1.65)	0%	26.20%	1.14 (0.99 to 1.30)
CBZ vs ZNS	1 (581)	1.30 (0.97 to 1.73)	NA[4]	100%	1.30 (0.97 to 1.73)
PHB vs PHT	5 (463)	1.07 (0.83 to 1.37)	27.7%	33.60%	1.24 (0.99 to 1.56)
PHB vs VPS*	2 (80)	0.71 (0.43 to 1.17)	9.10%	12.80%	**1.53 (1.20 to 1.94)**
PHB vs LTG	No direct evidence			0%	**1.63 (1.30 to 2.06)**
PHB vs OXC	No direct evidence			0%	**1.38 (1.04 to 1.83)**
PHB vs TPM	No direct evidence			0%	**1.42 (1.11 to 1.83)**
PHB vs GBP	No direct evidence			0%	**1.83 (1.42 to 2.35)**
PHB vs LEV	No direct evidence			0%	**1.44 (1.12 to 1.85)**
PHB vs ZNS	No direct evidence			0%	**1.64 (1.15 to 2.35)**
PHT vs VPS	5 (245)	0.96 (0.72 to 1.29)	0%	25.40%	**1.23 (1.02 to 1.48)**
PHT vs LTG	1 (90)	0.77 (0.38 to 1.54)	NA	6%	**1.31 (1.10 to 1.57)**

PHT vs OXC	2 (318)	1.46 (0.88 to 2.44)	23.9%	36.10%	1.11 (0.87 to 1.41)
PHT vs TPM	1 (53)	2.32 (0.95 to 5.70)	NA	4%	1.14 (0.93 to 1.40)
PHT vs GBP	No direct evidence			0%	**1.47 (1.20 to 1.80)**
PHT vs LEV	No direct evidence			0%	1.16 (0.95 to 1.41)
PHT vs ZNS	No direct evidence			0%	1.32 (0.96 to 1.82)
VPS vs LTG	3 (215)	**1.57 (1.23 to 2.00)**	39.4%	10%	1.07 (0.92 to 1.24)
VPS vs OXC	No direct evidence			0%	0.90 (0.72 to 1.14)
VPS vs TPM	2 (111)	1.18 (0.93 to 1.50)	0%	70.20%	0.93 (0.77 to 1.13)
VPS vs GBP	No direct evidence			0%	1.20 (0.99 to 1.44)
VPS vs LEV	1 (190)	1.27 (0.94 to 1.72)	NA	31%	0.94 (0.77 to 1.15)
VPS vs ZNS	No direct evidence			0%	1.08 (0.78 to 1.48)
LTG vs OXC	1 (499)	0.87 (0.23 to 3.25)	NA	5.50%	0.84 (0.69 to 1.03)
LTG vs TPM	1 (636)	0.73 (0.57 to 0.93)	NA	2.30%	0.87 (0.75 to 1.01)
LTG vs GBP	1 (647)	0.63 (0.07 to 5.42)	NA	4.40%	1.12 (0.96 to 1.30)
LTG vs LEV	1 (229)	0.84 (0.53 to 1.35)	NA	15.90%	0.88 (0.75 to 1.04)
LTG vs ZNS	No direct evidence			0%	1.01 (0.74 to 1.36)
OXC vs TPM	1 (487)	0.55 (0.15 to 2.06)	NA	5.40%	1.03 (0.84 to 1.27)
OXC vs GBP	1 (498)	0.73 (0.08 to 6.49)	NA	4.60%	**1.32 (1.08 to 1.63)**
OXC vs LEV	No direct evidence			0%	1.05 (0.83 to 1.32)
OXC vs ZNS	No direct evidence			0%	1.19 (0.84 to 1.69)
TPM vs GBP	1 (635)	1.31 (0.15 to 11.2)	NA	3.50%	**1.28 (1.09 to 1.51)**
TPM vs LEV	No direct evidence			0%	1.01 (0.83 to 1.23)
TPM vs ZNS	No direct evidence			0%	1.15 (0.84 to 1.59)
GBP vs LEV	No direct evidence			0%	**0.79 (0.65 to 0.96)**
GBP vs ZNS	No direct evidence			0%	0.90 (0.65 to 1.24)
LEV vs ZNS	No direct evidence			0%	1.14 (0.83 to 1.57)

1. See Table 19 for details of footnotes

Table 21 and Table 22 also show heterogeneity in the direct treatment effects. For three comparisons for individuals with partial seizures and for three comparisons for individuals with generalised seizures, substantial heterogeneity was present (I^2 greater than 50%). The heterogeneity in these comparisons seemed to originate from differences in trial designs contributing to the pooled result; i.e. pooling of trials recruiting different age groups and pooling trials with or without treatment strata (see Chapter 5.2.3.4 for further details).

For the comparisons for individuals with partial seizures, none of the treatment effects with substantial heterogeneity present were statistically significant so conclusions would not change if random-effects were applied. For the comparisons for individuals with generalised seizures, repeating analysis with random-effects did not change conclusions for two of the comparisons (CBZ vs VPS and PHT vs VPS) but for one comparison (CBZ vs PHB), when repeating analysis with random-effects there was no longer a statistically significant advantage to PHB: HR 0.59 (0.27 to 1.26).

Table 22: Pairwise and network meta-analysis results - Time-to-first seizure for individuals with partial seizures

Comparision[1]	Direct Evidence (Pairwise meta-analysis)			Direct plus Indirect Evidence (Network Meta-Analysis)	
	Number of studies (participants)	HR (95% CI)[2,3]	I^2 (%)	Proportion of direct evidence[5]	HR (95% CI)[2,3]
CBZ vs PHB	5 (237)	**0.55 (0.33 to 0.92)**	50.4%	35.50%	1.10 (0.80 to 1.51)
CBZ vs PHT	3 (150)	0.88 (0.51 to 1.54)	0%	26.60%	**0.76 (0.59 to 0.98)**
CBZ vs VPS	4 (411)	1.37 (0.98 to 1.92)	84.1%	10.40%	0.88 (0.76 to 1.03)
CBZ vs LTG	7 (302)	1.49 (0.94 to 2.35)	0%	0.30%	0.98 (0.70 to 1.37)
CBZ vs OXC	1 (9)	1.55 (0.38 to 6.31)	NA[4]	9%	1.09 (0.36 to 3.36)
CBZ vs TPM	2 (101)	1.19 (0.56 to 2.50)	62%	9%	1.15 (0.89 to 1.48)
CBZ vs GBP	1 (6)	2.83 (0.31 to 25.5)	NA	10.70%	0.79 (0.10 to 6.08)
CBZ vs LEV	2 (251)	1.04 (0.65 to 1.64)	0%	44.90%	1.19 (0.78 to 1.83)
PHB vs PHT	4 (161)	1.41 (0.76 to 2.62)	46.9%	20.30%	**0.69 (0.48 to 1.00)**
PHB vs VPS	2 (98)	1.87 (0.87 to 4.00)	69.8%	6.50%	0.80 (0.57 to 1.12)
PHB vs LTG	No direct evidence			0%	0.89 (0.56 to 1.42)
PHB vs OXC	No direct evidence			0%	1.00 (0.31 to 3.20)
PHB vs TPM	No direct evidence			0%	1.05 (0.70 to 1.56)
PHB vs GBP	No direct evidence			0%	0.72 (0.09 to 5.68)
PHB vs LEV	No direct evidence			0%	1.09 (0.64 to 1.85)
PHT vs VPS	4 (394)	1.11 (0.71 to 1.74)	0%	36.40%	1.16 (0.88 to 1.53)
PHT vs LTG	1 (91)	1.00 (0.40 to 2.46)	NA	16.20%	1.29 (0.85 to 1.97)
PHT vs OXC	2 (154)	0.60 (0.33 to 1.10)	49.7%	25.20%	1.44 (0.46 to 4.56)
PHT vs TPM	1 (150)	0.63 (0.18 to 2.26)	NA	9.80%	1.51 (1.06 to 2.15)
PHT vs GBP	No direct evidence			0%	1.05 (0.13 to 8.14)
PHT vs LEV	No direct evidence			0%	1.57 (0.96 to 2.58)
VPS vs LTG	3 (377)	0.64 (0.37 to 1.11)	23.2%	31.30%	1.11 (0.77 to 1.60)
VPS vs OXC	No direct evidence			0%	1.24 (0.40 to 3.84)
VPS vs TPM*	2 (441)	**0.42 (0.23 to 0.80)**	46.4%	21%	**1.30 (1.01 to 1.68)**
VPS vs GBP	No direct evidence			0%	0.90 (0.12 to 6.92)
VPS vs LEV	1 (512)	0.82 (0.48 to 1.40)	NA	34%	1.35 (0.86 to 2.13)
LTG vs OXC	1 (10)	0.94 (0.25 to 3.57)	NA	12.20%	1.12 (0.36 to 3.48)
LTG vs TPM	1 (14)	0.61 (0.28 to 1.30)	NA	13.10%	1.17 (0.78 to 1.77)
LTG vs GBP	1 (7)	1.72 (0.20 to 14.9)	NA	11.90%	0.81 (0.11 to 6.25)
LTG vs LEV	No direct evidence			0%	1.22 (0.71 to 2.10)
OXC vs TPM	1 (14)	1.90 (0.50 to 7.19)	NA	13.60%	1.05 (0.34 to 3.24)
OXC vs GBP	1 (7)	1.83 (0.20 to 16.5)	NA	13.30%	0.73 (0.08 to 6.49)
OXC vs LEV	No direct evidence			0%	1.09 (0.33 to 3.62)
TPM vs GBP	1 (11)	0.96 (0.11 to 8.29)	NA	13.20%	0.69 (0.09 to 5.32)
TPM vs LEV	No direct evidence			0%	1.04 (0.63 to 1.71)
GBP vs LEV	No direct evidence			0%	1.50 (0.19 to 12.0)

1. See Table 19 for details of footnotes

7.2.2.2 NMA results (direct plus indirect evidence)

Figure 16 shows how each AED performs compared to first-line treatment CBZ for individuals with partial seizures (ordered by treatment effect estimate); PHB is significantly better than CBZ and CBZ is significantly better than VPS, LTG and GBP. Figure 17 shows how each AED performs compared to first-line treatment LTG for individuals with partial seizures (ordered by treatment effect estimate); PHB, PHT and CBZ are significantly better than LTG. Figure 18 shows how each AED performs compared to first-line treatment VPS for individuals with generalised seizures (ordered by treatment effect estimate); VPS is significantly better than TPM.

Table 21 and Table 22 (above) show treatment effect estimates for all pairwise comparisons in the network combining direct with indirect evidence (NMA). In addition to the performance of the AEDs compared to the first-line treatments (as described above); for individuals with partial seizures, PHB and PHT seems to perform better than most other drugs and for individuals with generalised seizures, PHT seems to perform better than most other drugs. There were few notable differences between the newer drugs (OXC, TPM, GBP, LEV and ZNS) for either individuals with partial seizures or generalised seizures.

Figure 16: All AEDs compared to carbamazepine (CBZ) for time-to-first seizure, individuals with partial seizures

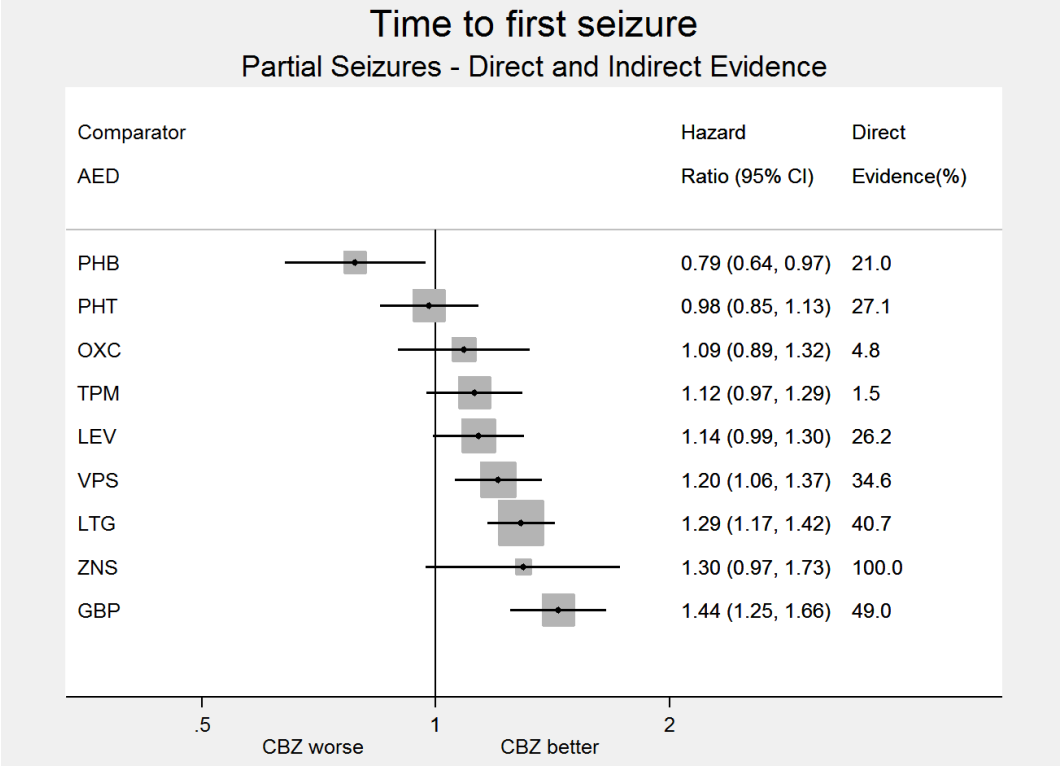


Figure 17: All AEDs compared to lamotrigine (LTG) for time-to-first seizure, individuals with partial seizures

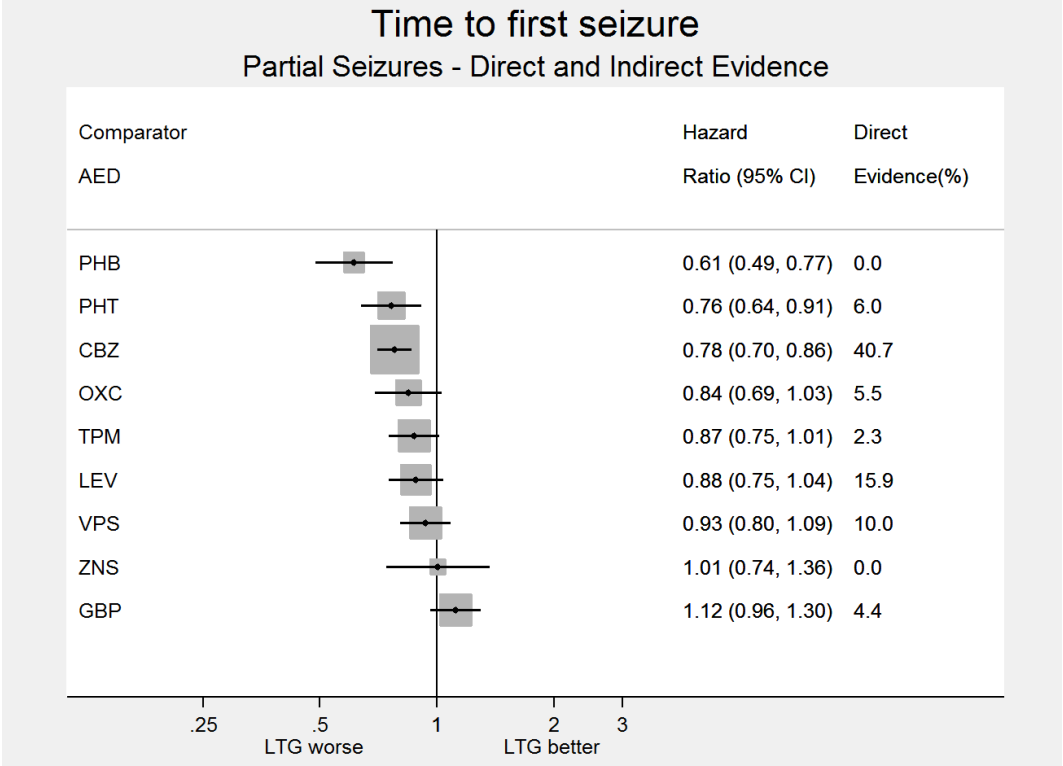


Figure 18: All AEDs compared to sodium valproate (VPS) for time-to-first seizure, individuals with generalised seizures

As described in Chapter 6.2.5, an I^2 statistic could not be directly calculated for the NMA but could be estimated. The estimated I^2 statistic was 0% and when repeating NMA with random-effects, calculated the τ^2 statistic was 9×10^{-21}. As no heterogeneity was present and τ^2 was negligible, numerical results for treatment effects and conclusions were identical.

7.2.2.3 Investigation of inconsistency

The 'design-by-treatment' inconsistency model was fitted to 17 variables and regressed on 23 designs, seven of which were multi-arm trials (up to five treatment arms). Accounting for the multi-arm trials, this resulted in a χ^2 test for inconsistency with 43 degrees of freedom which was not significant (χ^2 (43) = 38.2, p-value = 0.680, heterogeneity (τ) = 0.094).

Notably, for most pairwise comparisons, numerical results of direct evidence and NMA are similar, mostly in the same direction and CIs of estimates overlap. For all pairwise comparisons, results from NMA are more precise than results from direct evidence (in some cases much more precise where limited direct evidence exists, for example see LTG compared to GBP, Appendix 16, Figure 39). For the following comparisons; conclusions drawn from direct evidence and from NMA are different (see Table 21 and Table 22):

- Direct evidence shows a significant advantage to one of the AEDs and the NMA results show no significant difference between the AEDs: VPS vs LTG (partial seizures); CBZ vs PHB (generalised seizures).
- Direct evidence shows no significant difference between the AEDs and NMA shows a significant advantage for one of the AEDs: CBZ vs PHB, CBZ vs VPS, CBZ vs LTG, PHB vs VPS, PHT vs VPS, PHT vs LTG, OXC vs GBP (all partial seizures), CBZ vs PHT, PHB vs PHT (generalised seizures).
- No direct evidence exists between the AEDs while NMA shows a significant advantage for one of the AEDs: PHB vs LTG, PHB vs OXC, PHB vs TPM, PHB vs GBP, PHB vs LEV, PHB vs ZNS, PHT vs GBP, GBP vs LEV (all partial seizures).

CIs for the results from indirect evidence overlapped with the CIs from direct evidence and from NMA for all comparisons.

For the following comparisons, CIs for the results from direct evidence and from NMA do not overlap which indicates potential inconsistency is present (see Table 21 and Table 22, results marked with *): PHB vs VPS (partial seizures), VPS vs TPM (generalised seizures).

For the comparison of PHB vs VPS for individuals with partial seizures, from direct evidence, there is no significant difference between the drugs (HR 0.71 (0.43 to 1.17)), however from the NMA results, a statistically significant advantage is shown for PHB (HR 1.53 (1.20 to 1.94)).

For the comparison of PHB vs VPS for individuals with partial seizures, from direct evidence, there is no significant difference between the drugs (HR 0.71 (0.43 to 1.17)), however from the NMA results, a statistically significant advantage is shown for PHB (HR 1.53 (1.20 to 1.94)). For this comparison, only 12.8% of the network estimate is contributed from direct evidence and only 80 individuals contribute to this estimate. This small sample size and imprecision for the direct evidence is likely because VPS is not considered to be a first line treatment for partial seizures and although PHB is a broad spectrum agent for the treatment of many seizure types, it is no longer used as a first line treatment [68, 74].

For the comparison of VPS vs TPM for individuals with generalised seizures, from direct evidence only, there is a statistically significant advantage to TPM (HR 0.42 (0.23 to 0.80)), however from the NMA results, the direction of effect changes to a statistically significant advantage to VPS (HR 1.30 (1.01 to 1.68)). Furthermore, for this comparison, only 21% of the network estimate is contributed from direct evidence and a moderate amount of heterogeneity is present in this estimate (I^2=46%). The same two trials contribute evidence to this outcome as 'time-to-withdrawal of allocated treatment'; see Chapter 7.2.1.3 for discussion of the differences in design of these trials.

Furthermore, the 'design-by treatment' inconsistency model does not show any significant evidence of inconsistency within the network. Therefore, we are not concerned about any impact of this observed inconsistency of numerical results on the conclusions of the review.

7.2.3 Additional analyses and sensitivity analyses

Chapter 7.1.3 and Appendix 14 provide for full details and rationale of all additional analyses and sensitivity analyses conducted. For all additional analyses and sensitivity analyses, as in primary analysis, 95% CIs were very wide for some treatment comparisons for individuals with generalised seizures, due to small numbers of participants with generalised seizures randomised to some AEDs (such as GBP).

Additional and sensitivity analyses (where appropriate) were also conducted on the two remission outcomes. There were no changes in conclusions following any of these analyses, see the Cochrane IPD-NMA for further discussion [69].

7.2.3.1 Age adjusted analysis

An additional analysis was performed also adjusting for age in the original Cox PH model. Numerical results of these analyses were similar to results of the primary analysis; mostly the same to one or two decimal places for both individuals with partial seizures and individuals with generalised seizures. There were some changes in direction of effect size and some changes in the order or 'rank' of AEDs compared to the reference treatment and there were a few changes in conclusions following this sensitivity analysis, most notably (see Appendix 14; Figure 19 and Figure 20 for all numerical results):

- For individuals with partial seizures, LEV was no longer significantly better than CBZ and CBZ became significantly better than TPM for 'Time-to-withdrawal of allocated treatment.'
- For individuals with generalised seizures, CBZ was no longer significantly better than LTG and VPS (but CBZ became significantly better than TPM) for 'Time-to-withdrawal of allocated treatment.'

It should be noted that associations between age and treatment effect for commonly used AEDs have been shown in earlier Cochrane pairwise IPD-MAs and that age is also known to be associated with epilepsy type (and in turn with misclassification of epilepsy type) [67, 296]. Therefore, the results of this sensitivity analysis are likely to overlap with the results described below in Chapter 7.2.3.4.

7.2.3.2 Validity of proportional hazards assumption

For both 'time-to-withdrawal of allocated treatment' and 'time-to-first seizure,' at least one time-varying covariate in the Cox PH model was significant, therefore a sensitivity analysis was conducted using a parametric accelerated failure time (AFT) model (see Chapter 7.1.3).

For both outcomes, numerical results of these sensitivity analyses were similar to results of the primary analysis; mostly the same to one or two decimal places for both individuals with partial seizures and individuals with generalised seizures. There were some changes in direction of effect size and some changes in the order or 'rank' of AEDs compared to the reference treatment and there were a few changes in conclusions following this sensitivity analysis, most notably (see Appendix 14, Figure 21 and Figure 22 for all numerical results):

- For individuals with partial seizures, LEV became significantly better than VPS and OXC for 'Time-to-withdrawal of allocated treatment,' making LEV significantly better than all other AEDs except for LTG.
- For individuals with partial seizures, LTG became significantly better than GBP and PHB became better than PHT for 'Time-to-first seizure,' making PHB significantly better than all other AEDs.
- For both individuals with partial seizures and individuals with generalised seizures VPS was no longer significantly better than TPM (or any other treatment) for 'Time-to-first seizure' and that LEV became significantly better than VPS.

7.2.3.3 Inconsistencies in individual participant data provided

Appendix 14 provides full details and rationale of the sensitivity analyses conducted around inconsistencies in IPD and results of these additional analyses.

The IPD from one trial (Stephen 2007) was excluded from all analyses due to inconsistencies in provided data. Numerical results of these sensitivity analyses were similar to results of the primary analysis; mostly the same to one or two decimal places for both individuals with partial seizures and individuals with generalised seizures. There were no changes in conclusions for individuals with generalised seizures. For individuals with partial seizures, there were some changes in direction of effect size and some changes in the order or 'rank' of AEDs compared to the reference treatment and there were a few changes in conclusions following this sensitivity analysis, mostly notably for 'time-to-withdrawal of allocated treatment,' CBZ became significantly better than TPM and for 'time-to-first seizure,' CBZ became significantly better than LEV and VPS became significantly better than GBP (see Appendix 14, Figure 23 and Figure 24 for all numerical results).

The IPD from two trials, Reunanen 1996 and Placencia 1993, were each excluded (separately) from analysis of 'time-to-withdrawal of allocated treatment' due to the definition of withdrawal from allocated treatment. The IPD from one trial (Banu 2007), was excluded from analysis of 'time-to-first seizure' due to inconsistencies in provided data and IPD was also excluded from one trial from the analysis of 'time-to-first seizure' (Nieto-Barrera 2001) as seizure dates for the first four weeks of the trial were not provided. For all four of these analyses, numerical results were very similar compared to the primary analysis (the same to two decimal places for individuals with partial seizures and one or two decimal places for individuals with generalised seizures) and conclusions remained unchanged for the vast

majority of results. One notable change to conclusions was that LTG was no longer significantly better than CBZ for individuals with generalised seizures for 'time-to-withdrawal of allocated treatment' following the exclusion of the IPD from Reunanen 1996 (see Appendix 14, Figure 25, Figure 26, Figure 27, and Figure 28).

7.2.3.4 Misclassification of epilepsy type

Sensitivity analyses were performed to investigate the possibility of generalised seizures being misclassified; in the first analysis those with generalised seizures and age of onset greater than 30 are reclassified as having partial-onset seizures and in the second analysis generalised seizure types and age at onset greater than 30 and those with missing epilepsy type into an 'unclassified epilepsy type' group (see Chapter 7.1.3 for further details).

For 'time-to-withdrawal of allocated treatment,' for the first analysis, numerical results for individuals with generalised seizures were similar; there were some changes in direction of effect size and some changes in the order or 'rank' of AEDs compared to the reference treatment but no change in statistical significance for any estimate and no notable change to conclusions (Appendix 14, Figure 30).

For individuals with partial seizures, most numerical results were similar but the most notable change was that PHT was now significantly better than all other AEDs (Appendix 14, Figure 29). There was a large amount of heterogeneity present in this analysis; the estimated I^2 statistic was 98% and when repeating NMA with random-effects, calculated the τ^2 statistic was 7.074 and CIs of all treatment effect estimates were very wide so that no significant differences were present between any effect sizes (Appendix 14, Figure 33). There is no clear explanation as to why this sensitivity analysis has introduced a large amount of heterogeneity into analysis for this outcome but not for the other outcomes. Due to this uncertainty, interpretation of the numerical values of this sensitivity analysis is not encouraged.

For the second analysis of epilepsy type classification, for 'time to withdrawal of allocated treatment, numerical results of this sensitivity analysis were very similar compared to the primary analysis (the same to two decimal places for individuals with partial seizures and one or two decimal places for individuals with generalised seizures) and conclusions remained unchanged (see Appendix 14, Figure 31 and Figure 32 for numerical results).

For the 'time-to-first seizure', results of these sensitivity analysis were similar to the primary analysis (see Appendix 14; Figure 29, Figure 30, Figure 31 and Figure 32 for numerical results).

7.3 Discussion

7.3.1 Summary of main results

A total of 77 trials were identified in which 17,961 individuals with partial-onset or generalised-onset tonic clonic seizures (with or without other generalised types) were randomised to one of 10 AEDs commonly used as monotherapy. IPD was provided for at least one outcome of this review for 12,391 out of 17,961 eligible participants (69% of total data) from 36 out of the 77 eligible trials (47%); see Chapter 5 for further details of IPD requests.

NMA provided a total of 45 pairwise comparisons for individuals with partial seizures and 36 pairwise comparisons for individuals with generalised seizures (no participants with generalised seizures were randomised to ZNS). Direct estimates could be calculated for between half and two thirds of comparisons across the outcomes of the review. However for many of the comparisons data was contributed by only a single trial and/or by a small number of participants. Where synthesis of head-to-head data was possible, direct evidence was generally quite consistent and where substantial heterogeneity was present between trials ($I^2 > 50\%$), it is likely that the heterogeneity originated from variability in design of the trials such as synthesis of trials recruiting different age groups, synthesis of double-blind and open label trials and pooled of trials with and without treatment stratification.

NMA showed that for the primary outcome, 'time-to-withdrawal of allocated treatment,' for individuals with partial seizures, LTG and LEV were significantly better than first line treatment CBZ, which was significantly better than GBP and PHB. LTG was significantly better than all treatments except LEV. For individuals with generalised-onset seizures, first line treatment VPS performed significantly better than CBZ, TPM and PHB.

NMA also showed that for 'time-to-first seizure' for individuals with partial seizures, PHB was significantly better than both first line treatments CBZ and LTG; first line treatment CBZ performed significantly better than VPS, GBP and first line treatment LTG and PHT also performed significantly better than LTG. In general, the earliest licenced treatments (PHT and PHB) performed better than the other treatments for both epilepsy types.

Results from NMA were more precise than results from head-to-head comparisons, often much more precise for comparisons where limited direct evidence exists, reflecting the added precision of NMA over pairwise meta-analysis. Across outcomes for the majority of pairwise comparisons, numerical results of direct evidence and NMA were similar, mostly in

the same direction and confidence intervals of estimates overlapped and there was little indication of inconsistency between direct and NMA results. For the few pairwise comparisons where confidence intervals of direct estimates and NMA estimates did not overlap, generally direct evidence was limited and contributed only a small proportion of evidence to the NMA estimates.

Despite some methodological concerns in several trials contributing to analyses which may have induced bias into analyses or inconsistencies present within IPD, numerous additional and sensitivity analyses were performed to test the robustness of the results in the presence of these biases (see Chapter 5 and Chapter 7.2.3). Results of additional and sensitivity analyses were numerically similar and did not lead to any consistent changes to conclusions, therefore it is unlikely that any methodological inadequacies of individual trials has influenced the overall pooled NMA results.

7.3.2 Strengths and weaknesses

An IPD approach was taken to analysis due to the many advantages of such a 'gold-standard' approach. Particularly within this setting, an IPD approach allowed standardisation of definitions of outcomes across trials, and attrition and reporting biases were reduced from the re-analysis of unpublished data and calculation of additional outcomes which were not considered originally within trials. Furthermore, the use of IPD in this analysis allowed the consideration of the relationship between treatment effect and epilepsy type via an interaction term in the NMA and to present results separately according to epilepsy type in the context of the recommended first line treatment of the epilepsy type, such an approach which would not have been possible without the use of IPD.

This analysis includes 69% of eligible IPD from 47% of the eligible trials. Across the ten drugs, between 49% and 100% of IPD was provided. Data for the remaining 5570 participants from 41 trials could not be provided for a variety of reasons; see Chapter 5 for further discussion. Figure 9 in Chapter 5 shows network plots of pairwise comparisons in all included trials, trials providing IPD and trials without IPD. IPD was provided for all direct pairwise comparisons in the total network except for OXC compared to VPS and OXC compared to LEV. In fact, out of all drugs included in the network, the lowest proportion of IPD was received for OXC (49%) and the lack of data for these comparisons may have contributed to imprecision of some effect sizes relating to OXC (e.g. see Figure 15). Therefore, caution should be taken when interpreting results for OXC from these analyses. However, it should be noted that the 51%

of IPD missing for OXC mostly comes from trials for which we could not establish contact with an author or sponsor to request IPD. If additional data can be included in an update for OXC, the precision of these estimates to is likely to improve.

It is inevitable that the exclusion of 31% of eligible participants may have introduced some bias into results of analyses; further discussion of differences between studies providing and not providing IPD are discussed in Chapter 5.4.3. However, it is likely that the 69% of IPD provided for analyses is a representative sample of the total participants included in all eligible trials and that the benefits of an IPD approach outweigh the limitations.

The majority of IPD requested was provided directly but for one trial randomising 136 participants [319], data was requested via data sharing portal CSDR [91] and was provided via a remote secure data access system which allowed analysis in SAS based statistical software and export of analysis results. IPD from this trial could not be included with the entire individual participant dataset to fit the models outlined in Chapter 6.2 , therefore the results exported from the data access system were treated as AD in an additional analysis (see Chapter 8 for further discussion and results).

7.3.3 Clinical implications and relation to other studies

The majority of participants recruited into these trials were classified as experiencing partial seizures (66.8% of participants in all trials and 67.5% of participants with IPD provided); this majority is emphasised in the visual similarity of the network plot for individuals with partial seizures compared to the plot of all participants and reflected in the relative precision of the results of this review for partial seizures compared to generalised seizures (Figure 12). While a majority of partial seizures compared to generalised seizures is reflective of clinical practice (around 60% of individuals with epilepsy experience partial seizures [320]), the proportion of individuals with partial seizures recruited to the trials in this review is even greater.

The remaining participants were classified as experiencing generalised tonic-clonic seizures with or without other generalised seizure types (24.4% of participants in all trials and 26.5% of participants with IPD provided) or unclassified / missing epilepsy type (8.8% of participants in all trials and 6% of participants with IPD provided). Misclassification of epilepsy type is a recognised problem in epilepsy (whereby some individuals with generalised seizures have been mistakenly classed as having partial-onset seizures and vice-versa). The potential impact of this misclassification on results has been shown in our series of Cochrane IPD

reviews of monotherapy for epilepsy [59]. Investigation of misclassification within this analysis (reclassification of 1,164 participants with generalised seizures and age of onset of over 30 years, 36% of individuals originally classified at experiencing generalised seizures) did not show any important changes to treatment effect sizes and no changes to conclusions.

This does not, however, indicate that misclassification of epilepsy type has not occurred in these trials; rather that the primary analysis results are robust to any misclassification. Trials included in this analysis were published between 1981 and 2015 and a proportion of trials classified generalised and partial-onset seizures according to the ILAE classification of 1981 [321], rather than the revised classification in 1989 [322] or recently revised terminology [323], which may have led to misclassification. Furthermore, several trials were conducted in developing countries in Africa, Asia and Central or South America without access to the same facilities such as electroencephalograms (EEGs) or magnetic resonance image (MRI) scanners as trials conducted in the USA and Europe. Within these trials, it is likely that seizure type would have been classified clinically, which may have further contributed to misclassification.

In reality, it is likely that fewer than 20% of participants recruited into all of these trials experienced generalised seizures (17% of participants included in IPD analysis were classified as having generalised seizures following reclassification in sensitivity analysis), which is a lower proportion than would be expected in clinical practice [320]. For this reason, treatment effect sizes for generalised seizures, particularly those which are imprecise, should be treated as less applicable than the treatment effect sizes for partial seizures.

In order to provide more precise evidence, applicable to individuals with generalised seizures, it is important both ensure accurate seizure classification (as far as possible) and to increase the proportion of individuals with generalised seizures recruited into trials of AEDs to better reflect the 'real world' ratio of partial to generalised seizures. Increased recruitment of may not be straightforward, particularly as those with new onset generalised seizures are expected to be children and adolescents and recruitment of children into clinical trials comes with difficulties [324]. However, if targeted recruitment strategies could be implemented and the evidence base for individuals with generalised seizures increased this may better inform treatment decisions for this population, particularly for those of childbearing potential, for whom first line treatment sodium valproate may not be appropriate [74].

An NMA was published by representatives of the Cochrane Epilepsy Group in 2007 including IPD for over 6418 patients from 20 trials (also included in the current review) comparing

direct and indirect evidence from CBZ, PHB, PHT, VPS, LTG, OXC, TPM and GBP [285]. Results of this NMA showed for partial-onset seizures, LTG performed better than all other drugs in terms of treatment withdrawal but may not perform better than CBZ in terms of seizure control. PHB performed better than other drugs in terms of seizure control but at the expense of increased treatment failure. Overall for individuals with partial seizures; LTG, CBZ and OXC seemed to provide the best balance of seizure control and treatment failure. As in the current review, data for individuals with generalised seizures was limited and results suggested that VPS or PHT may provide the best combination of seizure control and treatment failure.

The present analysis was designed to update the information in the previous NMA with new evidence from trials published since 2007 and including evidence for two drugs which were licensed for use as monotherapy after 2007 (LEV and ZNS) [68]. The results of the present analysis generally agree with the results of the previous NMA in addition to providing evidence of the comparative effectiveness of the two new drugs within the spectrum of commonly used AEDs and further highlight that nearly 10 years on, data for individuals with generalised seizures is still limited.

7.3.4 Concluding remarks

Results of this analysis demonstrate that generally the earliest licenced AEDs such as PHT and PHB provide increased seizure control, in terms of delaying recurrence of first seizure and earlier remission, compared to newer AEDs. However, this comes at the expense of earlier treatment failure and it is newer AEDs such as LTG and LEV that perform the best in terms of treatment retention. Considering the optimum balance of efficacy (seizure control) and tolerability (treatment retention), for individuals with partial seizures, CBZ, LTG and LEV seem to be the best treatment options whereas for individuals with generalised tonic-clonic seizures (with or without other seizure types); VPS, LTG and LEV seem to be the best treatment options. ZNS, the most recently licenced AED for monotherapy treatment, may be an effective treatment option for individuals with partial-onset seizures; however further evidence from randomised controlled trials is needed and the effectiveness of this drug has yet to be evaluated in a published clinical trial for individuals with generalised seizures.

Overall, these results support the NICE guidelines that CBZ and LTG are suitable first-line treatments for individuals with partial-onset seizures and also demonstrates that LEV may be a suitable alternative. Results also support the use of VPS as the first-line treatment for individuals with generalised tonic-clonic seizures (with or without other seizure types) and

also demonstrates that LTG and LEV would be suitable alternative first-line treatments, particularly for those of child bearing potential, for whom VPS may not be an appropriate treatment option. Evidence for the relative effectiveness of other AEDs for individuals with generalised seizures is limited and of moderate quality; further evidence from randomised controlled trials is needed.

This review highlights the need for the design of future AED monotherapy trials that are well powered to detect a difference between particular AEDs while recruiting a sample of individuals' representative of the wider population in terms of age and seizure type. An approach to best reflect and inform clinical practice, as well as being statistically powerful, would be to recruit heterogeneous populations for whom epilepsy syndromes have been adequately defined, with testing for interaction between treatment and epilepsy syndrome. In view of potential problems of misclassification, syndromes will have to be well defined, with adequate checking mechanisms to ensure that classifications are accurate and a system to recognise uncertainty surrounding epilepsy syndromes in individuals within trials.

The choice of outcomes at the design stage of a trial and the presentation of the results of outcomes, particularly of a TTE nature, require very careful consideration. While the majority of trials of a monotherapy design do record and report outcomes measuring efficacy and tolerability of AEDs, there is little uniformity between the definition of the outcomes and the reporting of the summary statistics related to the outcomes [325] (see Chapter 3 of this thesis) making an AD approach to meta-analysis in reviews of monotherapy trials impossible. Where trial authors cannot or will not make IPD available for analysis, excluding a proportion of relevant evidence from the review was unavoidable but will inevitably have some impact upon the interpretation of results of the review and applicability of the evidence and conclusions. The ILAE recommends that trials of a monotherapy design should adopt a primary effectiveness outcome of 'time-to-withdrawal of allocated treatment' and should be of a duration of at least 48 weeks to allow for assessment of longer-term outcomes, such as remission [249]. If trials followed these recommendations, an AD approach to meta-analysis may be feasible, reducing the resources and time required from an IPD approach.

Chapter 8: Combining individual participant data with aggregate data in network meta-analysis

8.1 Introduction

8.1.1 Individual participant level compared to aggregate level approaches to meta-analysis

As outlined in Chapter 1.1.3, previous work has demonstrated that meta-analyses of the same studies taking IPD or AD approaches can produce different results [23, 43, 52-54]. A recent systematic review examined 39 meta-analyses with 190 comparisons taking both IPD and AD approaches for meta-analysis of the same studies [52]. Results of this systematic review showed that for 38 comparisons (20%) there was a disagreement in statistical significance between the IPD and AD approach to meta-analysis, with more IPD-MAs detecting a statistically significant result which was not confirmed in the AD-MA.

Conclusions from this work recommend that before embarking upon a resource intensive IPD approach to analysis, researchers should carefully consider the added benefits of IPD to their clinical question, and consider whether a less resource intensive AD-MA could provide an adequate (and mathematically equivalent) answer to the question [52, 53].

Specifically from a TTE setting, Haines and Hill [53] have demonstrated that in the context of repeated-measures data of accidental falls, a range of approaches may be taken to the statistical analysis at a trial level which influence the standard errors of the results and hence would influence the pooled result should these estimates be combined in AD-MA. The authors therefore argue that IPD-MA and AD-MA would fundamentally produce difference results in this setting and question whether AD-MA would ever be appropriate.

Similarly, Duchateau *et al* [164, 165] note differing results for TTE outcomes from IPD-MAs and AD-MAs in head and neck cancer, concluding that the most likely reason for the differences is due to IPD-MAs being based on exact TTE analyses whereas AD-MAs are based on mortality as a specific time point. The authors also note that where AD results may have been indirectly estimated from a Kaplan-Meier (KM) curve, the estimated number of events is likely to be an over-estimation of the true number of events reflected in the IPD, with the extent of overestimation increasing as an increasing number of patients are censored.

It is argued that where treatment-covariate interactions are of interest to meta-analysis that an IPD approach (i.e. modelling the interaction as a parameter within an IPD model) is generally superior to an AD approach (i.e. meta-regression) due to limitations of the latter approach such as low statistical power and ecological biases. In other words, study-level associations may not accurately reflect individual-level associations [26, 309, 326, 327].

Simmonds and Higgins [309] define three approaches for the investigation of treatment-covariate interactions in meta-analysis; a full IPD approach incorporating a treatment-covariate interaction term into a model, a meta-regression approach (i.e. a fully AD approach) and an intermediate 'meta-analysis of interactions' approach which uses IPD to estimate within-study treatment-covariate interactions and then combines the separate estimates for each study using standard meta-analysis techniques. The authors demonstrate that theoretically if the IPD model is specified correctly and assuming normally distributed participant responses, a full IPD approach will always have at least as much statistical power to detect treatment-covariate interactions as meta-regression or meta-analysis of interactions. Furthermore, the power of the latter two approaches depends on the distribution and heterogeneity of the covariate of interest. The authors also derive a series of Q statistics based on the distributions and heterogeneity of covariates to allow comparison of the power of the three approaches for detecting treatment-covariate interaction and potentially guiding a choice between the methodological approaches.

Although an IPD approach is still relatively rare within network meta-analysis (NMA) compared to an AD approach [18, 19, 44-47], several authors have highlighted the benefits of an IPD approach to NMA, particularly where detailed examination of heterogeneity, inconsistency and treatment-covariate interactions are of interest [19-21, 44-47, 328].

8.1.2 Combining IPD and AD in meta-analysis

As an alternative approach to performing either a complete IPD analysis or a complete AD analysis, methods have been developed which allow the combination of IPD and AD in meta-analysis [38-41, 329, 330] and network meta-analysis [19, 20, 44-46]; the latter are further discussed in Chapter 8.1.3.

Such methods have generally been developed for two general reasons; firstly, to increase the power and precision of an AD-MA by incorporating participant-level information from IPD [19, 39, 46] or secondly for the scenario when an IPD approach to analysis was intended but IPD is not available from a subset of trials for analysis. Where AD is available for some or all of these trials, this AD could supplement the IPD in analysis, allowing a larger (and more

complete) proportion of the relevant evidence base to be included in analysis [40]. These methods are particularly appealing where there is a concern that unavailability of IPD for a study may be related to the results of that study, hence potentially introducing bias into IPD-MA [24, 51]. It should, however, be noted that published AD may be prone to bias and not of the same quality as IPD [51, 331] so while incorporation of additional AD may increase precision, it may not necessarily improve the reliability of the overall analysis results. It is therefore recommended that sensitivity analyses are conducted comparing analyses of IPD only and IPD combined with AD in meta-analysis and consideration is given to potential reasons for unavailability of IPD and the relative quality of any published AD [36, 41, 51].

A systematic review of methods employed in IPD-MA conducted by Riley *et al* [40] showed that out of 199 applied IPD-MA identified, 33 published articles combined IPD and AD in meta-analysis and 30 clearly described the methods used to do this. In the majority of articles (27 out of 33 meta-analyses, 82%), IPD and AD were combined via a two-stage method to meta-analysis [32, 141, 142]; in other words, IPD were reduced to AD and combined with additional AD using standard meta-analytic techniques (see Chapter 2.3.2 for further description). Two-stage approaches have the potential disadvantage of losing participant-level information provided within the IPD, but as discussed further in Chapter 2.3.2, two-stage approaches and one-stage approaches often do produce the same results.

Further, specifically within a TTE setting, it has been shown that employing a two-stage method to combine IPD and AD in meta-analysis, where feasible, can have advantages over an IPD only approach such as increasing precision of resulting pooled estimates [141, 142].

Where treatment-covariate interactions are of interest, reducing IPD to AD for meta-analysis has the potential for ecological bias. In this case, the two-stage 'meta-analysis of interactions' approach defined by Simmonds and Higgins [309] may be applicable (see Chapter 6.2.4 and Chapter 8.1.1 for further details). The authors argue that this approach which remains within familiar meta-analysis frameworks may be more readily understood than complete IPD approaches; however, this approach would not readily extend to incorporation of AD with IPD, unless treatment-covariate interaction at an AD-level is available.

The remaining three articles identified by Riley *et al* [40] combined IPD and 'partially reconstructed IPD' which could be extracted from published literature, i.e. reconstruction of 2 x 2 tables for binary outcomes [41] or estimated survival times from KM curves for TTE outcomes [153, 157]. Messori *et al* [153] demonstrate that reconstructed IPD from KM curves can act as a good 'intermediate' method with advantages over an AD approach where

IPD is not available for every study or where resources to take a fully IPD approach are not available. The illustrative example of Messori *et al* [153] shows good correlation between estimated IPD and true IPD; however, it must be noted that reconstructed IPD may be less reliable than original IPD and may still be prone to ecological biases [40].

Although not used in any of the articles identified in Riley *et al* [40], multi-level or 'hierarchical related' regression modelling approaches allow for IPD and AD to be combined in the same meta-analysis model [38-41, 326, 329, 330, 332]; some examples were developed in the context of combining summary (e.g. geographic area) level and individual-level data from ecological studies rather than in a clinical context [326, 332]. However, such methodology is also applicable to evidence synthesis from a clinical setting.

This modelling approach allows the simultaneous or joint estimation of the likelihood from related regression models of each data source (IPD or AD) and a multi-level structure which allows the incorporation of AD alongside IPD via dummy variables to indicate the data source [38, 40, 329]. Such a structure allows for both data sources to contribute to overall treatment effect and any study-level covariates of interest but ensures that only studies providing IPD contribute to any participant-level covariates [40]. Simultaneous hierarchical estimation of related models for IPD and AD sources, which typically share common parameters, allows for both data sources to inform the common parameters which has the advantage of potentially reducing biases from both data sources; i.e. the inclusion of IPD may reduce ecological biases arising from AD and the combined analysis of IPD and AD together may increase statistical power, which may be particularly beneficial where only a small proportion of IPD is available [40, 326, 332]. However, it should be noted that applying models of the same form with treatment-covariate interactions to IPD and AD sources will only result in valid estimates of treatment effect according to the level of covariate if all individuals in the study have the same covariate value or if the relative effect modification of the covariate is the same at the individual and aggregate-level [19, 41].

Ravva *et al* [330] note that applying the same non-linear model to both IPD and AD sources to define common parameters may lead to 'aggregation bias,' a type of ecological bias where between-study effects are incorrectly interpreted as within-study effects. The authors describe a hierarchical linearization modelling technique and an application to pharmaceutical drug development (dose-response) to address this issue of aggregation bias by allowing AD model parameters to retain their original definition with respect to treatment-covariate interactions at the individual-level.

Specifically, meta-analysis models for combining IPD and AD have been developed for binary outcomes, modelling IPD and AD 'event-risk' responses as Bernoulli and Binomial distributions respectively [39, 41]. Such models provide pooled 'event-risk' estimates across studies, while accounting for within-study variability due to interactions between 'event-risk' and participant-level covariates of interest and can be applied within a Frequentist or Bayesian framework [39, 41].

Meta-analysis models for combining IPD and AD have also been developed by Goldstein *et al* [329], modelling a continuous response variable via multi-level linear or quadratic regression. Riley *et al* [38] build on the approach of Goldstein *et al* [329] and outline a series of one and two-step meta-analysis models to combine IPD and AD for continuous outcomes and demonstrate how these models can be used to incorporate participant-level covariates to estimate treatment-covariate interactions in relation to pooled treatment effect and between-study heterogeneity. The general framework of the one-step approach also extends to other data types (such as TTE data), multiple covariates, multiple correlated outcomes, non-linear interaction effects and incorporation of treatment-covariate interactions from the trials providing AD.

Recent work emphasises the importance of parameter specification when fitting one-stage meta-analysis models, whether analysing IPD only or incorporating IPD and AD, to ensure that within-study and between-study associations are separated to avoid inadvertent ecological biases [25, 38, 41].

8.1.3 Combining IPD and AD in network meta-analysis

As described in Chapter 8.1.2 for 'traditional' pairwise meta-analysis, a range of methods have been developed and described for combining AD with IPD in meta-analysis of various data types, allowing for the incorporation of treatment-covariate interactions. However, at the time of writing, methodology to combine IPD and AD in NMA across different data types and scenarios has been less widely researched. Such methods are arguably even more important in this setting, where more treatment comparisons are made across more eligible studies, the scope for unavailable IPD from a subset of studies is potentially larger [46].

Saramago *et al* [45] and Donegan *et al* [46] have described similar models for combining IPD and AD which extend the methodology of Sutton *et al* [39] for meta-analysis of a binary outcome to NMA of a binary outcome. Both approaches are performed in a Bayesian framework and extend to fixed or random-effects.

Saramago *et al* [45] describe NMA models for a binary outcome for scenarios where only IPD, only AD or a combination of IPD and AD are available for NMA. These models also allow the incorporation of individual or aggregate-level covariate data as treatment-covariate interactions and to explore within-study and across-study interactions.

Donegan *et al* [46] describe an approach for the joint synthesis of IPD and AD within NMA for a binary outcome via a 'shared parameter model' [46, 333], where available IPD and AD are entered as two separate datasets in a single model which allows both datasets to contribute to the shared model parameters. The model also extends to include multi-arm trials, treatment-covariate interactions and to explore within-study and across-study interactions (assuming independent, exchangeable or common treatment-covariate interactions for each treatment comparison). However, the authors experienced problems with model convergence when applying the proposed methods to explore within-study and across-study interactions to an illustrative example examining treatment success in malaria with a treatment-by-age interaction, likely due to the similarity in mean age across sites.

Both Saramago *et al* [45] and Donegan *et al* [46] conclude that the synthesis of IPD and AD to include as much available evidence as possible increases precision and the use of IPD in NMA has advantages over the usual AD approach to NMA.

Jansen and Cope [20] present an aggregate-level approach to NMA which allows for the incorporation of study-level covariates to adjust for confounding bias due to heterogeneity and inconsistency. This method is an extension of a method described in Chapter 2 [158] for NMA of TTE data, which is a multi-dimensional approach which does not require a PH assumption. The authors note the limitations of this AD approach that the methods described do not reflect individual level-effect modification. In a related publication, Jansen [19] hypothetically demonstrates the potential differences in NMA results without covariate adjustment, with covariate adjustment using AD and with covariate adjustment using IPD.

Jansen [19] presents two methods for the incorporation of IPD and AD in NMA for a binary outcome using non-linear models in the presence of a participant-level covariate. The first method can be considered an extension of the methods of Sutton *et al* [39] (outlined in Chapter 8.1.2) to NMA (combined with the methods of Cooper *et al* [21]). The second method uses a 'hierarchical related regression' as introduced by Jackson *et al* [326, 332] (also outlined in Chapter 8.1.2) which derives the AD model by integrating the underlying IPD model over the joint within-study distribution of covariates. Results of a simulation study show that under the majority of scenarios, the NMA models incorporating IPD and AD were

less associated with confounding bias than the NMA models using AD alone and generally the hierarchical related regression approach to combining IPD and AD in NMA was associated with less confounding bias than the extended approach of Sutton *et al* [39]. Jansen [19] concludes that the incorporation of even a small proportion of IPD into an AD-NMA to model participant-level covariates has values, particularly where there are concerns regarding heterogeneity, inconsistency or confounding bias.

Saramago *et al* [44] describe methodology for combining IPD and AD in NMA in a TTE setting; which extends the work of Sutton *et al* [39] and Saramago *et al* [45] for binary outcomes. Specifically the methods proposed by Saramago *et al* [44] combine individual event-time data with aggregated count data under the assumption that by specifying a parametric TTE distribution which can allow for HRs to be generated from the original count AD by estimating the cumulative hazard in each trial arm reporting count data for a given follow-up time [221]. The methods described also allow for the incorporation of treatment-covariate interactions from both IPD and AD sources.

8.1.4 Objective

The first objective of this Chapter is to directly extract aggregate TTE data for the outcomes of interest to the Cochrane Epilepsy IPD-NMA (outlined in Chapter 7) from the trials not providing IPD or to determine whether suitable AD can be estimated from other published summary statistics (as outlined in Chapter 2.3.2).

The second objective of this Chapter is to perform a combined NMA of IPD and any extracted or estimated AD and to compare results to those of the IPD-NMA. The principle aim of the combined analysis is to investigate whether the incorporation of AD changes the results and conclusions of the IPD-NMA, which could indicate that the 69% of IPD obtained for the NMA may not be representative of the entire evidence base.

8.2 Methods

8.2.1 Extraction of aggregate data from epilepsy studies

As outlined in Chapter 5 and Table 13, IPD was not provided from 5570 participants from 41 trials which were eligible for the Cochrane NMA (31% of total participant data). As noted in Chapter 5.3.2, if IPD was not available, any unpublished AD related to the outcomes of interest of the review was requested and would have been used in a combined analysis if any useable unpublished AD had been provided.

For each trial without IPD available, published literature (journal articles, clinicaltrials.gov entries etc.) were examined to determine whether any relevant AD could be extracted from this literature either directly or indirectly estimated (as described in Chapter 2.3.2 of this thesis). Aggregate data was extracted and/or estimated in the following order of preference:

- direct numerical estimates (e.g. a HR and measure of precision could be extracted for one or more outcomes of interest)
- indirect numerical estimates (see Chapter 2.3.2.1)
- indirect graphical estimates with numbers at risk provided (see Chapter 2.3.2.2)
- indirect graphical estimates without numbers at risk provided but estimated according to methods of Parmar *et al* [6] (see Chapter 2.3.2.2)

For graphical estimation, survival proportions were extracted by hand by SJN from an enlarged version of the published curve at an appropriate range of time points, according to the extent of follow-up of the trial or the intervals at which numbers at risk were reported. Indirect estimation was performed by entering extracted summary statistics or survival proportions into the macro-enabled spreadsheet developed by Tierney *et al* [128].

As outlined in Chapter 7.3.2, IPD was provided for one trial randomising 136 participants (referred to as Biton 2001 [319]) via a remote secure data access system which allowed analysis in SAS based statistical software and export of analysis results. IPD from this trial could not be included with the entire individual participant dataset to fit the models outlined in Chapter 6.2, therefore the results exported from the data access system were treated as AD in an additional analysis.

None of the 41 studies reported AD of any kind for the outcomes of time-to-12-month or time-to-6 month remission and none of the studies appeared to have measured either of the outcomes. The trial duration of 19 trials was less than 12 months and less than 6 months for five trials, therefore these remissions outcomes were not in the scope of the trials. Furthermore, the range of follow-up was not reported in seven trials so it was unclear whether time-to-12-month or time-to-6 month remission could have been measured. For one study (Biton 2001), AD for time-to-6 month remission for all participants and by epilepsy type could be calculated from IPD provided within the remote data access system. Results of combining AD from this single study to the IPD for time-to-6-month remission did not change conclusions and for brevity, results are not reported here. This trial was of less than a year duration so time-to-12-month remission was not in scope.

Table 23 and Table 24 summarise the aggregate data which could be extracted for the outcomes of interest to the Cochrane Epilepsy IPD-NMA.

Table 23: Aggregate data available for time-to-withdrawal of allocated treatment

Trial[1]	Number of participants[2]	Number of events	AD available[3]
Biton 2001*	136 (P=82; G=46)	32 (P=18; G=16)	Summary HR and 95% CI available (SAS remote data analysis system)
Brodie 2002	291	99	Summary HR and 95% CI available
Christie 1997	249	93	Graphical (no numbers at risk)
Forsythe 1991	64	22	Extracted approximate IPD (Table)
Gilad 2007*	64 (all partial)	11	Extracted approximate IPD (Graphical)
Rowan 2005	590	314	Graphical (partial numbers at risk available at yearly time points)
Saetre 2007	184	55	Summary HR and 95% CI available
Shakir 1981*	33 (P=14; G=19)	9 (P=3; G=6)	Extracted approximate IPD (Table)
Steinhoff 2005*	239 (P=176; G=63)	62 (P=53; G=9)	Graphical (no numbers at risk)

AD = aggregate data; G=generalised seizures; P = partial seizures, * = AD available by epilepsy type

1. See Appendix 10 for reference of the primary publication of each trial
2. Number of participants in the evaluable population or included in analysis for the outcome; for Biton 2001; epilepsy type was missing for eight participants
3. For 32 out of 41 trials (3720 participants, 21% of total participant data), 'time-to-withdrawal of allocated treatment' was not reported as an outcome of the trial so no AD was available.

For the outcome 'time-to-withdrawal of allocated treatment,' nine trials reported AD for 1850 participants across a range of drug comparisons; mostly of CBZ, LTG and VPS but also GBP, OXC and PHT. For 'time-to-first seizure,' six trials reported AD for 1369 participants across a range of drug comparisons; mostly CBZ and LTG but also GBP, LEV and VPS. For three trials, a summary HR and 95% CI were available for both outcomes and for one trial, summary statistics were estimated for both outcomes from published survival curves with partial numbers at risk available at yearly intervals).

For 'time-to-withdrawal of allocated treatment,' summary statistics could be estimated from two further studies with published survival curves without published numbers at risk. For the remaining trials, approximate IPD could be estimated from tables or graphs and used to calculate a summary HR and 95% CI. In two studies, due to the small number of events and very clear graphics, approximate event times could be extracted for each participant for both outcomes (Gilad 2007) or for time-to-first seizure (Consoli 2012).

Table 24: Aggregate data available for time-to-first seizure

Trial[1]	Number of participants[2]	Number of events[3]	AD available[4,5]
Biton 2001*	136 (P=82; G=46)	71 (P=44; G=27)	Summary HR and 95% CI available (SAS remote data analysis system)
Brodie 2002	291	138[3]	Summary HR and 95% CI available
Consoli 2012	104	9	Extracted approximate IPD (Graphical)
Gilad 2007*	64 (all partial)	26	Extracted approximate IPD (Graphical)
Rowan 2005	590	305[3]	Graphical (partial numbers at risk available at yearly time points)
Saetre 2007	184	84	Summary HR and 95% CI available

AD = aggregate data; G=generalised seizures; P = partial seizures, * = AD available by epilepsy type

1. See Appendix 10 for reference of the primary publication of each trial
2. Number of participants in the evaluable population or included in analysis for the outcome; for Biton 2001; epilepsy type was missing for eight participants
3. For two trials, the number of events was not reported and was estimated based on the reported proportions seizure free at the end of the study.
4. For 32 out of 41 trials (3022 participants, 17% of total participant data), 'time-to-first seizure' was not reported as an outcome of the trial so no AD was available.
5. For 3 out of 41 trials (1179 participants, 7% of total participant data), a 'time-to-first seizure' outcome was defined but was reported as mean or median time-to-first seizure or number of events only with no further statistical analysis so no usable AD was available.

Two studies presented times at which allocated drug was withdrawn and the reason for withdrawal in a table. Shakir 1981 presented 'time on trial drug' in months for each participant; therefore to calculate 'time-to-withdrawal of allocated treatment,' it was assumed, for example, that if 'time spent on trial drug' was five months, the individual spent five full months (152 full days) on the trial drug before withdrawal. Forsythe 1991 presented 'withdrawal and time of occurrence by month' and therefore to calculate 'Time-to-withdrawal of allocated treatment' we assumed that, for example, if withdrawal occurred during the fifth month, that withdrawal occurred halfway between the fifth and sixth month (i.e. participants spent 167 full days on treatment before withdrawal). This approach to analysis of these trials was taken in a pairwise IPD-MA including these two trials [59]; within that IPD-MA, sensitivity analysis was conducted examining the assumptions made of the withdrawal times in these trials. Results were similar following sensitivity analysis, therefore it was assumed that these assumptions were reasonable for the NMA.

It should be noted that the 'approximate' IPD which could be extracted could have been included as IPD in the complete IPD approach to the NMA. However, given the potential for ecological bias as noted by Riley *et al* [40], compared to IPD which was provided directly and

consistency checked (see Chapter 5), the 'approximate' IPD was reduced to a summary HR and 95% CI and treated as AD for the purpose of this analysis.

In relation to the findings of Chapter 3, the majority of trials did not report the TTE outcomes of interest for the Cochrane IPD-NMA (rather than outcomes were reported inconsistently). No trials reported remission as a TTE outcome in the trial publication. For the 32 trials which did not report relevant information for 'time-to-withdrawal of allocated treatment' (21% of total participant data), withdrawal information was generally reported as the proportion of participants withdrawing, rather than as a TTE outcome. Also, for 32 trials without any AD for 'time-to-first seizure' (17% of total participant data), the outcome was reported as the proportion of participants with seizure freedom or the change in seizure frequency rather than as a TTE outcome. However, for three trials (Korean LTG Study Group 2008, NCT01498822 and NCT01954121, recruiting 7% of total participant data), a 'time-to first seizure' outcome was defined but reported as the mean or median time-to-first seizure or the number of events only. Therefore this published data could not be used.

8.2.2 Methods for combining IPD and AD in NMA

As outlined in the objective (Chapter 8.1.4), the principle aim was to allow comparison of an NMA of combined IPD and AD to an NMA of IPD only to investigate whether the IPD-NMA is representative of the evidence base and to examine the robustness of results. To allow this comparison, methods for combining IPD and AD must use the same framework as the IPD only analysis (see Chapter 6.2). Use of a different approach to modelling (e.g. via a Bayesian framework [45]) would likely produce different numerical results to the IPD-NMA due to methodological differences, which may confound the impact of the AD on the IPD results.

Table 23 and Table 24 show that for four of the trials, aggregate summary statistics could be extracted or estimated according to epilepsy type but none of these aggregate summary statistics came from statistical models accounting for a treatment-by-epilepsy type interaction. Methods described in Chapter 6.2 of this thesis require that the epilepsy type of each individual was available, so, it would not be appropriate to combine summary statistics for all participants (regardless of epilepsy type) with IPD in these models.

Therefore the following approach was taken to allow IPD and AD to be combined in an NMA under a multivariate framework. Firstly, IPD was reduced to summary statistics separately by epilepsy type. In other words, models of the structures outlined in Equation 29 and Equation

30 (see Chapter 6.2.2) are fitted separately to each trial, producing separate trial-specific summary statistics of treatment effect for individuals with partial seizures and individuals with generalised seizures.

The summary statistics for each epilepsy type from each trial (estimated from the IPD) are then combined in separate NMAs by epilepsy type. This was achieved by producing a dataset of the summary statistics structured as a list of pairwise comparisons and converted from 'pairs' to 'augmented' format via the 'network' command within Stata version 14 [310] (see Appendix 13) and NMA is performed via 'mvmeta' as described in Chapter 6.2.5.

Secondly, these separate trial-specific summary statistics of treatment effect for individuals with partial seizures and individuals with generalised seizures estimated from the IPD are also combined with additional summary statistics extracted or estimated from published study reports. These summary statistics (estimated from the IPD and combined with additional AD) are then synthesised as described in the previous paragraph.

The results of this NMA using IPD only and the NMA with IPD and AD combined for each seizure type are then compared. For completeness, these results are also compared to the results of the IPD-NMAs from the models outlined in Chapter 6.2.2.

Table 23 and Table 24 also show that the majority of aggregate summary data available related to all participants in the trial, rather than separated by epilepsy type (see Chapter 8.3 for further discussion). Therefore to allow further investigation of any 'availability bias' in the results of the IPD-NMA, to incorporate as much additional published AD as possible, an additional analysis was conducted without separating epilepsy type.

In this approach, IPD was reduced to summary statistics without accounting for epilepsy type. In other words, a model of the structures outlined in Equation 29 (see Chapter 6.2.2) was fitted separately to each trial, producing separate trial-specific summary statistics of treatment effect for all individuals (regardless of epilepsy type). These summary statistics (firstly those estimated from the IPD only and secondly those estimated from the IPD combined with additional AD) are then synthesised in an NMA as described above.

It should be noted that results that do not accounting for epilepsy type, are of little clinical relevance given the known differences between AEDs in different epilepsy types and current clinical practice (see Chapter 1.2.1 for further details).

8.3 Results

As described in Chapter 6.2, 11,865 participants (66% of eligible participant data) contributed to the main analysis of 'time-to-withdrawal of allocated treatment' with a total of 4109 withdrawal events and 12,152 participants (68% of eligible participant data) contributed to the main analysis of 'time-to-first seizure' with a total of 6453 first seizure events.

An additional 336 participants (85 events) with partial seizures from four trials and an additional 128 participants (31 events) with generalised seizures from two trials contributed to 'time-to-withdrawal of allocated treatment.' In other words, an additional 2% of data was contributed by AD for individuals with partial seizures and an additional 0.7% of data for individuals with generalised seizures. An additional 146 participants (53 events) with partial seizures from two trials and an additional 46 participants (27 events) with generalised seizures from one trial contributed to for 'time-to-first seizure.' In other words, an additional 0.8% of data was contributed by AD for individuals with partial seizures and an additional 0.3% of data for individuals with generalised seizures. An additional 1850 participants (697 events) from nine trials contributed to 'time-to -withdrawal of allocated treatment' (i.e. an additional 10% of data) and an additional 1369 participants (633 events) from six trials contributed to 'time-to-first seizure' (i.e. an additional 8% of data).

NMA results compared to reference treatments and most commonly used treatments are discussed in the remainder of this section; results from all methods for all pairwise comparisons for both outcomes are available in Appendix 17.

8.3.1. Individuals with partial seizures

Results for all AEDs compared to reference treatment CBZ from NMAs of IPD only and IPD combined with AD (as outlined in Chapter 8.2.2) are presented in Table 25 and Table 26 respectively for 'time-to-withdrawal of allocated treatment' and 'time-to-first seizure.'

Comparing the results of the NMA which combines IPD and AD to the NMAs with IPD only, NMA results are very numerically similar, except for one change to statistical significance when incorporating AD to the comparison of CBZ vs TPM for 'time-to-withdrawal of allocated treatment' even though no additional AD for this comparison directly was added. It should also be noted that the results of all models for this comparison are the same to two decimal places. Overall, incorporating AD with IPD in NMA has had very little impact and the lack of notable difference is not surprising given the small amount of additional AD available by seizure type, i.e. adding AD contributes only up to an extra 2% of data to the outcomes.

Table 25: Network meta-analysis results also incorporating aggregate data: Time-to-withdrawal of allocated treatment for individuals with partial seizures.

Comparison[1]	IPD only[1,2]	IPD reduced to AD[1,3]	IPD reduced to AD, plus AD[1,3]
CBZ vs PHB	1.57 (1.20 to 2.05)	1.58 (1.21 to 2.07)	1.58 (1.21 to 2.07)
CBZ vs PHT	1.16 (0.93 to 1.45)	1.18 (0.94 to 1.48)	1.19 (0.95 to 1.49)
CBZ vs VPS	1.10 (0.90 to 1.35)	1.09 (0.90 to 1.32)	1.05 (0.88 to 1.25)
CBZ vs LTG	0.72 (0.63 to 0.83)	0.72 (0.63 to 0.83)	0.73 (0.65 to 0.84)
CBZ vs OXC	1.07 (0.84 to 1.37)	1.02 (0.81 to 1.28)	1.02 (0.82 to 1.28)
CBZ vs TPM	1.17 (0.99 to 1.38)	1.17 (0.99 to 1.38)	***1.17 (1.00 to 1.38)***
CBZ vs GBP	1.18 (1.01 to 1.39)	1.18 (1.00 to 1.38)	1.18 (1.01 to 1.39)
CBZ vs LEV	0.83 (0.70 to 0.99)	0.84 (0.71 to 0.99)	0.84 (0.71 to 0.99)
CBZ vs ZNS	1.08 (0.81 to 1.44)	1.08 (0.81 to 1.44)	1.08 (0.81 to 1.44)

1. Results presented are HR and 95% CIs, See Chapter 5.3.2 for abbreviations of drugs. Results highlighted in bold italics show a difference in statistical significance when published AD is combined with IPD compared to both of the IPD only models.
2. Results taken from the model outlined in Chapter 6.2.2
3. Results taken from the model outlined in Chapter 8.2.2

Table 26: Network meta-analysis results also incorporating aggregate data. Time-to-first seizure for individuals with partial seizures.

Comparison[1]	IPD only[1,2]	IPD reduced to AD[1,3]	IPD reduced to AD, plus AD[1,3]
CBZ vs PHB	0.77 (0.60 to 0.99)	0.80 (0.63 to 1.01)	0.79 (0.60 to 1.02)
CBZ vs PHT	0.97 (0.80 to 1.16)	1.05 (0.87 to 1.26)	1.03 (0.84 to 1.27)
CBZ vs VPS	1.19 (1.00 to 1.43)	1.28 (1.10 to 1.49)	1.21 (1.02 to 1.43)
CBZ vs LTG	1.20 (1.02 to 1.40)	1.22 (1.07 to 1.40)	1.22 (1.05 to 1.42)
CBZ vs OXC	1.06 (0.78 to 1.44)	1.01 (0.81 to 1.26)	1.01 (0.77 to 1.31)
CBZ vs TPM	0.99 (0.78 to 1.27)	1.05 (0.86 to 1.28)	1.02 (0.81 to 1.30)
CBZ vs GBP	1.41 (1.10 to 1.81)	1.41 (1.15 to 1.73)	1.42 (1.10 to 1.83)
CBZ vs LEV	1.19 (0.94 to 1.51)	1.21 (1.01 to 1.46)	1.20 (0.96 to 1.51)
CBZ vs ZNS	1.30 (0.86 to 1.95)	1.30 (0.91 to 1.84)	1.30 (0.85 to 1.98)

See Table 25 for abbreviations, definitions and footnotes.

8.3.2 Individuals with generalised seizures

Results for all AEDs compared to reference treatment VPS from NMAs of IPD only and IPD combined with AD (as outlined in Chapter 8.2.2) are presented in Table 27 and Table 28 respectively for 'time-to-withdrawal of allocated treatment' and 'time-to-first seizure.'

When comparing the results of the NMA which combines IPD and AD to the NMAs with IPD only, numerical results are mostly quite similar apart from one change to statistical significance when incorporating AD to the important comparison of VPS vs CBZ for 'time-to-first seizure' even though no additional AD for this comparison directly was added. Aside from this one change in conclusion, the lack of notable difference is not surprising given the

small amount of additional AD available by seizure type, i.e. adding AD contributes only up to an extra 0.8% of data to the outcomes.

Table 27: Network meta-analysis results also incorporating aggregate data. Time-to-withdrawal of allocated treatment for individuals with generalised seizures.

Comparison[1]	IPD only[1,2]	IPD reduced to AD[1,3]	IPD reduced to AD, plus AD[1,3]
VPS vs CBZ	***1.20 (0.89 to 1.62)***	1.50 (1.08 to 2.07)	1.46 (1.07 to 1.98)
VPS vs PHB	***1.95 (0.92 to 4.16)***	***2.12 (0.98 to 4.58)***	***2.09 (0.98 to 4.42)***
VPS vs PHT	1.26 (0.78 to 2.04)	1.10 (0.57 to 2.15)	1.13 (0.59 to 2.15)
VPS vs LTG	0.83 (0.51 to 1.36)	0.97 (0.55 to 1.72)	0.94 (0.54 to 1.65)
VPS vs OXC	1.00 (0.18 to 5.46)	0.77 (0.30 to 1.98)	0.78 (0.31 to 1.97)
VPS vs TPM	***1.08 (0.47 to 2.44)***	***1.27 (0.68 to 2.39)***	***1.26 (0.69 to 2.33)***
VPS vs GBP	1.11 (0.12 to 9.89)	1.20 (0.14 to 10.44)	1.19 (0.14 to 10.25)
VPS vs LEV	0.93 (0.48 to 1.82)	1.16 (0.62 to 2.18)	1.15 (0.63 to 2.10)

See Table 25 for abbreviations, definitions and footnotes.

Table 28: Network meta-analysis results also incorporating aggregate data. Time-to-first seizure for individuals with generalised seizures.

Comparison[1]	IPD only[1,2]	IPD reduced to AD[1,3]	IPD reduced to AD, plus AD[1,3]
VPS vs CBZ	***1.21 (1.05 to 1.40)***	1.20 (0.97 to 1.48)	***1.22 (1.01 to 1.49)***
VPS vs PHB	1.36 (0.95 to 1.95)	1.34 (0.86 to 2.08)	1.37 (0.89 to 2.10)
VPS vs PHT	1.06 (0.81 to 1.39)	0.93 (0.61 to 1.42)	1.10 (0.73 to 1.65)
VPS vs LTG	***1.52 (1.16 to 1.99)***	1.34 (0.86 to 2.07)	1.48 (0.96 to 2.29)
VPS vs OXC	1.67 (0.51 to 5.44)	1.59 (0.85 to 2.99)	1.64 (0.88 to 3.02)
VPS vs TPM	***1.15 (0.58 to 2.30)***	***1.15 (0.73 to 1.80)***	***1.19 (0.78 to 1.83)***
VPS vs GBP	0.58 (0.07 to 4.91)	0.55 (0.06 to 4.70)	0.58 (0.07 to 4.93)
VPS vs LEV	1.45 (0.93 to 2.28)	1.34 (0.86 to 2.09)	1.34 (0.88 to 2.05)

See Table 25 for abbreviations, definitions and footnotes.

8.3.3 All individuals (regardless of epilepsy type)

Results for all AEDs compared to commonly used treatment CBZ from NMAs of IPD only and IPD combined with AD (as outlined in Chapter 8.2.2) are presented in Table 29 and Table 30 respectively for 'time-to-withdrawal of allocated treatment' and 'time-to-first seizure.'

Comparing the results of the NMA which combines IPD and AD to the NMAs with IPD only, NMA results are very numerically similar and there are no changes in conclusions (i.e. the statistical significance of the results) across any of the comparisons of CBZ to the other AEDs. Although more AD for all individuals is available to incorporate into analysis than for analyses separated by epilepsy type, the proportion of AD incorporated compared to the amount of IPD available is still small, i.e. adding AD contributes only up to an extra 10% of data to the outcomes. Furthermore, as noted in Chapter 8.2.2, results produced by these methods

without accounting for epilepsy type are of little clinical relevance given the known differences between AEDs in different epilepsy types and current clinical practice.

Table 29: Network meta-analysis results also incorporating aggregate data. Time-to-withdrawal of allocated treatment for all individuals (regardless of epilepsy type)

Comparison[1]	IPD reduced to AD[1,2]	IPD reduced to AD, plus AD[1,2]
CBZ vs PHB	1.50 (1.14 to 1.96)	1.50 (1.14 to 1.97)
CBZ vs PHT	1.02 (0.83 to 1.26)	1.03 (0.84 to 1.27)
CBZ vs VPS	0.90 (0.76 to 1.05)	0.88 (0.76 to 1.02)
CBZ vs LTG	0.74 (0.64 to 0.86)	0.73 (0.65 to 0.83)
CBZ vs OXC	0.91 (0.70 to 1.18)	0.91 (0.72 to 1.16)
CBZ vs TPM	1.09 (0.92 to 1.30)	1.07 (0.90 to 1.27)
CBZ vs GBP	1.13 (0.91 to 1.39)	0.99 (0.82 to 1.18)
CBZ vs LEV	0.83 (0.69 to 1.00)	0.82 (0.68 to 1.00)
CBZ vs ZNS	1.08 (0.76 to 1.55)	1.08 (0.74 to 1.57)

1. Results presented are HR and 95% CIs, See Chapter 5.3.2 for abbreviations of drugs. Results highlighted in bold italics show a difference in statistical significance when published AD is combined with IPD
2. Results taken from the model outlined in Chapter 8.2.2

Table 30: Network meta-analysis results also incorporating aggregate data. Time-to-first seizure for all individuals (regardless of epilepsy type)

Comparison[1]	IPD reduced to AD[1,2]	IPD reduced to AD, plus AD[1,2]
CBZ vs PHB	0.88 (0.73 to 1.05)	0.87 (0.71 to 1.05)
CBZ vs PHT	1.00 (0.87 to 1.14)	0.99 (0.85 to 1.15)
CBZ vs VPS	1.06 (0.96 to 1.16)	1.01 (0.90 to 1.13)
CBZ vs LTG	1.22 (1.11 to 1.33)	1.24 (1.11 to 1.38)
CBZ vs OXC	1.05 (0.89 to 1.23)	1.04 (0.86 to 1.27)
CBZ vs TPM	1.09 (0.97 to 1.22)	1.10 (0.95 to 1.27)
CBZ vs GBP	1.39 (1.21 to 1.59)	1.36 (1.16 to 1.59)
CBZ vs LEV	1.21 (1.07 to 1.38)	1.19 (1.01 to 1.40)
CBZ vs ZNS	1.30 (0.97 to 1.73)	1.30 (0.92 to 1.82)

See Table 29 for abbreviations, definitions and footnotes.

8.4 Discussion

8.4.1 Summary of results and clinical implications

This Chapter presents approaches for combining IPD and AD in NMA, with or without accounting for epilepsy type. The principle aim of this Chapter was to investigate whether the incorporation of AD changes the results and conclusions of an NMA approach based on IPD only, which could indicate that the 69% of IPD obtained for the NMA may not be representative of the entire evidence base.

Nine trials reported AD for up to 1850 participants for the outcome 'time-to-withdrawal of allocated treatment' and six trials reported AD for up to 1369 participants for the outcome 'time-to-first seizure.' For three trials, a summary HR and 95% CI could be extracted directly for both outcomes and for two trials, summary statistics could be estimated from published survival curves for one or both outcomes using the methods described in Chapter 2.3.2.2. For the remaining trials, IPD could be approximately reconstructed from tables or survival curves and used to estimate an aggregate HR and 95% CI.

Out of 17,961 participants eligible for the IPD-NMA, 66% of IPD was available for the analysis of 'time-to-withdrawal of allocated treatment' and 68% was available for the analysis of 'time-to-first seizure.' The additional extracted or estimated AD contributed an extra 3.5% and 1% respectively for individuals with partial seizures, 1% and 0.5% respectively for individuals with generalised seizures and 13% and 10% respectively for all individuals, regardless of seizure type. Therefore, for both outcomes, even with additionally extracted AD, around 20% of participants are still missing from the NMA.

For all analyses, there seems to be very little impact of incorporating AD with IPD in NMA. The lack of notable difference compared to the IPD-only approach to the NMA is not surprising given the small amount of additional AD available by epilepsy type.

Of note in the context of this analysis is that for one of the trials which contributed AD to the combined IPD and AD-NMA for both outcomes (Biton 2001), IPD was provided for this trial but could not be treated as IPD within the analysis. IPD for Biton 2001 was requested via data sharing portal CSDR and provided via a remote data access system which allowed analysis in SAS-based statistical software and export of analysis results but prohibited exporting of the dataset. Therefore, it was not possible to combine this IPD with the other datasets to perform the IPD only analyses (described in Chapter 6 and Chapter 7 of this thesis) and the only option was to treat results exported from the data access system as AD in analysis.

As described above, there seems to be little impact on results following the addition of AD to the IPD analyses, therefore the restricted access format of this single trial does not seem to have impacted on the results of the NMA. However, it is a concern for updates of this NMA in particular and for future IPD syntheses in general, that the provision of data in different formats and the increased use of remote access systems may restrict the analyses that it is possible to perform across all eligible datasets and subsequently impact on results of syntheses and the scope of clinical questions that are able to be addressed.

8.4.2 Limitations of the methodological approach and future work

It was assumed for the purpose of these analyses that the AD extracted was reliable and consistent in outcome definition with the IPD. It was anticipated from the results of Chapter 3 that a limited amount of AD would be available for analysis and results of this Chapter show that at the most, data for 1850 out of 5570 participants (33%) from 10 out of 41 trials (24%) not providing IPD could be extracted or estimated. Further, for all 10 studies for which AD could be extracted or estimated, there were concerns regarding definitions of the outcomes (particularly for 'time-to-withdrawal of allocated treatment,' see Chapter 3 for further discussion), definitions of censoring, time origins of the outcomes or statistical analyses performed to estimate summary statistics or graphical figures. For the two studies, with approximate IPD provided in tables, assumptions had to be made about event times (examined in sensitivity analysis, see Chapter 7.2) and for data extracted graphically, graphs were generally of poor quality and without numbers at risk provided which will affect the precision of these estimated results.

Additionally, digitisation of survival curves was not used due to the low quality of some graphics from older publications and it was preferred for the aim of this analysis to use a consistent method across all studies for the extraction of graphical data. It should be noted that for practical rather than methodological objectives, where graphical quality allows digitisation of curves, such a method may result in more accurate estimates.

In this context, it can certainly be argued that the extracted or estimated AD was of lower quality than the IPD which has been consistency checked and prepared for analysis according to a pre-specified procedure to ensure consistency of outcome definition (see Chapter 5.3.3). Previous work has argued that incorporation of AD into IPD-MAs may only be justified where the amount of missing IPD is large and/or reasons for missing IPD are thought to be informative [38-41, 309]; in fact, methods for incorporating IPD and AD within NMA have generally been developed in the context of adding a small amount of IPD (from say one or two studies) to improve the precision of an AD-NMA [19, 44-46] rather than vice-versa, as was the objective of this Chapter. Sutton *et al* [39] have argued 'that even if IPD is available from only a selection of studies, assuming no selection bias in which studies have provided IPD, IPD analysis to explore treatment – covariate interactions may produce more reliable estimates than a more complete AD meta-regression.'

Although restricting analysis to IPD only when a subset of AD is available goes against the general principle of systematic reviewing of including 'all available evidence' [40], it is

questionable whether there was any benefit at all in this context of incorporating the small amount of AD available with the IPD in analysis, given the relative quality of the two sources, the small amount of additional AD available for analysis and no evidence of informative missingness when comparing studies with and without IPD (see Chapter 5.4.3).

Riley *et al* [38, 40] warn that a joint analysis of IPD and AD may 'distort the truth' where AD is far less reliable than IPD, and that careful consideration of the approach to synthesis is needed within each individual setting. Results described in this Chapter show very few notable differences for an NMA incorporating IPD and AD compared to IPD only analyses in the same framework and no evidence of an increase in heterogeneity or inconsistency when incorporating AD with IPD in the NMA. So within this context, the small amount of 'lower quality' AD does not seem to have 'distorted the truth' of the conclusions from the IPD-only analyses but it is unknown whether bias would have been introduced into the NMA if a larger proportion of AD had been available for a joint analysis.

Based on previous work conducted by the Cochrane Epilepsy group [59-67], including an earlier IPD-NMA using a subset of the present data [285] and the work outlined in Chapter 3 of this thesis, it was estimated that around 80% of IPD would be retrievable for analysis and very little useable AD would be available. Results of Chapter 5 to Chapter 7 of this thesis agree with these *a priori* assumptions that substantially more IPD than AD would contribute to a combined analysis, although the final retrieval rate of 69% of IPD was slightly lower than anticipated. Previous methods considering combined analyses of IPD and AD in NMA have generally been developed in the context of adding a small amount of IPD to improve the precision of an AD-NMA [19, 44-46] and such methods are more methodologically complex, requiring analysis within a Bayesian framework.

Therefore, following consideration, it was felt that the gain to the analysis and precision of results would not outweigh the methodological complexity within this context, particularly for communicating clinical results to the readership of Cochrane reviews for whom interpretations of Bayesian statistics such as credible intervals are quite unfamiliar. Hence a two-stage approach to the analyses outlined this Chapter (i.e. reducing IPD to AD and combining with additional AD using methods for AD-NMA) was specified in order to investigate, as far as possible, the extent of missing IPD on the clinical results of the NMA.

On the other hand, from a methodological point of view, it would be of interest in future work to investigate the benefits of a Bayesian framework as defined in previous work for this

example of combining IPD and AD, to determine if any further information can be gained throughout the 'borrowing of strength' of a Bayesian approach.

Furthermore, shortly after the submission of the protocol of the present IPD-NMA for publication on the Cochrane Database of Systematic reviews [68], we became aware of novel methodology proposed by Saramago *et al* [44] for combining individual event-time data with aggregated count data. These methods were not applied within this Chapter due to the differences in model distribution (parametric rather than semi-parametric) and framework (Bayesian rather than Frequentist) which were outside of the scope of this Chapter, which aimed to investigate whether the IPD-NMA is representative of the evidence base and to examine the robustness of results.

These methods perhaps hold the most potential for future work in the context of this example. As outlined above, a more accessible two-stage method was preferred over methodologically complex methods to combine up to 68% of IPD with up to 13% of AD, resulting in a combined IPD and AD-NMA which still had around 20% of eligible patients missing. The methods of Saramago *et al* [44], assuming a parametric distribution, allow for HRs to be generated from the original count AD by estimating the cumulative hazard in each trial arm reporting count data for a given follow-up time. The example illustrated by the authors relates to high compression treatments for venous leg ulcers and the time-to-event outcome of interest is time-to-healing (with corresponding aggregate count data outcome of number healed) and the authors have IPD for 841 participants from two trials (43% of total data) and AD for 1105 participants from 14 trials (57% of total data).

Examination of the outcomes and summary statistics reported within the epilepsy monotherapy trials not providing IPD or any time-to-event AD showed that an additional 16 trials (recruiting 2806 participants) reported the 'number of participants seizure-free' and an additional 13 trials (recruiting 2398 participants) reported the 'number of participants withdrawing from treatment.' Therefore, in principle, using the methods of Saramago *et al* [44] may allow for up to 93% of eligible data to be included in NMA for 'time-to-withdrawal of allocated treatment' and up to 91% of data for 'time-to-first seizure.'

However, before undertaking such an analysis, further consideration would have to be given to the follow-up times and the measurement times of the count data within each trial. For example, some trials report seizure freedom at a series of time points throughout the trial, other trials report only seizure freedom during the maintenance period etc. Additionally, in order to synthesise 'time-to-withdrawal of allocated treatment' and 'number of

withdrawals,' consideration would have to be given to reasons for withdrawals which may not be as clearly articulated within published reported compared to participant-specific withdrawal reasons provided within IPD. Furthermore, within this context where treatment-epilepsy type interactions are of interest, while the methods of Saramago *et al* [44] do extend to the incorporation of treatment-covariate interactions, for the majority of studies, AD (whether TTE or count data) tends to be published for all individuals, rather than separated by epilepsy type or adjusted for a treatment-epilepsy type interaction.

Therefore, as this Chapter has shown for the 'simple' two-stage methods incorporating up to an additional 2% of epilepsy-type specific AD has very little impact on numerical results compared to an IPD only approach, the added benefits of the more complex methodology described by Saramago *et al* [44] may also be very minimal for this example.

8.4.3 Concluding remarks

In conclusion, this Chapter demonstrates the numerical results and conclusions of the IPD-NMA described in Chapter 6 and Chapter 7 seem robust to the incorporation of a small amount of additional published AD with IPD in the NMA model.

Despite best efforts to include as much relevant evidence as possible, including published aggregate data, NMAs presented within this Chapter are still missing between 20 and 30% of eligible data which almost inevitably will result in bias to some extent and must be taken into account when interpreting clinical results.

The provision of accessible, standardised and high-quality data (whether provided at the aggregate or IPD level) is essential to allow updates of this IPD-NMA as further information becomes available, particularly for recently licenced and future treatment options.

It appears that two IPD-MA projects submitted to the YODA project have already been prevented due to the restrictive remote access to data [93]. It is of concern that the increased use of remote access systems for sharing IPD with researchers will further restrict the analyses that it is possible to perform across all eligible datasets and subsequently have even further impact on results of syntheses and the scope of clinical questions that are able to be addressed. Additional flexibility within data sharing platforms, such as the ability to temporarily download IPD from the remote systems to perform syntheses, under the additional protection of legal documents such as data sharing agreements to prevent misuse of data may offer a solution.

Chapter 9: Discussion and Conclusions

9.1 Summary of main findings of the thesis

Evidence syntheses are highly regarded techniques for the quantitative summary of evidence from a number of sources [8, 34]. In comparison to traditional aggregate-level data approaches, an IPD approach to meta-analysis has been widely regarded as the 'gold-standard' for many years [22, 23], with a sharp increase in the number of IPD-MAs published in the last decade [33-35]. Recent work has also shown the benefit of an IPD approach for NMA [19, 44-48, 328].

Within many clinical settings, outcomes of interest are measured as a time to an event. Synthesis of TTE data is particularly common in the field of oncology but also important in other settings, such as measuring the retention time on treatment or the remission time from seizures for people with epilepsy. A range of methods for the meta-analytic synthesis of TTE data have been proposed over several decades and applied to a wide range of clinical and methodological scenarios; Chapter 2 of this thesis presents a literature review of this methodology according to the level of data required for the approach (IPD or AD).

It is well documented within the field of oncology that the necessary published information required to perform AD-MA of TTE data is often not reported or is reported inconsistently [31, 144, 160, 235-237]. Therefore, a range of accessible and user-friendly methods have been developed with the aim of making use of more commonly reported summary statistics and published survival curves to indirectly estimate HRs and associated variances [6, 128]. However, whether these methods can be used in practice has been questioned since many alternative summary statistics are also not reported or published graphical figures are of inadequate quality [6, 141, 144]. Chapter 3 of this thesis summarises previous investigations of the reporting of aggregate TTE data in oncology and presents a novel systematic review of the reporting of TTE outcomes and associated statistics in epilepsy monotherapy studies.

This is believed to be the first systematic review of this topic outside of the field of oncology, reflecting reporting standards across 35 years and 24 speciality and general medicine journals. In line with previous work, results of this systematic review reveal concerning reporting inadequacies relating to the definition, analysis and reporting of TTE outcomes within these epilepsy monotherapy trials. In fact, the findings for some areas of reporting, particularly relating to definitions of outcomes, seem to be worse than previous reviews in oncology. These results also support the rationale of the Cochrane Epilepsy Group of taking

an IPD approach to synthesis of AED monotherapy trials, with nine IPD-MAs [59-67] and two IPD-NMAs [68, 69, 285] published to date, as a fully AD approach to synthesis based only on the published information presented in these trials would not be feasible or recommended to inform clinical practice.

Increasing popularity of IPD approaches to meta-analysis over the last decade has resulted in rapid development of methodology allowing questions of growing clinical, statistical and computational complexity to be addressed via IPD-MA models [36]; Chapter 2 of this thesis summarises IPD-MA methodology specific to TTE outcomes. There has also been a rapid increase in the uptake of methods; Chapter 4 of this thesis shows that the number of systematic and non-systematic IPD-MAs published per year has increased to an average of 105 between 2009 and 2015 compared to 49 per year between 2005 and 2009 [34].

However, despite the benefits of an IPD approach to synthesis and the increased use of such an approach, in practice retrieving all IPD to perform a re-analysis can require a considerable amount of time, cost and personnel and can be computationally intensive in the case of large individual participant datasets [22, 49]. Furthermore, retrieval of all relevant IPD is not always possible for a variety of reasons (IPD may have been destroyed or lost, original investigators may be unwilling to collaborate etc.) and only a proportion of IPD may be available for re-analysis. This leaves the IPD-MA at potential risk of 'availability bias' where the subset of IPD available is not representative of the evidence base. In this case, a combined synthesis of IPD and AD may be a feasible option to increase precision and reduce 'availability bias;' methodology for the combined synthesis of IPD and AD is summarised in Chapter 8.

The culture of clinical trial data sharing has changed in recent years, with a shift in attitudes towards the support of data sharing and many calls for improved data transparency and data sharing initiatives introduced across the research community as a whole [77-87]. The way that pharmaceutical clinical trial data is shared for secondary research, such as syntheses, has also changed with the launch of data-sharing platforms such as CSDR, YODA and SOAR since 2013 [91, 93, 95]. Such platforms allow researchers to request access to IPD via a structured process of selecting studies of interest, submitting a scientific research proposal, signing of a data sharing agreement by the researcher and sponsor and finally access to de-identified IPD and related documentation which can be analysed remotely in a SAS analytic environment and analysis results exported from the environment. While such initiatives and platforms should make access to IPD easier and faster, the impacts of these changes in attitudes and data sharing methods may not become clear for some time.

Chapter 4 presents a novel systematic review of data retrieval within IPD-MAs. Out of 760 IPD-MAs using systematic methods to identify eligible studies published between 1987 and 2015, only 188 (25%) retrieved 100% of the eligible IPD for analysis and only 324 (43%) retrieved at least 80% of relevant IPD. Chapter 4 also shows that up to 2015, there is no evidence of an improvement in IPD retrieval rates over time but that IPD-MAs that included only randomised trials, had an authorship policy, included fewer eligible participants and were conducted outside of the Cochrane Database of Systematic Reviews were associated with a high or complete IPD retrieval rate.

Chapter 5 reflects upon the IPD requesting experiences of the Cochrane Epilepsy group over the last 20 years, from a time when IPD approaches to synthesis were relatively novel up to the end of 2016, throughout the new era of data sharing initiatives. Chapter 5 shows that the earliest Cochrane Epilepsy IPD-MA published in 2000 included IPD from 63% of total trials and 83% of total participants [67], a good retrieval rate in the wider context of all IPD-MAs as presented in Chapter 4; however the latest Cochrane Epilepsy IPD-NMA included IPD from 47% of total trials and 69% of total participants [69]. This reflects a decline in the IPD retrieval rate from requests made between 1995 and 2005 to requests made between 2012 and 2015.

Chapter 4 also highlights that reporting inadequacies in IPD-MAs of all clinical contexts are not uncommon, with 257 out of the 760 IPD-MAs (34%) not reporting sufficient information to calculate the IPD retrieval rate and in 58% of IPD-MAs that failed to retrieve 100% of eligible IPD, there were no specific reasons provided for the unavailability of data. Furthermore, in around a quarter of IPD-MAs that failed to retrieve 100% of eligible IPD, there was a complete lack of discussion or acknowledgement of any biases that may have been introduced by the missing IPD.

Chapter 6 presents the statistical methodology and Chapter 7 presents clinical results for an IPD-NMA of ten AEDs used in monotherapy for 12,391 participants from 36 clinical trials that IPD was successfully retrieved from, data for 69% of eligible participants from 47% of eligible trials as outlined in Chapter 5. Outcomes considered within this IPD-NMA included 'time-to-withdrawal of allocated treatment' and 'time-to-first seizure after randomisation' and the NMA also incorporated a treatment-covariate interaction between the antiepileptic drug and epilepsy type (partial or generalised seizures). Clinical results of this IPD-NMA support current NICE guidelines [74] and suggest some alternative treatment options for those individuals for which the first-line recommended treatments are not suitable.

Chapter 8 presents methodological approaches for including AD with IPD in NMA for this example. As predicted by the results of the systematic review presented in Chapter 3, few of the epilepsy monotherapy studies without IPD available reported any suitable AD which could be extracted or indirectly estimated (numerically or graphically, via the methods outlined in Chapter 2). The additional extracted or estimated AD contributed only up to an extra 3.5% or 1% of data to the NMA for individuals with partial seizures and individuals with generalised seizures respectively and the incorporation of this additional AD with the IPD in NMA had a negligible impact on results.

On the other hand, the methodological approach to the relationship between treatment and epilepsy type did have an impact on the results; while different approaches produced very similar numerical results and mostly identical conclusions for individuals with partial seizures (the majority epilepsy type, around 70%), numerical results for individuals with generalised seizures (the minority epilepsy type, around 25%) change quite substantially, as well as some changes in the statistical conclusions, with difference analysis approaches for epilepsy type.

9.2. Implications for practice and research

The implications of this thesis fall into three topics; the methodological implications of the findings around IPD retrieval and around conduct and reporting of IPD syntheses, and the clinical implications of the illustrative example of the IPD-NMA of antiepileptic drugs.

9.2.1 Clinical implications

The findings of this thesis are underpinned by the application of methodology to a large Cochrane IPD-NMA of ten AEDs and investigation of the treatment-by-epilepsy type (partial or generalised seizures) interaction.

The results of the IPD-NMA demonstrate that, in line with current NICE guidelines [74], that CBZ and LTG are suitable first-line treatments for individuals with partial seizures but also adds new information that LEV may be a suitable alternative. Results for individuals with partial seizures are robust to additional and sensitivity analyses including investigation of different methodological approaches for modelling the relationship between treatment effect and epilepsy type, accounting for any observed inconsistencies in the IPD provided and following the incorporation of additional AD with IPD into the NMA model.

The majority of participants recruited into the trials included in the IPD-NMA were classified as experiencing partial seizures, which is reflective of clinical practice that around 60% of individuals with epilepsy experience partial seizures [320]. However, the proportion of individuals with partial seizures recruited to the trials in the IPD-NMA is greater than would be expected in clinical practice with 67.5% of included participants classified as experiencing partial seizures, around 26.5% of participants experiencing generalised seizures and the remaining 6% experiencing seizures of a type which is difficult to classify. Additionally, there was an indication that up to 36% of individuals classified as experiencing generalised seizures may have had their seizure type misclassified, so the true proportion of individuals experiencing generalised seizures within the IPD-NMA may be as small as 17%.

Due to this imbalance in the two epilepsy types, results for individuals with generalised seizures were less precise and less robust to different approaches to statistical analysis. Results of the main analysis for individuals with generalised seizures are also in line with current NICE guidelines, supporting the use of VPS as a first-line treatment but also adds new information that LTG and LEV may be suitable alternative first-line treatments.

The findings of this thesis and the IPD-NMA provide recommendations for the design and conduct of future AED monotherapy trials. It is essential that future trials are adequately powered to detect a difference between particular AEDs while recruiting a sample of individuals representative of the wider population in terms of age and epilepsy type. The latter is particularly important in order that future syntheses can provide more precise and robust evidence for individuals with generalised seizures as additional potential treatment options become available. Given that current clinical practice reflects a different selection of preferred treatment options of individuals with different epilepsy types [74], it is also recommended that future trials should incorporate interactions between treatment effect and epilepsy types within statistical analysis and clearly report treatment and interaction effects according to different epilepsy types within the trials.

Furthermore within these trials, the choice of outcomes at the design stage and the presentation of the results of outcomes require careful consideration. The ILAE recommend that trials of a monotherapy design should adopt a primary effectiveness outcome of 'time-to-withdrawal of allocated treatment' and should be of a duration of at least 48 weeks to allow for assessment of longer-term outcomes, such as remission [249]. If trials followed these recommendations, an AD approach to meta-analysis may be feasible, reducing the resources and time required from an IPD approach.

In a wider clinical context, the findings of this thesis highlight the importance of an IPD approach when considering a clinically complex treatment-covariate interaction, in addition to the essential requirement of transparent reporting at all levels, whether at the trial publication level, IPD level or data request level (e.g. with respect to reasons why IPD are not available), to allow the appropriate conclusions to be drawn from the clinical question.

9.2.2 Methodological implications: IPD retrieval rate

Reviews of the literature conducted for this thesis indicate that in recent years, IPD approaches to synthesis are rapidly gaining in popularity, with the majority of new methodological work relating to meta-analysis requiring the analysis of IPD and it is now estimated that over 100 new IPD-MAs are being published each year. In parallel, initiatives to promote open and transparent sharing of clinical trial data continue to gain in momentum over recent years.

However, the findings of this thesis show that these substantial changes in culture and methodological practice do not seem to be mirrored by improved IPD retrieval rates. This may be due, in part, to the increasing uptake of IPD-MAs across a wide range of clinical areas and settings and increasing use of sophisticated systematic searching methods, such as those employed by the Cochrane Collaboration, which uncover grey literature where IPD may be difficult to obtain.

There has been a decline in the IPD retrieval rate from data requests made by the Cochrane Epilepsy group between 2012 and 2015 compared to requests made between 1995 and 2005. A concern is that 'prohibitive costs' have prevented the sharing of pharmaceutical data for recent requests made by the Cochrane Epilepsy group and there seems to be an emerging association between the resources of the data provider and the provision of IPD.

Changes in methods of sharing pharmaceutical and the additional step of rigorous data checking and de-identification have anecdotally resulted in the provision of cleaner datasets compared to datasets provided in previous requests to the Cochrane Epilepsy group. While this is beneficial to the researcher, who is required to spend less time and resource checking data and resolving problems or inconsistencies, these changes are associated with additional costs and resources to the data provider.

Recent work has demonstrated that preparation of academic clinical trial data for external sharing may take up to 50 hours and the associated cost may be as high as £3000 per trial [300]. Such time and financial costs may be even higher for older trials, or trials which are

particularly large or of a complex data structure. It can be argued that it is unreasonable for data providers to incur all of these costs, without discernible acknowledgement or reward, particularly for academic data providers for whom the intellectual property of their data may be of great value to the academic institution and to individuals within the institution. Furthermore, results of this thesis have also demonstrated that implementation of an authorship policy, an incentive to participate in IPD-MAs, is associated with a complete or high retrieval rate of IPD compared to IPD-MAs without an authorship policy. Therefore, collaboration between researchers and data providers, whether financial, shared authorship or otherwise, may assist in sharing costs and resources and also the benefits of any secondary research. In turn, collaboration of personnel and resources may potentially maximise retrieval rates of IPD.

Of further concern in the present era of data transparency is the continued reporting of non-specific reasons for unavailability of IPD. Three out of the 35 requests made by the Cochrane Epilepsy group between 2012 and 2015 received a negative response but no specific reason was given as to why data could not be provided and metrics available on the CSDR website also list non-specific reasons for why data could not be provided for 15% of requests. While it is inevitable that some IPD-MAs will not be able to include all relevant IPD for perfectly valid reasons, where a specific reason for the unavailability of IPD is not provided, it is difficult to make a judgement regarding the presence of availability bias which has implications for the interpretation of the results and conclusions of the synthesis.

Findings of this thesis also suggest that the current format of the data sharing platforms such as CSDR and YODA may not be suitable for syntheses due to the restrictive access to the data provided and associated legal obligations which prevent collating all available data in a single location for analysis. The impact that the current restrictions of data sharing platforms may have on future clinical and statistical analyses, potentially rendering some analysis approaches impossible within the framework, are of great concern. Additional flexibility within data sharing platforms, such as the ability for 'approved' researchers to temporarily download IPD to perform syntheses, may offer a solution.

9.2.3 Methodological implications: conduct and reporting of IPD syntheses

The findings of this thesis reveal concerning inadequacies of reporting in many areas which have implications to the conduct of evidence synthesis and many of which have contributed to the conduct of a large IPD-NMA in epilepsy.

Concerns regarding inadequate reporting of aggregate TTE data has been well documented since 1995 [31] and there does not seem to have been any improvement over time, with several other reviews of this topic within oncology trials, the most recent published in 2013 [242], showing equally poor reporting. Findings of this thesis emphasise the calls from previous reviews [31, 242] for the urgent development of minimum reporting standards for TTE analyses. In the continuing absence of the development of such standards, use of the suggested guidelines from previous work [31, 242] by journal editors and peer reviewers when considering study publications using TTE analyses would greatly improve reporting rates and in turn facilitate the conduct of AD-MAs and syntheses with TTE endpoints.

Findings of this thesis also demonstrate inadequacy of reporting in the published results of IPD-MAs, particularly regarding study and participant numbers contributing to different stages of the IPD-MA, reasons for unavailability of IPD and consideration of potential availability bias introduced by missing IPD. It is unknown whether the lack of discussion or additional analyses to investigate availability biases are due to inadequate or selective reporting, or whether meta-analysts simply were not aware that biases could have been introduced into their analyses by a missing subset of eligible IPD.

A recent scoping review of published IPD-NMAs, a relatively new research field compared to IPD-MAs, has also revealed several recurrent areas of inadequate reporting, such as an evaluation of consistency assumptions, existence of a study protocol, and methods used to request, collect, and manage IPD and management of missing data [48].

These findings highlight that improvements are needed in the conduct and reporting of IPD syntheses. It is highly recommended that all syntheses, whether of an AD or IPD level and whether systematic or prospective, follow a registered protocol [284] and that any deviations from the protocol are clearly described and justified. It is particularly important for IPD approaches of a systematic nature, that researchers demonstrate awareness that despite their best efforts, it may not be possible to obtain all eligible IPD for analysis and that it is clearly outlined how biases related to missing IPD will be evaluated.

It is also absolutely essential that the number of eligible studies and participants, how much data was requested and obtained with clear reasons for non-availability of IPD, preferably via a flow diagram, are transparently reported. Such information is not only important for the interpretation of reported results and conclusions, but is also of value for future research; providing ‘best practice’ guidance for researchers embarking upon IPD synthesis.

Proper uptake of new PRISMA-IPD guidelines for the conduct and reporting of IPD-MA [283], in addition to guidance on the use of IPD-MA to synthesise the results of RCTs [282] should lead to improvements in the highlighted areas of conduct and reporting that are currently inadequate. There has also been a recent call for the development of PRISMA-IPD-NMA guidelines, based on the current PRISMA-IPD and PRISMA-NMA guidelines for the conduct and reporting of IPD-MAs and NMAs respectively [48]. Such guidelines would be a welcome addition to the evidence synthesis literature and may prevent inadequacies of reporting of IPD-NMAs becoming commonplace, as this research field continues to develop.

While existing guidelines such as PRISMA, PRISMA-IPD and PRISMA-NMA provide a minimum set of reporting items within each synthesis design (i.e. AD, IPD or NMA); there are currently no formal recommendations, from PRISMA or from Cochrane methods groups, to guide authors when choosing the most appropriate synthesis approach to the clinical question and what information should be reported regarding the rationale for the analysis approach. Previous work has recommended that researchers should carefully consider the added benefits of IPD whether a less resource intensive AD-MA could provide an adequate (and mathematically equivalent) answer to the question [52, 53]. Ideally, the rationale for an IPD approach to synthesis should be clearly described and justified in a registered protocol.

Previous work has also suggested that the additional methodological complexity of incorporating both IPD and AD sources within synthesis may not necessarily provide any additional benefit to analysis and in fact may introduce biases into the combined analysis where AD is of a lower quality than the IPD [38, 40]. Although relatively simple methodology was used to incorporate IPD and AD for the illustrative example outlined in this thesis, findings are in line with previous work and suggest that there was very little impact to incorporating IPD and AD in this context, given that the majority of evidence came from IPD sources compared to AD sources. Therefore, it should be recommended that researchers consider the added benefits of a combined IPD and AD analysis over an IPD-only approach, taking into account the relative quality and available proportion of each data source. Again, ideally, any planned methodology for combining IPD and AD in synthesis should be clearly outlined in a registered protocol.

9.3 Limitations and future work

The specific limitations to each element of the work presented in this thesis have been discussed in detail in the relevant chapters. This section will focus on the overall limitations of the work presented in this thesis and make suggestions for future work which may be able to address some of the limitations.

The first major limitation is that many of the findings from this thesis are based on published information; such as the results reported in epilepsy monotherapy trials, IPD retrieval rates and characteristics of IPD-MAs and metrics from the CSDR and YODA websites.

Due to the size of the cohorts included in the novel systematic reviews presented in this thesis, particularly the systematic review of IPD-MAs, it was out of the scope of the work to contact original investigators individually to request additional or unpublished information.

In the context of Chapter 3, where results showed inconsistent choices of outcomes and inadequate reporting of TTE analyses in epilepsy trials, it would be of value to gain further insight into potential reasons for these findings. For example, how the trial outcomes were selected and why so few trials report the primary outcome according to the definition recommended by the ILAE [249]; perhaps clinicians do not agree with current guidelines and find recommended outcomes too complex or prefer alternative outcomes? Such insights could help to provide updated guidelines that more trials may adhere to, hence improving the consistency of outcome reporting across trials and facilitating synthesis of trials.

Furthermore, it was not possible within the systematic review of IPD-MAs in Chapter 4 to examine data requesting and collection methods in detail as this level of information was not provided within published journal articles. This is a great limitation of the work as the approach to collecting data is likely to be very influential on the proportion of data retrieved. Therefore gaining further insight into the approaches taken by different research teams with respect to the routes of communications to make data requests, wording of requests, number of attempts at a request, any deadline set for data to be provided etc. would be of great interest and would be valuable to inform the conduct of future IPD syntheses.

While word limits of journals do not always allow for detailed descriptions of methodology, including data collection methods for IPD-MAs, such information could be provided as online appendices or published within registered protocols, for example via PROSPERO [284] where word limits are not as restrictive. The detailed data collection methods for the IPD-NMA described in this thesis are outlined in Chapter 5 and published in an article in the British

Medical Journal [278]. While the methods employed to collect data for this IPD-NMA may not necessarily reflect the 'best practice' to obtain the maximum amount of IPD as possible; we hope providing this additional level of practical information regarding our methodology may further aid with interpretation of the results of the IPD-NMA and may act as a starting point for research groups planning future IPD syntheses.

The second main limitation of the work in this thesis is the relatively simplistic 'two-stage' approach to the IPD-NMA (Chapter 6) and the NMAs combining IPD and AD (Chapter 8). As discussed in Chapter 6.3.2, at the time that the IPD-NMA analysis was planned [68], methodology in this field was still relatively rare, therefore the proposed methodology was adapted from a previous IPD-NMA in epilepsy conducted by the group [285]. Although additional IPD-NMA methodology is now available and additional applied IPD-NMAs have been published [48], we are still not aware of any existing methodology which would allow for one-stage IPD-NMA of a TTE outcome with a treatment-by-covariate interaction, separated into within and across-study interactions which also allows the incorporation of AD within a single model.

The majority of methods for NMA, particularly IPD-NMA and models which combine IPD and AD in NMA have been developed within a hierarchical Bayesian modelling framework. When planning this IPD-NMA, intended for a Cochrane readership, it was decided that analysis and presentation of results within a Frequentist framework would be preferable for accessible interpretation of results. It would be of interest for further work to investigate the benefits of Bayesian framework for this example; particularly whether any gain in precision in results for individuals with generalised seizures would be possible from the additional 'borrowing of strength' across the network of a Bayesian approach. Also it would be of interest to further research whether a Bayesian framework is the only feasible option to incorporate all requirements of this example together in a one-stage NMA approach (i.e. analysis of IPD only or the simultaneous analysis of IPD and AD, analysis of a TTE outcome, consideration of within and across-study treatment-covariate interactions).

Important considerations for further work would also be whether existing IPD-NMA methods developed within a Bayesian framework could also be translated into a Frequentist framework; hence potentially making results of complex IPD-NMAs more accessible to a wider range of non-expert readers, such as the readers of Cochrane reviews.

9.4 Concluding remarks

The work of this thesis has provided a detailed insight into the conduct of an IPD-NMA in epilepsy and highlighted many inadequacies of the conduct and reporting of AD and IPD syntheses across a wide range of clinical disciplines. It is essential for clinical research of all sources, whether an original trial or synthesis, that the study is well-designed, following a registered protocol, adhering to any relevant recommendations or minimum reporting guidelines and that all results are reported transparently. Recommendations taken from previous work and from this thesis and improved conduct and reporting of clinical research, whether trials or syntheses, could have a valuable impact on evidence based medicine.

The work of this thesis was undertaken during a time of great change within the research community regarding how clinical trial data is shared for secondary research. These changes will have substantial impacts on how IPD syntheses such as meta-analyses and network meta-analyses can be conducted in the future. While the full extent of the impact of this new era of data transparency may not become apparent for some time, the work undertaken in this has identified some of the early benefits and importantly some of challenges and restrictions of new methods of sharing clinical trial data. If emerging limitations, particularly those related to restricted access to data, can be addressed while data sharing platforms are still relatively new and under development, in the future, data sharing platforms are likely to be a valuable tool rather than a hindrance to IPD syntheses.

References

1. Pignon J-P, Aupérin A, Hill C. [Meta-analyses on individual patient data and treatment evaluation in oncology]. Bull Cancer 2007;94(11):957-64.
2. Gotzsche PC. Why we need a broad perspective on meta-analysis. It may be crucially important for patients. BMJ 2000;321(7261):585-6.
3. Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. Res Synth Methods 2012;3(2):80-97.
4. Early Breast Cancer Trialists' Collaborative Group. Statistical methods: Treatment of early breast cancer. Oxford: Oxford University Press; 1990.
5. Peto R. Why do we need systematic overviews of randomized trials? (Transcript of an oral presentation, modified by the editors). Stat Med 1987;6(3):233-40.
6. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med 1998;17(24):2815-34.
7. Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. Stat Med 1991;10(11):1665-77.
8. OCEBM Levels of Evidence Working Group. The Oxford 2011 Levels of Evidence. http://www.cebm.net/index.aspx?o=5653 (accessed 09/01/2017); 2011.
9. Whitehead A. Meta-Analysis of Controlled Clinical Trials. Chichester: Wiley; 2002.
10. Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. Prog Cardiovasc Dis 1985;27(5):335-71.
11. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. Methods for Meta-Analysis in Medical Research. Chichester: John Wiley & Sons Ltd.; 2000.
12. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-88.
13. Williamson PR, Smith CT, Hutton JL, Marson AG. Aggregate data meta-analysis with time-to-event outcomes. Stat Med 2002;21(22):3337-51.
14. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005;331(7521):897-900.
15. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997;50(6):683-91.
16. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004;23(20):3105-24.
17. Lu G, Ades A. Assessing Evidence Inconsistency in Mixed Treatment Comparisons. J Am Stat Assoc 2006;101(474):447-59.
18. Efthimiou O, Debray TP, van Valkenhoef G, Trelle S, Panayidou K, Moons KG, et al. GetReal in network meta-analysis: a review of the methodology. Res Synth Methods 2016.
19. Jansen JP. Network meta-analysis of individual and aggregate level data. Res Synth Methods 2012;3(2):177-90.
20. Jansen JP, Cope S. Meta-regression models to address heterogeneity and inconsistency in network meta-analysis of survival outcomes. BMC Med Res Methodol 2012;12:152.
21. Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. Stat Med 2009;28(14):1861-81.
22. Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data Stat Med 1995;14(19):2057-79.
23. Stewart LA, Parmar MKB. Meta-analysis of the literature or of individual patient data: is there a difference? Lancet 1993;341(8842):418-22.
24. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. Eval Health Prof 2002;25(1):76-97.

25. Hua H, Burke DL, Crowther MJ, Ensor J, Tudur Smith C, Riley RD. One-stage individual participant data meta-analysis models: estimation of treatment-covariate interactions must avoid ecological bias by separating out within-trial and across-trial information. Stat Med 2017;36(5):772-89.
26. Lambert PC, Sutton AJ, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. J Clin Epidemiol 2002;55(1):86-94.
27. Sutton AJ, Duval SJ, Tweedie RL, Abrams KR, Jones DR. Empirical assessment of effect of publication bias on meta-analyses. BMJ 2000;320(7249):1574-7.
28. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. BMJ 2010;340:c365.
29. Tierney JF, Pignon JP, Gueffyier F, Clarke M, Askie L, Vale CL, et al. How individual participant data meta-analyses have influenced trial design, conduct, and analysis. J Clin Epidemiol 2015;68(11):1325-35.
30. Vale CL, Rydzewska LH, Rovers MM, Emberson JR, Gueyffier F, Stewart LA. Uptake of systematic reviews and meta-analyses based on individual participant data in clinical practice guidelines: descriptive study. BMJ 2015;350:h1088.
31. Altman DG, De Stavola BL, Love SB, Stepniewska KA. Review of survival analyses published in cancer journals. Br J Cancer 1995(2):511-8.
32. Simmonds MC, Higgins JP, Stewart LA, Tierney JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. Clin Trials 2005;2(3):209-17.
33. Simmonds M, Stewart G, Stewart L. A decade of individual participant data meta-analyses: A review of current practice. Contemp Clin Trials 2015;45(Pt A):76-83.
34. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ 2010;340:c221.
35. Huang Y, Mao C, Yuan J, Yang Z, Di M, Tam WW, et al. Distribution and epidemiological characteristics of published individual patient data meta-analyses. PLoS One 2014;9(6):e100151.
36. Debray TP, Moons KG, van Valkenhoef G, Efthimiou O, Hummel N, Groenwold RH, et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. Res Synth Methods 2015;6(4):293-309.
37. Tudur Smith C, Williamson PR, Marson AG. Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. Stat Med 2005;24(9):1307-19.
38. Riley RD, Lambert PC, Staessen JA, Wang J, Gueyffier F, Thijs L, et al. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. Stat Med 2008;27(11):1870-93.
39. Sutton AJ, Kendrick D, Coupland CA. Meta-analysis of individual- and aggregate-level data. Stat Med 2008;27(5):651-69.
40. Riley RD, Simmonds MC, Look MP. Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods. J Clin Epidemiol 2007;60(5):431-9.
41. Riley RD, Steyerberg EW. Meta-analysis of a binary outcome using individual participant data and aggregate data. Res Synth Methods 2010;1(1):2-19.
42. Debray TP, Riley RD, Rovers MM, Reitsma JB, Moons KG. Individual participant data (IPD) meta-analyses of diagnostic and prognostic modeling studies: guidance on their use. PLoS Med 2015;12(10):e1001886.
43. Lyman GH, Kuderer NM. The strengths and limitations of meta-analyses based on aggregate data. BMC Med Res Methodol 2005;5:14.
44. Saramago P, Chuang LH, Soares MO. Network meta-analysis of (individual patient) time to event data alongside (aggregate) count data. BMC Med Res Methodol 2014;14:105.
45. Saramago P, Sutton AJ, Cooper NJ, Manca A. Mixed treatment comparisons using aggregate and individual participant level data. Stat Med 2012;31(28):3516-36.

46. Donegan S, Williamson P, D'Alessandro U, Garner P, Smith CT. Combining individual patient data and aggregate data in mixed treatment comparison meta-analysis: Individual patient data may be beneficial if only for a subset of trials. Stat Med 2013;32(6):914-30.
47. Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing the consistency assumption by exploring treatment by covariate interactions in mixed treatment comparison meta-analysis: individual patient-level covariates versus aggregate trial-level covariates. Stat Med 2012;31(29):3840-57.
48. Veroniki AA, Straus SE, Soobiah C, Elliott MJ, Tricco AC. A scoping review of indirect comparison methods and applications using individual patient data. BMC Med Res Methodol 2016;16:47.
49. Crowther MJ, Riley RD, Staessen JA, Wang J, Gueyffier F, Lambert PC. Individual patient data meta-analysis of survival data using Poisson regression models. BMC Med Res Methodol 2012;12:34.
50. The Cochrane Collaboration. Review Manager (RevMan) Copenhagen: The Nordic Cochrane Centre; 2014.
51. Ahmed I, Sutton AJ, Riley RD. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. BMJ 2012;344:d7762.
52. Tudur Smith C, Marcucci M, Nolan SJ, Iorio A, Sudell M, Riley R, et al. Individual patient data meta-analyses compared with meta-analyses based on aggregate data. Cochrane Database Syst Rev 2016(9):10.1002/14651858.MR000007.pub3.
53. Haines TP, Hill AM. Inconsistent results in meta-analyses for the prevention of falls are found between study-level data and patient-level data. J Clin Epidemiol 2011(2):154-62.
54. Williamson PR, Marson AG, Tudur C, Hutton JL, Chadwick D. Individual patient data meta-analysis of randomized anti-epileptic drug monotherapy trials. J Eval Clin Pract 2000;6(2):205-14.
55. Mathew T, Nordstrom K. On the equivalence of meta-analysis using literature and using individual patient data. Biometrics 1999;55(4):1221-3.
56. Olkin I, Sampson A. Comparison of Meta-Analysis Versus Analysis of Variance of Individual Patient Data. Biometrics 1998;54(1):317-22.
57. Collett D. Modelling Survival Data in Medical Research. Second ed. London: Chapman & Hall; 2003.
58. Cox DR. Regression Models and Life-Tables. J R Stat Soc Series B Stat Methodol 1972;34(2):187-220.
59. Nolan SJ, Marson AG, Weston J, Tudur Smith C. Phenytoin versus valproate monotherapy for partial onset seizures and generalised onset tonic-clonic seizures: an individual participant data review. Cochrane Database Syst Rev 2016;4:CD001769.
60. Nolan SJ, Sudell M, Tudur Smith C, Marson AG. Topiramate versus carbamazepine monotherapy for epilepsy: an individual participant data review. Cochrane Database Syst Rev 2016(1):CD012065.
61. Nolan SJ, Tudur Smith C, Weston J, Marson AG. Lamotrigine versus carbamazepine monotherapy for epilepsy: an individual participant data review. Cochrane Database Syst Rev 2016(11):CD001031.
62. Nolan SJ, Marson AG, Weston J, Tudur Smith C. Carbamazepine versus phenobarbitone monotherapy for epilepsy: an individual participant data review. Cochrane Database Syst Rev 2016;12:CD001904.
63. Nolan SJ, Marson AG, Pulman J, Tudur Smith C. Phenytoin versus valproate monotherapy for partial onset seizures and generalised onset tonic-clonic seizures. Cochrane Database Syst Rev 2013;8:CD001769.
64. Nolan SJ, Muller M, Tudur Smith C, Marson AG. Oxcarbazepine versus phenytoin monotherapy for epilepsy. Cochrane Database Syst Rev 2013;5:CD003615.
65. Nolan SJ, Tudur Smith C, Pulman J, Marson AG. Phenobarbitone versus phenytoin monotherapy for partial onset seizures and generalised onset tonic-clonic seizures. Cochrane Database Syst Rev 2013;1:CD002217.

66. Nevitt SJ, Marson AG, Weston J, Tudur Smith C. Carbamazepine versus phenytoin monotherapy for epilepsy: an individual participant data review. Cochrane Database Syst Rev 2017;2:Cd001911.
67. Marson AG, Williamson PR, Hutton JL, Clough HE, Chadwick DW. Carbamazepine versus valproate monotherapy for epilepsy. Cochrane Database Syst Rev 2000(3):CD001030.
68. Nolan SJ, Sudell M, Weston J, Tudur Smith C, Marson AG. Antiepileptic drug monotherapy for epilepsy: a network meta-analysis. Cochrane Database Syst Rev 2014(12):CD011412.
69. Nevitt SJ, Sudell M, Weston J, Tudur Smith C, Marson AG. Antiepileptic drug monotherapy for epilepsy: a network meta-analysis of individual participant data. Cochrane Database Syst Rev 2017(6):CD011412.
70. World Health Organisation. Epilepsy. http://www.who.int/mediacentre/factsheets/fs999/en/ (accessed 23/05/2017). 2016.
71. Cockerell OC, Johnson AL, Sander JW, Hart YM, Shorvon SD. Remission of epilepsy: results from the National General Practice Study of Epilepsy. Lancet 1995;346(8968):140-4.
72. Electronic Medicines Compendium (eMC). Electronic Medicines Compendium (eMC); 2016 accessed (23/05/2017)]; Available from: https://www.medicines.org.uk/emc/
73. U.S. Food and Drug Administration. U.S. Food and Drug Administration.; 2016 (accessed 23/05/2017)]; Available from: http://www.fda.gov/
74. National Institute for Health and Care Excellence. The epilepsies: the diagnosis and management of the epilepsies in adults and children and primary and secondary care. NICE guidelines [CG137]. https://www.nice.org.uk/guidance/cg137 (accessed 23/05/2017); 2012.
75. Meador K, Reynolds MW, Crean S, Fahrbach K, Probst C. Pregnancy outcomes in women with epilepsy: a systematic review and meta-analysis of published pregnancy registries and cohorts. Epilepsy Res 2008;81(1):1-13.
76. Epilepsy Foundation of America I. Seizure Medication List; 2016 (accessed 23/05/2017)]; Available from: http://www.epilepsy.com/learn/treating-seizures-and-epilepsy/seizure-and-epilepsy-medicines/seizure-medication-list.
77. Rathi V, Dzara K, Gross CP, Hrynaszkiewicz I, Joffe S, Krumholz HM, et al. Sharing of clinical trial data among trialists: a cross sectional survey. BMJ 2012;345:e7570.
78. Zarin DA. Participant-level data and the new frontier in trial transparency. N Engl J Med 2013;369(5):468-9.
79. Tudur Smith C, Hopkins C, Sydes MR, Woolfall K, Clarke M, Murray G, et al. How should individual participant data (IPD) from publicly funded clinical trials be shared? BMC Med 2015;13:298.
80. Nisen P, Rockhold F. Access to patient-level data from GlaxoSmithKline clinical trials. N Engl J Med 2013;369(5):475-8.
81. Ross JS, Gross CP, Krumholz HM. Promoting transparency in pharmaceutical industry-sponsored research. Am J Public Health 2012;102(1):72-80.
82. Ross JS, Lehman R, Gross CP. The importance of clinical trial data sharing: toward more open science. Circ Cardiovasc Qual Outcomes 2012;5(2):238-40.
83. Krumholz HM, Peterson ED. Open access to clinical trials data. JAMA 2014;312(10):1002-3.
84. Lo B. Sharing clinical trial data: maximizing benefits, minimizing risk. JAMA 2015;313(8):793-4.
85. European Medicines Agency. European Medicines Agency policy on publication of clinical data for medicinal products for human use (POLICY/0070). http://www.ema.europa.eu/docs/en_GB/document_library/Other/2014/10/WC500174796.pdf (Accessed 23/05/2017); 2014.
86. Taichman DB, Backus J, Baethge C, Bauchner H, de Leeuw PW, Drazen JM, et al. Sharing Clinical Trial Data--A Proposal from the International Committee of Medical Journal Editors. N Engl J Med 2016;374(4):384-6.
87. Taichman DB, Sahni P, Pinborg A, Peiperl L, Laine C, James A, et al. Data Sharing Statements for Clinical Trials - A Requirement of the International Committee of Medical Journal Editors. N Engl J Med 2017;376(23):2277-9.

88. Rathi VK, Strait KM, Gross CP, Hrynaszkiewicz I, Joffe S, Krumholz HM, et al. Predictors of clinical trial data sharing: exploratory analysis of a cross-sectional survey. Trials 2014;15:384.
89. Savage CJ, Vickers AJ. Empirical study of data sharing by authors publishing in PLoS journals. PLoS One 2009;4(9):e7078.
90. Jones CW, Roberts BW, Platts-Mills TF. Patient Perspectives on Sharing Deidentified Trial Data. Ann Intern Med 2016;165(10):748-9.
91. Clinical Study Data Request. https://clinicalstudydatarequest.com/ (accessed 30/05/2017).
92. Hrynaszkiewicz I, Norton ML, Vickers AJ, Altman DG. Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers. BMJ 2010;340:c181.
93. YODA: the Yale School of Medicine's Open Data Access (YODA) Project (accessed 23/05/2017)]; Available from: http://yoda.yale.edu/.
94. Krumholz HM, Waldstreicher J. The Yale Open Data Access (YODA) Project--A Mechanism for Data Sharing. N Engl J Med 2016;375(5):403-5.
95. Pencina MJ, Louzao DM, McCourt BJ, Adams MR, Tayyabkhan RH, Ronco P, et al. Supporting open access to clinical trial data for researchers: The Duke Clinical Research Institute-Bristol-Myers Squibb Supporting Open Access to Researchers Initiative. Am Heart J 2016;172:64-9.
96. Mbuagbaw L, Foster G, Cheng J, Thabane L. Challenges to complete and useful data sharing. Trials 2017;18(1):71.
97. New science data-sharing rules are two scoops of disappointment (access 13/06/2017)]; Available from: https://www.statnews.com/2017/06/06/data-sharing-rules-disappoint/.
98. Empty rhetoric over data sharing slows science (accessed 13/06/2017)]; Available from: https://www.nature.com/news/empty-rhetoric-over-data-sharing-slows-science-1.22133.
99. Navar AM, Pencina MJ, Rymer JA, Louzao DM, Peterson ED. Use of Open Access Platforms for Clinical Trial Data. JAMA 2016;315(12):1283-4.
100. Hadjichrysanthou C, Cauet E, Lawrence E, Vegvari C, de Wolf F, Anderson RM. Understanding the within-host dynamics of influenza A virus: from theory to clinical implications. J R Soc Interface 2016;13(119).
101. Borges AH, Lundh A, Tendal B, Bartlett JA, Clumeck N, Costagliola D, et al. Nonnucleoside Reverse-transcriptase Inhibitor- vs Ritonavir-boosted Protease Inhibitor-based Regimens for Initial Treatment of HIV Infection: A Systematic Review and Metaanalysis of Randomized Trials. Clin Infect Dis 2016;63(2):268-80.
102. Le Noury J, Nardo JM, Healy D, Jureidini J, Raven M, Tufanaru C, et al. Restoring Study 329: efficacy and harms of paroxetine and imipramine in treatment of major depression in adolescence. BMJ 2015;351:h4320.
103. Kent DM, Nelson J, Dahabreh IJ, Rothwell PM, Altman DG, Hayward RA. Risk and treatment effect heterogeneity: re-analysis of individual participant data from 32 large clinical trials. Int J Epidemiol 2016.
104. Mospan GA, Wargo KA. 5-Day versus 10-Day Course of Fluoroquinolones in Outpatient Males with a Urinary Tract Infection (UTI). J Am Board Fam Med 2016;29(6):654-62.
105. Storgaard H, Gluud LL, Bennett C, Grondahl MF, Christensen MB, Knop FK, et al. Benefits and Harms of Sodium-Glucose Co-Transporter 2 Inhibitors in Patients with Type 2 Diabetes: A Systematic Review and Meta-Analysis. PLoS One 2016;11(11):e0166125.
106. Schwartz LM, Halloran ME, Rowhani-Rahbar A, Neuzil KM, Victor JC. Rotavirus vaccine effectiveness in low-income settings: An evaluation of the test-negative design. Vaccine 2017;35(1):184-90.
107. Voysey M, Kelly DF, Fanshawe TR, Sadarangani M, O'Brien KL, Perera R, et al. The Influence of Maternally Derived Antibody and Infant Age at Vaccination on Infant Vaccine Responses : An Individual Participant Meta-analysis. JAMA Pediatr 2017;171(7):637-46.
108. Radua J, Grunze H, Amann BL. Meta-Analysis of the Risk of Subsequent Mood Episodes in Bipolar Disorder. Psychother Psychosom 2017;86(2):90-8.
109. Ebrahim S, Sohani ZN, Montoya L, Agarwal A, Thorlund K, Mills EJ, et al. Reanalyses of randomized clinical trial data. JAMA 2014;312(10):1024-32.

110. Geifman N, Bollyky J, Bhattacharya S, Butte AJ. Opening clinical trial data: are the voluntary data-sharing portals enough? BMC Med 2015;13:280.
111. Strom BL, Buyse M, Hughes J, Knoppers BM. Data sharing, year 1--access to data from industry-sponsored clinical trials. N Engl J Med 2014;371(22):2052-4.
112. Homer P, Gross M. Data Unleashed: Cooperation Among Competitors. Pharmaceutical Executive. 2016.
113. Mayo-Wilson E, Doshi P, Dickersin K. Are manufacturers sharing data as promised? BMJ 2015;351.
114. Nolan SJ. ClinicalStudyDataRequest.com and the SAS Data Access System: An Academic Researcher's Experience. In: The 5th Clinical Trial Data Transparency Forum. , 2015; SAS Germany Headquarters; Heidelberg, Germany.
115. Clarke M, Stewart L, Pignon JP, Bijnens L. Individual patient data meta-analysis in cancer. Br J Cancer 1998;77(11):2036-44.
116. University of Oxford. EBCTCG: Early Breast Cancer Trialists' Collaborative Group. https://www.ctsu.ox.ac.uk/research/ebctcg (accessed 09/01/2017); 2017.
117. Coleman R, Powles T, Paterson A, Gnant M, Anderson S, Diel I, et al. Adjuvant bisphosphonate treatment in early breast cancer: meta-analyses of individual patient data from randomised trials. Lancet 2015;386(10001):1353-61.
118. Parmar MKB, Machin D. Survival Analysis: a Practical Approach. Second ed. Chichester: Wiley; 1995.
119. Armitage P, Berry G. Statistical Methods in Medical Research. Third ed. Oxford: Blackwell Scientific Publications; 1994.
120. Peto R, Pike MC. Conservatism of the approximation sigma (O-E)2-E in the logrank test for survival data or tumor incidence data. Biometrics 1973;29(3):579-84.
121. Deeks JJ, Higgins JPT, Altman DG, (editors). Chapter 9: Analysing data and undertaking meta-analyses. In: Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011.
122. Yuan X, Anderson SJ. Meta-analysis methodology for combining treatment effects from Cox proportional hazard models with different covariate adjustments. Biom J 2010;52(4):519-37.
123. Bowden J, Tierney JF, Simmonds M, Copas AJ, Higgins JPT. Individual patient data meta-analysis of time-to-event outcomes: one-stage versus two-stage approaches for estimating the hazard ratio under a random effects model. Res Synth Methods 2011;2(3):150-62.
124. Greenland S, Salvan A. Bias in the one-step method for pooling study results. Stat Med 1990;9(3):247-52.
125. Tudur Smith C, Williamson PR. A comparison of methods for fixed effects meta-analysis of individual patient data with time to event outcomes. Clin Trials 2007;4(6):621-30.
126. Earle CC, Pham B, Wells GA. An assessment of methods to combine published survival curves. Medical Decis Making 2000;20(1):104-11.
127. Koller MT, Stijnen T, Steyerberg EW, Lubsen J. Meta-analyses of chronic disease trials with competing causes of death may yield biased odds ratios. J Clin Epidemiol 2008;61(4):365-72.
128. Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. Trials 2007(1):16.
129. Franklin J, Pluetschow A, Paus M. Measures of the quality of individual patient survival data [abstract]. XI Cochrane Colloquium: Evidence, Health Care and Culture; 2003 Oct 26-31; Barcelona, Spain 2003:9.
130. Fine HA, Dear KBG, Loeffler JS, Mc Black PL, Canellos GP. Meta-analysis of radiation therapy with and without adjuvant chemotherapy for malignant gliomas in adults. Cancer 1993;71(8):2585-97.
131. Sedgwick P, Joekes K. Interpreting hazard ratios. BMJ 2015;351:h4631.
132. Keene ON. Alternatives to the hazard ratio in summarizing efficacy in time-to-event studies: an example from influenza trials. Stat Med 2002;21(23):3687-700.
133. Otwombe KN, Petzold M, Martinson N, Chirwa T. A review of the study designs and statistical methods used in the determination of predictors of all-cause mortality in HIV-infected cohorts: 2002-2011. PLoS One 2014;9(2):e87356.

134. Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data. Oxford: Oxford University Press; 1980.
135. Boutitie F, Gueyffier F, Pocock SJ, Boissel JP. Assessing treatment-time interaction in clinical trials with time to event data: a meta-analysis of hypertension trials. Stat Med 1998;17(24):2883-903.
136. Rondeau V, Michiels S, Liquet B, Pignon JP. Investigating trial and treatment heterogeneity in an individual patient data meta-analysis of survival data by means of the penalized maximum likelihood approach. Stat Med 2008;27(11):1894-910.
137. A'Hern RP, Smith IE, Ebbs SR. Chemotherapy and survival in advanced breast cancer: the inclusion of doxorubicin in Cooper type regimens. Br J Cancer 1993;67(4):801-5.
138. A'Hern RP, Ebbs SR, Baum MB. Does chemotherapy improve survival in advanced breast cancer? A statistical overview. Br J Cancer 1988;57(6):615-8.
139. Caubet J-F, Tosteson TD, Dong EW, Naylon EM, Whiting GW, Ernstoff MS, et al. Maximum androgen blockade in advanced prostate cancer: A meta-analysis of published randomized controlled trials using nonsteroidal antiandrogens. Urology 1997;49(1):71-8.
140. Tsiatis AA. The asymptotic joint distribution of the efficient scores test for the proportional hazards model calculated over time. Biometrika 1981;68(1):311-5.
141. Tudur C, Williamson PR, Khan S, Best LY. The value of the aggregate data approach in meta-analysis with time-to-event outcomes. J R Stat Soc Ser A 2001;164(2):357-70.
142. Collette L, Duchateau L, Suciu S, Bijnens L, Sylvester R. Including literature data in individual patient data meta-analyses for time-to-event endpoints. European Organization for Research and Treatment of Cancer (EORTC) Data Center. Draft paper; 1998. p. 1-48.
143. Glenny A, Furness S, Worthington H. The calculation of hazard ratios published in reports of randomised controlled trials; are they valid? Poster presentation. In: 19th Cochrane Colloquium, 2011 Oct 19-22; Madrid, Spain.
144. Hirooka T, Hamada C, Yoshimura I. A note on estimating treatment effect for time-to-event data in a literature-based meta-analysis. Methods Inf Med 2009;48(2):104-12.
145. Dear KB. Iterative generalized least squares for meta-analysis of survival data at multiple times. Biometrics 1994;50(4):989-1002.
146. Voest EE, van Houwelingen JC, Neijt JP. A meta-analysis of prognostic factors in advanced ovarian cancer with median survival and overall survival (measured with the log (relative risk)) as main objectives. Eur J Cancer Clin Oncol 1989;25(4):711-20.
147. Shore T, Weinerman B, Nelson N. A meta-analysis of stages I and II Hodgkin's disease. Cancer 1990;65(5):1155-60.
148. Hunink MG, Wong JB. Meta-analysis of failure-time data with adjustment for covariates. Medical Decis Making 1994;14(1):59-70.
149. Srinivasan C, Zhou M. A Note on Pooling Kaplan-Meier Estimators. Biometrics 1993;49(3):861-4.
150. Albers M, Romiti M, De Luccia N, Brochado-Neto FC, Nishimoto I, Pereira CAB. An updated meta-analysis of infrainguinal arterial reconstruction in patients with end-stage renal disease. J Vasc Surg 2007;45(3):536-42.
151. Hoogendoorn M, Hoogenveen RT, Rutten-van Molken MP, Vestbo J, Feenstra TL. Case fatality of COPD exacerbations: a meta-analysis and statistical modelling approach. Eur Respir J 2011;37(3):508-15.
152. Coplen SE, Antman EM, Berlin JA, Hewitt P, Chalmers TC. Efficacy and safety of quinidine therapy for maintenance of sinus rhythm after cardioversion. A meta-analysis of randomized control trials [published erratum appears in Circulation 1991 Feb;83(2):714]. Circulation 1990;82(4):1106-16.
153. Messori A, Trippoli S, Vaiani M, Cattel F. Survival meta-analysis of individual patient data and survival meta-analysis of published (aggregate) data: Is there an intermediate approach between these two opposite options? Clin Drug Investig 2000(5):309-16.
154. Messori A, Rampazzo R. Meta-analysis of clinical trials based on censored end-points: simplified theory and implementation of the statistical algorithms on a microcomputer. Comput Methods Programs Biomed 1993(4):261-7.

155. Arends LR, Hunink MG, Stijnen T. Meta-analysis of summary survival curve data. Stat Med 2008(22):4381-96.
156. Vale CL, Tierney JF, Stewart LA. Effects of adjusting for censoring on meta-analyses of time-to-event outcomes. Int J Epidemiol 2002(1):107-11.
157. Guyot P, Ades AE, Ouwens MJNM, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol 2012;12:9-.
158. Ouwens MJNM, Philips Z, Jansen JP. Network meta-analysis of parametric survival curves. Res Synth Methods 2010;1(3-4):258-71.
159. Jansen JP. Network meta-analysis of survival data with fractional polynomials. BMC Med Res Methodol 2011;11:61.
160. Michiels S, Piedbois P, Burdett S, Syz N, Stewart L, Pignon JP. Meta-analysis when only the median survival times are known: a comparison with individual patient data results. Int J Technol Assess Health Care 2005(1):119-25.
161. Riley RD, Abrams KR, Sutton AJ, Lambert PC, Jones DR, Heney D, et al. Reporting of prognostic markers: current problems and development of guidelines for evidence-based practice in the future. Br J Cancer 2003;88(8):1191-8.
162. Imamura M, MacLennan S, Dahm P, Neuberger M, Omar M, MacLennan G. The hazards of extracting hazard ratios from survival curves for meta-analyses in observational studies. Poster presentation. In: 19th Cochrane Colloquium, 2011 Oct 19-22; Madrid, Spain.
163. D'Amico R, Deeks J, Altman D. Effect of length of follow-up on meta-analysis [abstract]. In: Sixth International Cochrane Colloquium; 1998 Oct 22-26; Baltimore, MD, USA, 1998: p. 51.
164. Duchateau L, Collette L, Sylvester R, Pignon JP. Estimating number of events from the Kaplan-Meier curve for incorporation in a literature-based meta-analysis: what you don't see you can't get! Biometrics 2000;56(3):886-92.
165. Duchateau L, Pignon JP, Bijnens L, Bertin S, Bourhis J, Sylvester R. Individual patient- versus literature-based meta-analysis of survival data: time to event and event rate at a particular time can make a difference, an example based on head and neck cancer. Control Clin Trials 2001(5):538-47.
166. Bennett DA. Review of analytical methods for prospective cohort studies using time to event data: single studies and implications for meta-analysis. Stat Methods Med Res 2003;12(4):297-319.
167. Stewart GB, Altman DG, Askie LM, Duley L, Simmonds MC, Stewart LA. Statistical analysis of individual participant data meta-analyses: a comparison of methods and recommendations for practice. PLoS One 2012;7(10):e46042.
168. Mathew T, Nordström K. Comparison of One-Step and Two-Step Meta-Analysis Models Using Individual Patient Data. Biom J 2010;52(2):271-87.
169. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. Stat Med 2017;36(5):855-75.
170. Arends LR, Voko Z, Stijnen T. Combining multiple outcome measures in a meta-analysis: an application. Stat Med 2003(8):1335-53.
171. O'Quigley J, Stare J. Proportional hazards models with frailties and random effects. Stat Med 2002;21(21):3219-33.
172. Sargent DJ. A General Framework for Random Effects Survival Analysis in the Cox Proportional Hazards Setting. Biometrics 1998;54(4):1486-97.
173. Williamson PR, Clough HE, Hutton JL, Marson AG, Chadwick DW. Statistical issues in the assessment of the evidence for an interaction between factors in epilepsy trials. Stat Med 2002;21(18):2613-22.
174. Tudur Smith C, Williamson PR, Marson AG. An overview of methods and empirical comparison of aggregate data and individual patient data results for investigating heterogeneity in meta-analysis of time-to-event outcomes. J Eval Clin Pract 2005;11(5):468-78.
175. Hozo I, Djulbegovic B. Individual-patient data meta-analysis (IPD MA) in the presence of competing risks [abstract]. In: XIII Cochrane Colloquium; 2005 Oct 22-26; Melbourne, Australia, 2005.

176. Thall PF, Lachin JM. Assessment of stratum-covariate interactions in Cox's proportional hazards regression model. Stat Med 1986;5(1):73-83.
177. Hess KR. Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions. Stat Med 1994;13(10):1045-62.
178. Sargent DJ, Zee B, Milan C, Torri V, G. F. Meta-analysis of individual patient survival data using random effects models. In: Stangl D, Berry DA. Meta-analysis in medicine and health policy. New York: Marcel Dekker; 2000. p. 255-75.
179. Yau KK, McGilchrist CA. ML and REML estimation in survival analysis with time dependent correlated frailty. Stat Med 1998;17(11):1201-13.
180. Yamaguchi T, Ohashi Y, Matsuyama Y. Proportional hazards models with random effects to examine centre effects in multicentre cancer clinical trials. Stat Methods Med Res 2002;11(3):221-36.
181. Gustafson P. A Bayesian analysis of bivariate survival data from a multicentre cancer clinical trial. Stat Med 1995;14(23):2523-35.
182. Matsuyama Y, Sakamoto J, Ohashi Y. A Bayesian hierarchical survival model for the institutional effects in a multi-centre cancer clinical trial. Stat Med 1998;17(17):1893-908.
183. Bennett MM, Crowe BJ, Price KL, Stamey JD, Seaman JW, Jr. Comparison of Bayesian and frequentist meta-analytical approaches for analyzing time to event data. J Biopharm Stat 2013;23(1):129-45.
184. Thompson S, Kaptoge S, White I, Wood A, Perry P, Danesh J, et al. Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies. Int J Epidemiol 2010;39(5):1345-59.
185. Fisher DJ, Copas AJ, Tierney JF, Parmar MK. A critical review of methods for the assessment of patient-level interactions in individual participant data meta-analysis of randomized trials, and guidance for practitioners. J Clin Epidemiol 2011;64(9):949-67.
186. Michiels S, Baujat B, Mahe C, Sargent DJ, Pignon JP. Random effects survival models gave a better understanding of heterogeneity in individual patient data meta-analyses. J Clin Epidemiol 2005;58(3):238-45.
187. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. Br J Cancer 1977;35(1):1-39.
188. Katsahian S, Latouche A, Mary J-Y, Chevret S, Porcher R. Practical methodology of meta-analysis of individual patient data using a survival outcome. Contemp Clin Trials 2008;29(2):220-30.
189. Hougaard P. Analysis of Multivariate Survival Data. New York: Springer-Verlag; 2000.
190. Vaida F, Xu R. Proportional hazards model with random effects. Stat Med 2000;19(24):3309-24.
191. Simmonds MC, Higgins JP, Stewart LA. Random-effects meta-analysis of time-to-event data using the expectation-maximisation algorithm and shrinkage estimators. Res Synth Methods 2013;4(2):144-55.
192. Massonnet G, Janssen P, Burzykowski T. Fitting conditional survival models to meta-analytic data by using a transformation toward mixed-effects models. Biometrics 2008;64(3):834-42.
193. Crowther MJ, Look MP, Riley RD. Multilevel mixed effects parametric survival models using adaptive Gauss-Hermite quadrature with application to recurrent events and individual participant data meta-analysis. Stat Med 2014;33(22):3844-58.
194. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Stat Med 2002;21(15):2175-97.
195. Simmonds MC, Tierney J, Bowden J, Higgins JP. Meta-analysis of time-to-event data: a comparison of two-stage methods. Res Synth Methods 2011;2(3):139-49.
196. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemother Rep 1966;50(3):163.

197. Pennells L, Kaptoge S, White IR, Thompson SG, Wood AM, Emerging Risk Factors C. Assessing risk prediction models using individual participant data from multiple studies. Am J Epidemiol 2014;179(5):621-32.
198. Andersen PK, Gill RD. Cox's Regression Model for Counting Processes: A Large Sample Study. Ann Stat 1982;10(4):1100-20.
199. McGilchrist CA. REML estimation for survival models with frailty. Biometrics 1993;49(1):221-5.
200. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis: A Global Perspective. Upper Saddle River: Prentice Hall 2009.
201. Becker BJ, L.V. H, Pigott TD. Multivariate Meta-Analysis. San Diego, CA: Academic Press; 2000.
202. Rondeau V, Pignon J-P, Michiels S. A joint model for the dependence between clustered times to tumour progression and deaths: A meta-analysis of chemotherapy in head and neck cancer. Stat Methods Med Res 2011.
203. Riley RD, Thompson JR, Abrams KR. An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown. Biostatistics 2008;9(1):172-86.
204. Riley RD, Abrams KR, Sutton AJ, Lambert PC, Thompson JR. Bivariate random-effects meta-analysis and the estimation of between-study correlation. BMC Med Res Methodol 2007;7:3.
205. Riley RD, Abrams KR, Lambert PC, Sutton AJ, Thompson JR. An evaluation of bivariate random-effects meta-analysis for the joint synthesis of two correlated outcomes. Stat Med 2007;26(1):78-97.
206. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. Stat Med 2002;21(4):589-624.
207. Daniels MJ, Hughes MD. Meta-analysis for the evaluation of potential surrogate markers. Stat Med 1997;16(17):1965-82.
208. Alonso A, Molenberghs G. Evaluating time to cancer recurrence as a surrogate marker for survival from an information theory perspective. Stat Methods Med Res 2008;17(5):497-504.
209. Ghosh D, Taylor JMG, Sargent DJ. Meta-analysis for Surrogacy: Accelerated Failure Time Models and Semicompeting Risks Modeling. Biometrics 2012;68(1):226-32.
210. Renfro LA, Shi Q, Sargent DJ, Carlin BP. Bayesian adjusted R2 for the meta-analytic evaluation of surrogate time-to-event endpoints in clinical trials. Stat Med 2012;31(8):743-61.
211. Burzykowski T, Molenberghs G, Buyse M, Geys H, Renard D. Validation of Surrogate End Points in Multiple Randomized Clinical Trials with Failure Time End Points. J R Stat Soc Ser C Appl Stat 2001;50(4):405-22.
212. Renard D, Geys H, Molenberghs G, Burzykowski T, Buyse M, Vangeneugden T, et al. Validation of a longitudinally measured surrogate marker for a time-to-event endpoint. J Appl Stat 2003;30:235-47.
213. Raudenbush SW, Becker BJ, Kalaian H. Modeling multivariate effect sizes. Psychol Bull 1988;103(1):111-20.
214. Sargent DJ, Patiyil S, Yothers G, Haller DG, Gray R, Benedetti J, et al. End points for colon cancer adjuvant trials: observations and recommendations based on individual patient data from 20,898 patients enrolled onto 18 randomized trials from the ACCENT Group. J Clin Oncol 2007(29):4569-74.
215. Fiocco M, Stijnen T, Putter H. Meta-analysis of time-to-event outcomes using a hazard-based approach: Comparison with other models, robustness and meta-regression. Comput Stat Data Anal 2012;56(5):1028-37.
216. Fiocco M, Putter H, Van Houwelingen JC. A new serially correlated gamma-frailty process for longitudinal count data. Biostatistics 2009;10(2):245-57.
217. Fiocco M, Putter H, van Houwelingen JC. Meta-analysis of pairs of survival curves under heterogeneity: a Poisson correlated gamma-frailty approach. Stat Med 2009;28(30):3782-97.
218. Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F, Colditz GA. Meta-analysis of multiple outcomes by regression with random effects. Stat Med 1998;17(22):2537-50.

219. Jackson D, Rollins K, Coughlin P. A multivariate model for the meta-analysis of study level survival data at multiple times. Res Synth Methods 2014;5(3):264-72.
220. White IR, Barrett JK, Jackson D, Higgins JP. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. Res Synth Methods 2012;3(2):111-25.
221. Woods BS, Hawkins N, Scott DA. Network meta-analysis on the log-hazard scale, combining count and hazard ratio statistics accounting for multi-arm trials: a tutorial. BMC Med Res Methodol 2010;10:54.
222. Welton NJ, Willis SR, Ades AE. Synthesis of survival and disease progression outcomes for health technology assessment of cancer therapies. Res Synth Methods 2010;1(3-4):239-57.
223. Cope S, Jansen JP. Quantitative summaries of treatment effect estimates obtained with network meta-analysis of survival curves to inform decision-making. BMC Med Res Methodol 2013;13:147.
224. Wei Y, Royston P, Tierney JF, Parmar MK. Meta-analysis of time-to-event outcomes from randomized trials using restricted mean survival time: application to individual participant data. Stat Med 2015;34(21):2881-98.
225. Lueza B, Rotolo F, Bonastre J, Pignon JP, Michiels S. Bias and precision of methods for estimating the difference in restricted mean survival time from an individual patient data meta-analysis. BMC Med Res Methodol 2016;16:37.
226. Simes RJ. Confronting publication bias: a cohort design for meta analysis. Stat Med 1987(1):11-29.
227. Zang J, Xu J, Xiang C, He J, Zou S. Multi-level model synthesis of median survival time in meta-analysis. Epidemiology 2015;26(1):e2-3.
228. Siannis F, Barrett JK, Farewell VT, Tierney JF. One-stage parametric meta-analysis of time-to-event outcomes. Stat Med 2010;29(29):3030-45.
229. Uno H, Claggett B, Tian L, Inoue E, Gallo P, Miyata T, et al. Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. J Clin Oncol 2014;32(22):2380-5.
230. Royston P, Parmar MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol 2013;13:152.
231. Zhao L, Claggett B, Tian L, Uno H, Pfeffer MA, Solomon SD, et al. On the restricted mean survival time curve in survival analysis. Biometrics 2016;72(1):215-21.
232. Moodie PF, Nelson NA, Koch GG. A non-parametric procedure for evaluating treatment effect in the meta-analysis of survival data. Stat Med 2004(7):1075-93.
233. Barrett JK, Farewell VT, Siannis F, Tierney J, Higgins JP. Two-stage meta-analysis of survival data from individual participants using percentile ratios. Stat Med 2012;31(30):4296-308.
234. Combescure C, Foucher Y, Jackson D. Meta-analysis of single-arm survival studies: a distribution-free approach for estimating summary survival curves with random effects. Stat Med 2014;33(15):2521-37.
235. Floriani I, D'Amico R, Telaro E, Tinazzi A, Torri V. Evaluation of the summary statistics used for meta-analyses using literature survival data in oncology [abstract]. In: Second International Conference Scientific Basis of Health Services & Fifth Annual Cochrane Colloquium; 1997 Oct 8-12; Amsterdam, The Netherlands, 1997.
236. Mathoulin-Pelissier S, Gourgou-Bourgade S, Bonnetain F, Kramar A. Survival End Point Reporting in Randomized Cancer Clinical Trials: A Review of Major Journals. J Clin Oncol 2008;26(22):3721-6.
237. Arkenau H-T, Nordman I, Dobbins T, Ward R. Reporting time-to-event endpoints and response rates in 4 decades of randomized controlled trials in advanced colorectal cancer. Cancer 2011;117(4):832-40.
238. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis part II: multivariate data analysis--an introduction to concepts and methods. Br J Cancer 2003;89(3):431-6.
239. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis Part III: multivariate data analysis -- choosing a model and assessing its adequacy and fit. Br J Cancer 2003;89(4):605-11.

240. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part IV: further concepts and methods in survival analysis. Br J Cancer 2003;89(5):781-6.
241. Pocock SJ, Clayton TC, Altman DG. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. Lancet 2002;359(9318):1686-9.
242. Abraira V, Muriel A, Emparanza JI, Pijoan JI, Royuela A, Plana MN, et al. Reporting quality of survival analyses in medical journals still needs improvement. A minimal requirements proposal. J Clin Epidemiol 2013;66(12):1340-6.e5.
243. Batson S, Greenall G, Hudson P. Review of the Reporting of Survival Analyses within Randomised Controlled Trials and the Implications for Meta-Analysis. PLoS One 2016;11(5):e0154870.
244. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA 1996;276(8):637-9.
245. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The Revised CONSORT Statement for Reporting Randomized Trials: Explanation and Elaboration. Ann Intern Med 2001;134(8):663-94.
246. Moher D, Hopewell S, Schulz K, Montori V, Gøtzsche P, Devereaux P, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340.
247. Guyot P, Welton NJ, Ouwens MJNM, Ades AE. Survival time outcomes in randomized, controlled trials and meta-analyses: the parallel universes of efficacy and cost-effectiveness. Value Health 2011;14(5):640-6.
248. Higgins JPT, Green S, (editors). Cochrane Handbook of Systematic Reviews of Interventions: Version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011.
249. ILAE Commission on Antiepileptic Drugs. Considerations on Designing Clinical Trials to Evaluate the Place of New Antiepileptic Drugs in the Treatment of Newly Diagnosed and Chronic Patients with Epilepsy. Epilepsia 1998;37(7):799-803.
250. Kwan P, Brodie MJ, Kalviainen R, Yurkewicz L, Weaver J, Knapp LE. Efficacy and safety of pregabalin versus lamotrigine in patients with newly diagnosed partial seizures: a phase 3, double-blind, randomised, parallel-group trial. Lancet Neurol 2011;10(10):881-90.
251. European Medicines Agency Committee for Proprietary Medicinal Products (CPMP). CHMP Guideline on clinical investigation of medicinal products in the treatment of epileptic disorders. CHMP; 2010 (accessed 23/05/2017)]; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/01/WC500070043.pdf
252. Glauser T, Ben-Menachem E, Bourgeois B, Cnaan A, Chadwick D, Guerreiro C, et al. ILAE treatment guidelines: Evidence-based analysis of antiepileptic drug efficacy and effectiveness as initial monotherapy for epileptic seizures and syndromes. Epilepsia 2006;47(7):1094-120.
253. Glauser T, Ben-Menachem E, Bourgeois B, Cnaan A, Guerreiro C, Kälviäinen R, et al. Updated ILAE evidence review of antiepileptic drug efficacy and effectiveness as initial monotherapy for epileptic seizures and syndromes. Epilepsia 2013;54(3):551-63.
254. ICH harmonised tripartite guideline (E9). International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceutical for Human Use. Statistical principles for clinical trials; 1998 (accessed 23/05/2017)]; Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf.
255. Committee for Proprietary Medicinal Products (CPMP). Note for guidance on clinical investigation of medicinal products in the treatment of epileptic disorders. CPMP; 2000 (accessed 23/05/2017)]; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003529.pdf.
256. de Silva M, MacArdle B, McGowan M, Hughes E, Stewart J, Neville BG, et al. Randomised comparative monotherapy trial of phenobarbitone, phenytoin, carbamazepine, or sodium valproate for newly diagnosed childhood epilepsy. Lancet 1996;347(9003):709-13.

257. Placencia M, Sander JW, Shorvon SD, Roman M, Alarcon F, Bimos C, et al. Antiepileptic drug treatment in a community health care setting in northern Ecuador: a prospective 12-month assessment. Epilepsy Res 1993;14(3):237-44.
258. Cereghino JJ, Brock JT, Van Meter JC, Penry JK, Smith LD, White BG. Carbamazepine for epilepsy. A controlled prospective evaluation. Neurology 1974;24(5):401-10.
259. Shakir RA, Johnson RH, Lambie DG, Melville ID, Nanda RN. Comparison of sodium valproate and phenytoin as single drug treatment in epilepsy. Epilepsia 1981;22(1):27-33.
260. Verity CM, Hosking G, Easter DJ. A multicentre comparative trial of sodium valproate and carbamazepine in paediatric epilepsy. The Paediatric EPITEG Collaborative Group. Dev Med Child Neurol 1995;37(2):97-108.
261. Brodie MJ, Perucca E, Ryvlin P, Ben-Menachem E, Meencke HJ, Levetiracetam Monotherapy Study G. Comparison of levetiracetam and controlled-release carbamazepine in newly diagnosed epilepsy. Neurology 2007;68(6):402-8.
262. Ramsay E, Faught E, Krumholz A, Naritoku D, Privitera M, Schwarzman L, et al. Efficacy, tolerability, and safety of rapid initiation of topiramate versus phenytoin in patients with new-onset epilepsy: a randomized double-blind clinical trial. Epilepsia 2010;51(10):1970-7.
263. Heller AJ, Chesterman P, Elwes RD, Crawford P, Chadwick D, Johnson AL, et al. Phenobarbitone, phenytoin, carbamazepine, or sodium valproate for newly diagnosed adult epilepsy: a randomised comparative monotherapy trial. J Neurol Neurosurg Psychiatry 1995;58(1):44-50.
264. Marson AG, Al-Kharusi AM, Alwaidh M, Appleton R, Baker GA, Chadwick DW, et al. The SANAD study of effectiveness of carbamazepine, gabapentin, lamotrigine, oxcarbazepine, or topiramate for treatment of partial epilepsy: an unblinded randomised controlled trial. Lancet 2007;369(9566):1000-15.
265. Marson AG, Al-Kharusi AM, Alwaidh M, Appleton R, Baker GA, Chadwick DW, et al. The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial. Lancet 2007;369(9566):1016-26.
266. Reynolds EH, Shorvon SD. Monotherapy or polytherapy for epilepsy? Epilepsia 1981;22(1):1-10.
267. Baulac M, Brodie MJ, Patten A, Segieth J, Giorgi L. Efficacy and tolerability of zonisamide versus controlled-release carbamazepine for newly diagnosed partial epilepsy: a phase 3, randomised, double-blind, non-inferiority trial. Lancet Neurol 2012;11(7):579-88.
268. Leone MA, Giussani G, Nolan SJ, Marson AG, Beghi E. Immediate antiepileptic drug treatment, versus placebo, deferred, or no treatment for first unprovoked seizure. Cochrane Database Syst Rev 2016(5).
269. Mattson RH, Cramer JA, Collins JF, Smith DB, Delgado-Escueta AV, Browne TR, et al. Comparison of carbamazepine, phenobarbital, phenytoin, and primidone in partial and secondarily generalized tonic-clonic seizures. N Engl J Med 1985;313(3):145-51.
270. Mattson RH, Cramer JA, Collins JF. A comparison of valproate with carbamazepine for the treatment of complex partial seizures and secondarily generalized tonic-clonic seizures in adults. The Department of Veterans Affairs Epilepsy Cooperative Study No. 264 Group. N Engl J Med 1992;327(11):765-71.
271. Mikkelsen B, Berggreen P, Joensen P, Kristensen O, Kohler O, Mikkelsen BO. Clonazepam (Rivotril) and carbamazepine (Tegretol) in psychomotor epilepsy: a randomized multicenter trial. Epilepsia 1981;22(4):415-20.
272. Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. BMC Med 2010;8:24.
273. Jaspers GJ, Degraeuwe PL. A failed attempt to conduct an individual patient data meta-analysis. Syst Rev 2014;3:97.
274. Hrobjartsson A. Why did it take 19 months to retrieve clinical trial data from a non-profit organisation? BMJ 2013;347:f6927.

275. Vickers AJ. Whose data set is it anyway? Sharing raw data from randomized trials. Trials 2006;7:15.
276. Kovalchik SA. Survey finds that most meta-analysts do not attempt to collect individual patient data. J Clin Epidemiol 2012;65(12):1296-9.
277. Nolan SJ, Marson AG, Tudur Smith C. The changing world of data sharing and data transparency: what does this mean for individual participant data reviews? In: 23rd Cochrane Colloquium, 2015 3-7 October 2015; Vienna, Austria.
278. Nevitt SJ, Marson AG, Davie B, Reynolds S, Williams L, Smith CT. Exploring changes over time and characteristics associated with data retrieval across individual participant data meta-analyses: systematic review. BMJ 2017;357:j1390.
279. Papke LE, Wooldridge J. Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. National Bureau of Economic Research Cambridge, Mass., USA; 1993.
280. Stengel D, Bauwens K, Rademacher G, Ekkernkamp A, Guthoff C. Emergency ultrasound-based algorithms for diagnosing blunt abdominal trauma. Cochrane Database Syst Rev 2013(7):Cd004446.
281. Zanchetti A, Mancia G. Searching for information from unreported trials--amnesty for the past and prospective meta-analyses for the future. J Hypertens 1998;16(2):125.
282. Tierney JF, Vale C, Riley R, Smith CT, Stewart L, Clarke M, et al. Individual Participant Data (IPD) Meta-analyses of Randomised Controlled Trials: Guidance on Their Use. PLoS Med 2015;12(7):e1001855.
283. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, et al. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. JAMA 2015;313(16):1657-65.
284. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. Syst Rev 2012;1:2.
285. Tudur Smith C, Marson AG, Chadwick DW, Williamson PR. Multiple treatment comparisons in epilepsy monotherapy trials. Trials 2007;8:34.
286. Engel J, Pedley TA, Aicardi J. Antiepileptic Drugs. In: Epilepsy: A Comprehensive Textbook. In: Lippincott Williams & Wilkins; 2008. vol 3 p. 1488-9.
287. Liporace JD, Sperling MR, Dichter MA. Absence seizures and carbamazepine in adults. Epilepsia 1994;35(5):1026-8.
288. Snead OC, 3rd, Hosey LC. Exacerbation of seizures in children by carbamazepine. N Engl J Med 1985;313(15):916-21.
289. Chadwick DW, Anhut H, Greiner MJ, Alexander J, Murray GH, Garofalo EA, et al. A double-blind trial of gabapentin monotherapy for newly diagnosed partial seizures. International Gabapentin Monotherapy Study Group 945-77. Neurology 1998;51(5):1282-8.
290. Brodie MJ, Richens A, Yuen AW. Double-blind comparison of lamotrigine and carbamazepine in newly diagnosed epilepsy. UK Lamotrigine/Carbamazepine Monotherapy Trial Group.[Erratum appears in Lancet 1995 Mar 11;345(8950):662]. Lancet 1995;345(8948):476-9.
291. Tudur Smith C, Marson AG, Williamson PR. Phenytoin versus valproate monotherapy for partial onset seizures and generalized onset tonic-clonic seizures. Cochrane Database Syst Rev 2001(4):CD001769.
292. Taylor S, Tudur Smith C, Williamson PR, Marson AG. Phenobarbitone versus phenytoin monotherapy for partial onset seizures and generalized onset tonic-clonic seizures. Cochrane Database Syst Rev 2001(4):CD002217.
293. Tudur Smith C, Marson AG, Williamson PR. Carbamazepine versus phenobarbitone monotherapy for epilepsy. Cochrane Database Syst Rev 2003(1):CD001904.
294. Tudur Smith C, Marson AG, Clough HE, Williamson PR. Carbamazepine versus phenytoin monotherapy for epilepsy. Cochrane Database Syst Rev 2002(2):CD001911.
295. Muller M, Marson AG, Williamson PR. Oxcarbazepine versus phenytoin monotherapy for epilepsy. Cochrane Database Syst Rev 2006(2):CD003615.
296. Gamble CL, Williamson PR, Marson AG. Lamotrigine versus carbamazepine monotherapy for epilepsy. Cochrane Database Syst Rev 2006(1):CD001031.

297. Higgins JPT, Lasserson T, Chandler J, Tovey D, R. C. Methodological Expectations of Cochrane Intervention Reviews. London: Cochrane; 2016.
298. European Union Clinical Trials Directive. Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32001L0020&from=EN (accessed 23/05/2017); 2001.
299. Chaimani A, Higgins JP, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. PLoS One 2013;8(10):e76654.
300. Tudur Smith C, Nolan S, Appelbe D, Appleton R, Dixon P, Harrison J, et al. Resource implications of preparing Individual Participant Data from a clinical trial to share with external researchers Trials 2017;18(1):319.
301. Laurie AL, Chen Y, Chou R, Fu R. Meta-analysis of the Impact of Patient Characteristics on Estimates of Effectiveness and Harms of Recombinant Human Bone Morphogenetic Protein-2 in Lumbar Spinal Fusion. Spine 2016;41(18):E1115-23.
302. Noshchenko A, Lindley EM, Burger EL, Cain CM, Patel VV. What Is the Clinical Relevance of Radiographic Nonunion After Single-Level Lumbar Interbody Arthrodesis in Degenerative Disc Disease?: A Meta-Analysis of the YODA Project Database. Spine 2016;41(1):9-17.
303. El Emam K. Guide to the De-Identification of Personal Health Information: CRC Press (Auerbach); 2013.
304. Privitera MD, Brodie MJ, Mattson RH, Chadwick DW, Neto W, Wang S, et al. Topiramate, carbamazepine and valproate monotherapy: double-blind comparison in newly diagnosed epilepsy. Acta Neurol Scand 2003;107(3):165-75.
305. Trinka E, Marson AG, Van Paesschen W, Kälviäinen R, Marovac J, Duncan B, et al. KOMET: An unblinded, randomised, two parallel-group, stratified trial comparing the effectiveness of levetiracetam with controlled-release carbamazepine and extended-release sodium valproate as monotherapy in patients with newly diagnosed epilepsy. J Neurol Neurosurg Psychiatry 2012.
306. SAS. Insitute Inc. SAS version 9.3 for Windows. Cary, NC, USA: SAS Institute Inc.; 2011.
307. StataCorp. Stata Statistical Software Release 14. College Station, TX: StataCorp LP; 2015.
308. White IR. Multivariate random-effects meta-analysis. Stata J 2009;9(1):40-56.
309. Simmonds MC, Higgins JP. Covariate heterogeneity in meta-analysis: criteria for deciding between meta-regression and individual patient data. Stat Med 2007;26(15):2982-99.
310. White IR. Network meta-analysis. Stata J 2015;15(4):951-85.
311. Jackson D, White IR, Riley RD. Quantifying the impact of between-study heterogeneity in multivariate meta-analyses. Stat Med 2012;31(29):3805-20.
312. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21(11):1539-58.
313. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol 2011;64(2):163-71.
314. Higgins J, P. T., Thompson S, G., Deeks J, J. , Altman D, G. . Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-60.
315. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. Stat Med 2010;29(7-8):932-44.
316. Higgins JP, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. Res Synth Methods 2012;3(2):98-110.
317. Malafosse A, Genton P, Hirsch E, Marescaux C, Broglin D, R. B. Idiopathic generalised epilepsies: clinical, experimental and genetic. Eastleigh: John Libbey and Company; 1994.
318. Nolan SJ, Marson AG, Weston J, Tudur Smith C. Carbamazepine versus phenobarbitone monotherapy for epilepsy: an individual participant data review. Cochrane Database Syst Rev 2015;7:CD001904.

319. Biton V, Mirza W, Montouris G, Vuong A, Hammer AE, Barrett PS. Weight change associated with valproate and lamotrigine monotherapy in patients with epilepsy. Neurology 2001;56(2):172-7.
320. National Institute of Neurological Disorders and Stroke (NINDS). The Epilepsies and Seizures: Hope Through Research. https://catalog.ninds.nih.gov/pubstatic//15-156/15-156.pdf (accessed 24/01/2017).
321. Proposal for revised clinical and electroencephalographic classification of epileptic seizures. From the Commission on Classification and Terminology of the International League Against Epilepsy. Epilepsia 1981;22(4):489-501.
322. Proposal for revised classification of epilepsies and epileptic syndromes. Commission on Classification and Terminology of the International League Against Epilepsy. Epilepsia 1989;30(4):389-99.
323. Berg AT, Berkovic SF, Brodie MJ, Buchhalter J, Cross JH, van Emde Boas W, et al. Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005-2009. Epilepsia 2010;51(4):676-85.
324. Joseph PD, Craig JC, Caldwell PH. Clinical trials in children. Br J Clin Pharmacol 2015;79(3):357-69.
325. Nolan SJ, Sutton L, Marson A, Tudur Smith C. Consistency of outcome and statistical reporting of time-to-event data: the impact on Cochrane Reviews and meta-analyses in epilepsy. In: 21st Cochrane Colloquium: Better Knowledge for Better Health, 2013; Quebec City, Canada
326. Jackson C, Best N, Richardson S. Improving ecological inference using individual-level data. Stat Med 2006;25(12):2136-59.
327. Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. Stat Med 2002;21(3):371-87.
328. Saramago P, Woods B, Weatherly H, Manca A, Sculpher M, Khan K, et al. Methods for network meta-analysis of continuous outcomes using individual patient data: a case study in acupuncture for chronic pain. BMC Med Res Methodol 2016;16(1):131.
329. Goldstein H, Yang M, Omar R, Turner R, Thompson S. Meta-analysis using multilevel models with an application to the study of class size effects. J R Stat Soc Ser C Appl Stat 2000;49(3):399-412.
330. Ravva P, Karlsson MO, French JL. A linearization approach for the model-based analysis of combined aggregate and individual patient data. Stat Med 2014;33(9):1460-76.
331. Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. Int J Epidemiol 2005;34(1):79-87.
332. Jackson C, Richardson S. Hierarchical related regression for combining aggregate and individual data in studies of socio-economic disease risk factors. J R Stat Soc Ser A 2008;171(1):159-78.
333. Dias S, Welton, N. J., Sutton, A. J., & Ades, A. E. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials.; 2011 (accessed 30/05/2017)]; Available from: http://research-information.bristol.ac.uk/files/7215331/TSD2_General_meta_analysis.final.08.05.12.pdf.

Appendices

Appendix 1: Search strategies and reference list for Chapter 2

Search strategy for Cochrane Methodology Register (searched up to July 2012)

#1 (meta-analysis):ti,ab,kw in Methods Studies

#2 (time-to-event):ti,ab,kw in Methods Studies

#3 (survival analysis):ti,ab,kw in Methods Studies

#4 (survival data):ti,ab,kw in Methods Studies

#5 (survival studies):ti,ab,kw in Methods Studies

#6 (failure time):ti,ab,kw in Methods Studies

#7 (#2 or #3 or #4 or #5 or #6)

#8 (#1 and #7)

#9 (review methodology):ti,ab,kw in Methods Studies

#10 (#8 and #9)

#11 (longitudinal):ti,ab,kw in Methods Studies

#12 (#10 and not #11)

Search strategy for MEDLINE (searched from 1946 up to 24 January 2017)

1 *Meta-Analysis as Topic/

2 (meta-analysis or meta-analyses).ti,ab.

3 1 or 2

4 "time-to-event".ti,ab.

5 (failure time or failure time data).ti,ab.

6 (survival analys$ or survival data or survival study).ti,ab.

7 *Survival Analysis/

8 4 or 5 or 6 or 7

9 3 and 8

10 Models, Statistical/

11 Proportional Hazards Models/

12 10 or 11

13 9 and 12

References of articles identified for Chapter 2

115 articles were identified by electronic searching and hand searching for the methodological review described in Chapter 2 (see Chapter 2.3.1 for further details).

1. Abel UR, Edler L. A pitfall in the meta-analysis of hazard ratios. Control Clin Trials 1988;9(2):149-51.
2. A'Hern RP, Ebbs SR, Baum MB. Does chemotherapy improve survival in advanced breast cancer? A statistical overview. Br J Cancer 1988;57(6):615-8.
3. A'Hern RP, Smith IE, Ebbs SR. Chemotherapy and survival in advanced breast cancer: the inclusion of doxorubicin in Cooper type regimens. Br J Cancer 1993;67(4):801-5.
4. Albers M, Romiti M, De Luccia N, Brochado-Neto FC, Nishimoto I, Pereira CAB. An updated meta-analysis of infrainguinal arterial reconstruction in patients with end-stage renal disease. J Vasc Surg 2007;45(3):536-42.
5. Alonso A, Molenberghs G. Evaluating time to cancer recurrence as a surrogate marker for survival from an information theory perspective. Stat Methods Med Res 2008;17(5):497-504.
6. Arends LR, Hunink MG, Stijnen T. Meta-analysis of summary survival curve data. Stat Med 2008(22):4381-96.
7. Arends LR, Voko Z, Stijnen T. Combining multiple outcome measures in a meta-analysis: an application. Stat Med 2003(8):1335-53.
8. Babbs CF. Statistical analysis of joint short-term and long-term survival in resuscitation research. Resuscitation 2007;75(2):323-31.
9. Barrett JK, Farewell VT, Siannis F, Tierney J, Higgins JP. Two-stage meta-analysis of survival data from individual participants using percentile ratios. Stat Med 2012;31(30):4296-308.
10. Bennett DA. Review of analytical methods for prospective cohort studies using time to event data: single studies and implications for meta-analysis. Stat Methods Med Res 2003;12(4):297-319.
11. Bennett MM, Crowe BJ, Price KL, Stamey JD, Seaman JW, Jr. Comparison of Bayesian and frequentist meta-analytical approaches for analyzing time to event data. J Biopharm Stat 2013;23(1):129-45.
12. Boutitie F, Gueyffier F, Pocock SJ, Boissel JP. Assessing treatment-time interaction in clinical trials with time to event data: a meta-analysis of hypertension trials. Stat Med 1998;17(24):2883-903.
13. Bowden J, Tierney JF, Simmonds M, Copas AJ, Higgins JPT. Individual patient data meta-analysis of time-to-event outcomes: one-stage versus two-stage approaches for estimating the hazard ratio under a random-effects model. Res Synth Methods 2011;2(3):150-62.
14. Burzykowski T, Molenberghs G, Buyse M, Geys H, Renard D. Validation of Surrogate End Points in Multiple Randomized Clinical Trials with Failure Time End Points. J R Stat Soc Ser C Appl Stat 2001;50(4):405-22.
15. Buyse M, Piedbois P. On the relationship between response to treatment and survival time. Stat Med 1996;15(24):2797-812.
16. Caubet J-F, Tosteson TD, Dong EW, Naylon EM, Whiting GW, Ernstoff MS, et al. Maximum androgen blockade in advanced prostate cancer: A meta-analysis of published randomized controlled trials using nonsteroidal antiandrogens. Urology 1997;49(1):71-8.
17. Clarke M, Stewart L, Pignon JP, Bijnens L. Individual patient data meta-analysis in cancer. Br J Cancer 1998;77(11):2036-44.
18. Collette L, Duchateau L, Suciu S, Bijnens L, Sylvester R. Including literature data in individual patient data meta-analyses for time-to-event endpoints. European Organization for Research and Treatment of Cancer (EORTC) Data Center. Draft paper; 1998. p. 1-48.

19. Combescure C, Foucher Y, Jackson D. Meta-analysis of single-arm survival studies: a distribution-free approach for estimating summary survival curves with random-effects. Stat Med 2014;33(15):2521-37.
20. Cooper N, Abrams K, Sutton A. Methodology for meta-analysis of mean time to event outcomes in the presence of censoring [abstract]. In: XI Cochrane Colloquium: Evidence, Health Care and Culture; 2003 Oct 26-31; Barcelona, Spain, 2003: p. 5.
21. Cope S, Jansen JP. Quantitative summaries of treatment effect estimates obtained with network meta-analysis of survival curves to inform decision-making. BMC Med Res Methodol 2013;13:147.
22. Coplen SE, Antman EM, Berlin JA, Hewitt P, Chalmers TC. Efficacy and safety of quinidine therapy for maintenance of sinus rhythm after cardioversion. A meta-analysis of randomized control trials [published erratum appears in Circulation 1991 Feb;83(2):714]. Circulation 1990;82(4):1106-16.
23. Crowther MJ, Look MP, Riley RD. Multilevel mixed effects parametric survival models using adaptive Gauss-Hermite quadrature with application to recurrent events and individual participant data meta-analysis. Stat Med 2014;33(22):3844-58.
24. Crowther MJ, Riley RD, Staessen JA, Wang J, Gueyffier F, Lambert PC. Individual patient data meta-analysis of survival data using Poisson regression models. BMC Med Res Methodol 2012;12:34.
25. D'Amico R, Deeks J, Altman D. Effect of length of follow-up on meta-analysis [abstract]. In: Sixth International Cochrane Colloquium; 1998 Oct 22-26; Baltimore, MD, USA, 1998: p. 51.
26. Daniels MJ, Hughes MD. Meta-analysis for the evaluation of potential surrogate markers. Stat Med 1997;16(17):1965-82.
27. Dear KB. Iterative generalized least squares for meta-analysis of survival data at multiple times. Biometrics 1994;50(4):989-1002.
28. Duchateau L, Collette L, Suciu S, Sylvester R. The impact of including survival information from literature on the overall x2 in an IPD meta-analysis. Second Symposium on Systematic Reviews: Beyond the Basics; 1999 Jan 5-7; Oxford, UK [serial online] 1999
29. Duchateau L, Collette L, Sylvester R, Pignon JP. Estimating number of events from the Kaplan-Meier curve for incorporation in a literature-based meta-analysis: what you don't see you can't get! Biometrics 2000;56(3):886-92.
30. Duchateau L, Pignon JP, Bijnens L, Bertin S, Bourhis J, Sylvester R. Individual patient-versus literature-based meta-analysis of survival data: time to event and event rate at a particular time can make a difference, an example based on head and neck cancer. Control Clin Trials 2001;22(5):538-47.
31. Duchateau L, Pignon JP, Collette L, Sylvester R. Incorporating number of events as estimated from the Kaplan Meier curves in a literature based meta-analysis: When the going gets tough, should we get going? [abstract]. 7th Annual Cochrane Colloquium; 1999 Oct 5-9; Rome, Italy [serial online] 1999:8
32. Earle CC, Pham B, Wells GA. An assessment of methods to combine published survival curves. Medical Decis Making 2000;20(1):104-11.
33. Early Breast Cancer Trialists' Collaborative Group. Statistical methods: Treatment of early breast cancer. Oxford: Oxford University Press; 1990.
34. Eggleston B, Nolan T, Wallace D, Pappas P. Comparing time-to-event analysis methodologies in a study of the treatment of acquired immune deficiency syndrome associated cryptococcal meningitis [abstract]. Clinical Trials [serial online] 2008 (4):405-6
35. Fine HA, Dear KBG, Loeffler JS, Mc Black PL, Canellos GP. Meta-analysis of radiation therapy with and without adjuvant chemotherapy for malignant gliomas in adults. Cancer 1993;71(8):2585-97.
36. Fiocco M, Putter H, Van Houwelingen JC. A new serially correlated gamma-frailty process for longitudinal count data. Biostatistics 2009;10(2):245-57.

37. Fiocco M, Putter H, van Houwelingen JC. Meta-analysis of pairs of survival curves under heterogeneity: a Poisson correlated gamma-frailty approach. Stat Med 2009;28(30):3782-97.
38. Fiocco M, Stijnen T, Putter H. Meta-analysis of time-to-event outcomes using a hazard-based approach: Comparison with other models, robustness and meta-regression. Comput Stat Data Anal 2012;56(5):1028-37.
39. Fisher DJ, Copas AJ, Tierney JF, Parmar MK. A critical review of methods for the assessment of patient-level interactions in individual participant data meta-analysis of randomized trials, and guidance for practitioners. J Clin Epidemiol 2011;64(9):949-67.
40. Floriani I, D'Amico R, Telaro E, Tinazzi A, Torri V. Evaluation of the summary statistics used for meta-analyses using literature survival data in oncology [abstract]. In: In: Second International Conference Scientific Basis of Health Services & Fifth Annual Cochrane Colloquium, 1997; Oct 8-12; Amsterdam, The Netherlands: p. 97.
41. Franklin J, Pluetschow A, Paus M. Measures of the quality of individual patient survival data [abstract]. XI Cochrane Colloquium: Evidence, Health Care and Culture; 2003 Oct 26-31; Barcelona, Spain 2003:9.
42. Ghosh D, Taylor JMG, Sargent DJ. Meta-analysis for Surrogacy: Accelerated Failure Time Models and Semicompeting Risks Modeling. Biometrics 2012;68(1):226-32.
43. Glenny A, Furness S, Worthington H. The calculation of hazard ratios published in reports of randomised controlled trials; are they valid? Poster presentation. In: 19th Cochrane Colloquium, 2011 Oct 19-22; Madrid, Spain.
44. Gustafson P. A Bayesian analysis of bivariate survival data from a multicentre cancer clinical trial. Stat Med 1995;14(23):2523-35.
45. Guyot P, Ades AE, Ouwens MJNM, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol 2012;12:9-.
46. Guyot P, Welton NJ, Ouwens MJNM, Ades AE. Survival time outcomes in randomized, controlled trials and meta-analyses: the parallel universes of efficacy and cost-effectiveness. Value Health 2011;14(5):640-6.
47. Haines TP, Hill AM. Inconsistent results in meta-analyses for the prevention of falls are found between study-level data and patient-level data. J Clin Epidemiol 2011(2):154-62.
48. Hirooka T, Hamada C, Yoshimura I. A note on estimating treatment effect for time-to-event data in a literature-based meta-analysis. Methods Inf Med 2009;48(2):104-12.
49. Hoogendoorn M, Hoogenveen RT, Rutten-van Molken MP, Vestbo J, Feenstra TL. Case fatality of COPD exacerbations: a meta-analysis and statistical modelling approach. Eur Respir J 2011;37(3):508-15.
50. Hozo I, Djulbegovic B. Individual-patient data meta-analysis (IPD MA) in the presence of competing risks [abstract]. In: XIII Cochrane Colloquium; 2005 Oct 22-26; Melbourne, Australia, 2005.
51. Hua H, Burke DL, Crowther MJ, Ensor J, Tudur Smith C, Riley RD. One-stage individual participant data meta-analysis models: estimation of treatment-covariate interactions must avoid ecological bias by separating out within-trial and across-trial information. Stat Med 2017;36(5):772-89
52. Hunink MG, Wong JB. Meta-analysis of failure-time data with adjustment for covariates. Medical Decis Making 1994;14(1):59-70.
53. Imamura M, MacLennan S, Dahm P, Neuberger M, Omar M, MacLennan G. The hazards of extracting hazard ratios from survival curves for meta-analyses in observational studies. Poster presentation. In: 19th Cochrane Colloquium, 2011 Oct 19-22; Madrid, Spain.
54. Jackson D, Rollins K, Coughlin P. A multivariate model for the meta-analysis of study level survival data at multiple times. Res Synth Methods 2014;5(3):264-72.
55. Jansen JP, Cope S. Meta-regression models to address heterogeneity and inconsistency in network meta-analysis of survival outcomes. BMC Med Res Methodol 2012;12:152.

56. Jansen JP. Network meta-analysis of survival data with fractional polynomials. BMC Med Res Methodol 2011;11:61.
57. Katsahian S, Latouche A, Mary J-Y, Chevret S, Porcher R. Practical methodology of meta-analysis of individual patient data using a survival outcome. Contemp Clin Trials 2008;29(2):220-30.
58. Keene ON. Alternatives to the hazard ratio in summarizing efficacy in time-to-event studies: an example from influenza trials. Stat Med 2002;21(23):3687-700.
59. Koller MT, Stijnen T, Steyerberg EW, Lubsen J. Meta-analyses of chronic disease trials with competing causes of death may yield biased odds ratios. J Clin Epidemiol 2008;61(4):365-72.
60. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemother Rep 1966;50(3):163.
61. Massonnet G, Janssen P, Burzykowski T. Fitting conditional survival models to meta-analytic data by using a transformation toward mixed-effects models. Biometrics 2008;64(3):834-42.
62. Messori A, Rampazzo R. Meta-analysis of clinical trials based on censored end-points: simplified theory and implementation of the statistical algorithms on a microcomputer. Comput Methods Programs Biomed 1993(4):261-7.
63. Messori A, Trippoli S, Vaiani M, Cattel F. Survival meta-analysis of individual patient data and survival meta-analysis of published (aggregate) data: Is there an intermediate approach between these two opposite options? Clin Drug Investig 2000(5):309-16.
64. Michiels S, Baujat B, Mahe C, Sargent DJ, Pignon JP. Random-effects survival models gave a better understanding of heterogeneity in individual patient data meta-analyses. J Clin Epidemiol 2005;58(3):238-45.
65. Michiels S, Piedbois P, Burdett S, Syz N, Stewart L, Pignon JP. Meta-analysis when only the median survival times are known: a comparison with individual patient data results. Int J Technol Assess Health Care 2005(1):119-25.
66. Michiels S, Syz N, Piedbois P, Burdett S, Stewart L, Pignon JP. Is it possible to perform a meta-analysis when only the median survival time is known? A comparison with individual patient data results [abstract]. 10th Cochrane Colloquium; 2002 Jul 31- Aug 3; Stavanger, Norway. [serial online] 2002:30-1.
67. Monti GE, Frankena K. Survival analysis on aggregate data to assess time to sero-conversion after experimental infection with Bovine Leukemia virus. Prev Vet Med 2005;68(2-4):241-62.
68. Moodie PF, Nelson NA, Koch GG. A non-parametric procedure for evaluating treatment effect in the meta-analysis of survival data. Stat Med 2004(7):1075-93.
69. O'Quigley J, Stare J. Proportional hazards models with frailties and random-effects. Stat Med 2002;21(21):3219-33.
70. Otwombe KN, Petzold M, Martinson N, Chirwa T. A review of the study designs and statistical methods used in the determination of predictors of all-cause mortality in HIV-infected cohorts: 2002-2011. PLoS One 2014;9(2):e87356.
71. Ouwens MJNM, Philips Z, Jansen JP. Network meta-analysis of parametric survival curves. Res Synth Methods 2010;1(3-4):258-71.
72. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med 1998;17(24):2815-34.
73. Pennells L, Kaptoge S, White IR, Thompson SG, Wood AM, Emerging Risk Factors C. Assessing risk prediction models using individual participant data from multiple studies. Am J Epidemiol 2014;179(5):621-32.
74. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. Br J Cancer 1977;35(1):1-39.
75. Pham B, Earle C. Comparison of four methods for metaanalysis of survival data [abstract]. Sixth International Cochrane Colloquium; 1998 Oct 22-26; Baltimore, MD, USA [serial online] 1998:62

76. Renfro LA, Shi Q, Sargent DJ, Carlin BP. Bayesian adjusted R2 for the meta-analytic evaluation of surrogate time-to-event endpoints in clinical trials. Stat Med 2012;31(8):743-61.
77. Riley RD, Abrams KR, Sutton AJ, Lambert PC, Jones DR, Heney D, et al. Reporting of prognostic markers: current problems and development of guidelines for evidence-based practice in the future. Br J Cancer 2003;88(8):1191-8.
78. Rondeau V, Michiels S, Liquet B, Pignon JP. Investigating trial and treatment heterogeneity in an individual patient data meta-analysis of survival data by means of the penalized maximum likelihood approach. Stat Med 2008;27(11):1894-910.
79. Rondeau V, Pignon J-P, Michiels S. A joint model for the dependence between clustered times to tumour progression and deaths: A meta-analysis of chemotherapy in head and neck cancer. Stat Methods Med Res 2011.
80. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Stat Med 2002;21(15):2175-97.
81. Royston P, Parmar MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol 2013;13:152.
82. Saramago P, Chuang LH, Soares MO. Network meta-analysis of (individual patient) time to event data alongside (aggregate) count data. BMC Med Res Methodol 2014;14:105.
83. Saramago P, Sutton AJ, Cooper NJ, Manca A. Mixed treatment comparisons using aggregate and individual participant level data. Stat Med 2012;31(28):3516-36.
84. Sargent DJ, Patiyil S, Yothers G, Haller DG, Gray R, Benedetti J, et al. End points for colon cancer adjuvant trials: observations and recommendations based on individual patient data from 20,898 patients enrolled onto 18 randomized trials from the ACCENT Group. J Clin Oncol 2007(29):4569-74.
85. Sargent DJ, Zee B, Milan C, Torri V, G. F. Meta-analysis of individual patient survival data using random-effects models. In: Stangl D, Berry DA. Meta-analysis in medicine and health policy. New York: Marcel Dekker; 2000. p. 255-75.
86. Sargent DJ. A General Framework for Random-effects Survival Analysis in the Cox Proportional Hazards Setting. Biometrics 1998;54(4):1486-97.
87. Shore T, Weinerman B, Nelson N. A meta-analysis of stages I and II Hodgkin's disease. Cancer 1990;65(5):1155-60.
88. Siannis F, Barrett JK, Farewell VT, Tierney JF. One-stage parametric meta-analysis of time-to-event outcomes. Stat Med 2010;29(29):3030-45.
89. Simes RJ. Confronting publication bias: a cohort design for meta analysis. Stat Med 1987(1):11-29.
90. Simmonds MC, Higgins JP, Stewart LA, Tierney JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. Clin Trials 2005;2(3):209-17.
91. Simmonds MC, Tierney J, Bowden J, Higgins JPT. Meta-analysis of time-to-event data: a comparison of two-stage methods. Research Synthesis Methods 2011;2(3):139-49.
92. Simmonds MC, Higgins JP, Stewart LA. Random-effects meta-analysis of time-to-event data using the expectation-maximisation algorithm and shrinkage estimators. Res Synth Methods 2013;4(2):144-55.
93. Srinivasan C, Zhou M. A Note on Pooling Kaplan-Meier Estimators. Biometrics 1993;49(3):861-4.
94. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. Methods for Meta-Analysis in Medical Research. Chichester: John Wiley & Sons Ltd.; 2000.
95. Thompson S, Kaptoge S, White I, Wood A, Perry P, Danesh J, et al. Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies. Int J Epidemiol 2010;39(5):1345-59.

96. Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. Trials 2007(1):16.
97. Tudur C, Williamson PR, Khan S, Best LY. The value of the aggregate data approach in meta-analysis with time-to-event outcomes. J R Stat Soc Ser A 2001;164(2):357-70.
98. Tudur Smith C, Williamson PR, Marson AG. An overview of methods and empirical comparison of aggregate data and individual patient data results for investigating heterogeneity in meta-analysis of time-to-event outcomes. J Eval Clin Pract 2005;11(5):468-78.
99. Tudur Smith C, Williamson PR, Marson AG. Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. Stat Med 2005;24(9):1307-19.
100. Tudur Smith C, Williamson PR. A comparison of methods for fixed-effects meta-analysis of individual patient data with time to event outcomes. Clin Trials 2007;4(6):621-30.
101. Vaida F, Xu R. Proportional hazards model with random-effects. Stat Med 2000;19(24):3309-24.
102. Vale C, Burdett S, Tierney J, Stewart L. Does adjusting for censoring improve the reliability of estimates in meta-analyses of survival outcomes? [abstract]. 9th Annual Cochrane Colloquium; 2001Oct 9-13; Lyon, France. [serial online] 2001:12-3
103. Vale CL, Tierney JF, Stewart LA. Effects of adjusting for censoring on meta-analyses of time-to-event outcomes. Int J Epidemiol 2002(1):107-11.
104. Voest EE, van Houwelingen JC, Neijt JP. A meta-analysis of prognostic factors in advanced ovarian cancer with median survival and overall survival (measured with the log (relative risk)) as main objectives. Eur J Cancer Clin Oncol 1989;25(4):711-20.
105. Wei Y, Royston P, Tierney JF, Parmar MK. Meta-analysis of time-to-event outcomes from randomized trials using restricted mean survival time: application to individual participant data. Stat Med 2015;34(21):2881-98.
106. Wells GA, Earle CC, Pham B, Yetisir E, Coyle D. Assessment of methodologies to combine survival curves in meta-analysis: applications to oncology [abstract]. 15th Annual Meeting of the International Society of Technology Assessment in Health Care; 1999 Jun 20-23; Edinburgh, UK 1999:68.
107. Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. Stat Med 1991;10(11):1665-77.
108. Williamson PR, Clough HE, Hutton JL, Marson AG, Chadwick DW. Statistical issues in the assessment of the evidence for an interaction between factors in epilepsy trials. Stat Med 2002;21(18):2613-22.
109. Williamson PR, Marson AG, Tudur C, Hutton JL, Chadwick D. Individual patient data meta-analysis of randomized antiepileptic drug monotherapy trials. Journal of Evaluation in Clinical Practice 2000(2):205-14.
110. Williamson PR, Smith CT, Hutton JL, Marson AG. Aggregate data meta-analysis with time-to-event outcomes. Stat Med 2002;21(22):3337-51.
111. Woods BS, Hawkins N, Scott DA. Network meta-analysis on the log-hazard scale, combining count and hazard ratio statistics accounting for multi-arm trials: a tutorial. BMC Med Res Methodol 2010;10:54.
112. Yamaguchi T, Ohashi Y, Matsuyama Y. Proportional hazards models with random-effects to examine centre effects in multicentre cancer clinical trials. Stat Methods Med Res 2002;11(3):221-36.
113. Yuan X, Anderson SJ. Meta-analysis methodology for combining treatment effects from Cox proportional hazard models with different covariate adjustments. Biom J 2010;52(4):519-37.
114. Zang J, Xu J, Xiang C, He J, Zou S. Multi-level model synthesis of median survival time in meta-analysis. Epidemiology 2015;26(1):e2-3.
115. Zavala D, Carvajal AM. Software for data transformation for time-to-event outcomes [abstract]. In: XV Cochrane Colloquium; 2007 Oct 23-27; Sao Paulo, Brazil, 2007: p. 154.

Appendix 2: Search strategies used in Chapter 3

Search Strategy of Cochrane Epilepsy Group's Specialised Register (search conducted 14/09/2012)

#1 (Carbamazepine or Tegretol or Carbagen or CBZ or Clobazam or Frisium or Urbanol or Onfi or CLB or Eslicarbazepine or Zebinix or Exalief or Stedesa or ESL or Ethosuximide or Emeside or Zarontin or ESM or Gabapentin or Fanatrex or Gabarone or Neogab or Gralise or Neurontin or Nupentin or GBP or Lamotrigine or Lamict* or Lamotrine or Lamitrin or Lamogine or Lamitor or LTG or Levetiracetam or Keppra or LEV or Oxcarbazepine or Trileptal or OXC or Phenobarbit* or Luminal or PB or Phenytoin or Epanutin or Phenytek or Dilantin or Eptoin or Diphenin* or Diphenylhydantoin or PHT or Pregabalin or Lyrica or PGB or Primidone or Mysoline or Prysoline or Liskantin or Desitin or Resimatil or Mylepsinum or Sertan or PRM or Remacemide or Ecovia or RMC or Sultiam* or Sulthiam* or Ospolot or STM or Tiagabine or Gabitril or TGB or Topiramate or Topamax or TPM or Valpro* or Convulex or Depak* or Depacon or Valparin or Stavzor or Epilim or Epiject or Episenta or Epival or Orlept or Orfiril or Selenica or VPA or Vigabatrin or Sabril or VGB or Zonisamide or Zonegran or Excegran or ZNS) AND (INREGISTER) [REFERENCE] [STANDARD]

#2 ((adjunct* or "add-on" or "add on") NOT monotherap*) AND (INREGISTER) [REFERENCE] [STANDARD]

#3 (#1 NOT #2) AND (INREGISTER) [REFERENCE] [STANDARD]

Appendix 3: Data extraction form used in Chapter 3

Reporting of Time-to-Event outcomes and analyses in Epilepsy Monotherapy Studies

Data Extraction Form

STAGE 1 – OUTCOMES

Name of reviewer: Laura ☐ Sarah ☐

Date of extraction: ___ / ___ / _____

Study Authors:

Study Title:

Journal:

Year of Study:

Study ID (first author and year):

List all outcomes reported in the study:

Is at least one time-to-event outcome reported? Yes ☐ No ☐

Unclear ☐ If Unclear, why?

IF YES OR UNCLEAR, CONTINUE TO STAGE 2: TIME-TO-EVENT STUDIES

IF NO, EXCLUDE STUDY FROM STAGE 2 (DATA EXTRACTION IS COMPLETE)

STAGE 2- TIME TO EVENT OUTCOMES (STUDIES)

Number of treatment arms:

List of treatment arms (including doses for dose-controlled studies):

Did the study receive pharmaceutical sponsorship/support or have pharmaceutical involvement (e.g. authors affiliated with pharmaceutical companies)?

Yes ☐ No ☐ Unclear ☐

If yes, provide details of pharmaceutical involvement:

Type of control:	Placebo or no treatment	☐	Other AED	☐
	Dose-controlled (same AED)	☐		
	Other active control (specify)	☐		
Study sponsorship:	Pharmaceutical	☐	Academic	☐
	Government	☐	Other (specify)	☐
	Unclear	☐	Details (if unclear):	
Blinding	Single-blind	☐	Double-blind	☐
	Open-Label / Unblinded	☐	Other blinding	☐
	Unclear	☐		

If study was blinded, who was blinded? If open label, was a reason for no blinding given?

Design of study: Superiority* ☐ Equivalence ☐

Non-inferiority ☐ Unclear ☐

*Superiority design for Epilepsy monotherapy studies as defined by ILAE 2006 guidelines

Reference: Glauser T, Ben Menachem E, Bourgeois B, et al. ILAE Treatment Guidelines: Evidence-based Analysis of Antiepileptic Drug Efficacy and Effectiveness as Initial Monotherapy for Epileptic Seizures and Syndromes. Epilepsia 2006;**47**(7):1094-120)

Details:

Population (tick all that apply):

Refractory/ drug resistant seizures ☐ Newly diagnosed seizures ☐

Relapsed seizures (after AED discontinuation) ☐ Unclear ☐

Other (specify): ☐

General information: are the following clearly reported?

Number of participants randomised: Yes ☐ No ☐

Number of participants completing study: Yes ☐ No ☐

Number of participants excluded from the study / from analysis:

Yes ☐ No ☐

Details of any reasons for exclusions:

Are all above participant numbers specified by treatment arms?

Yes ☐ No ☐

Details of anything unclear or not reported:

Time frame of the study: are the following clearly reported?

Length of accrual period of participants (specify):

Length of titration / baseline / dose-escalation period (specify):

Length of maintenance period (specify):

Study Duration (specify):

Other study period: (e.g. pre-treatment period, open-label extension period etc.):

Frequency / number of follow-up visits (e.g. monthly visits, 4 visits per participant etc.):

Minimum follow-up time:	Yes	☐	No	☐
Maximum follow-up time:	Yes	☐	No	☐
Mean follow-up time:	Yes	☐	No	☐
Median follow-up time:	Yes	☐	No	☐

If yes, is the method used to calculate median follow-up time reported (specify):

Details of anything unclear or not reported in time frame of the study:

All primary / secondary outcomes: are the following clearly reported?

Number of outcomes (specify):

Number of time-to-event outcomes (specify):

Are all reported outcomes clearly defined? Yes ☐ No ☐

If No specify details:

Is a single primary outcome identified? Yes ☐ No ☐

If yes, specify this primary outcome (and whether the outcome is time-to-event, binary, continuous etc.):

If yes, is a sample size calculation described relating to the primary outcome?

Yes ☐ No ☐

Is time-to-withdrawal of allocated treatment (retention time) reported?

Yes ☐ No ☐

If yes, are competing risks taken into account (e.g. withdrawal due to poor seizure control and/or adverse events) Yes ☐ No ☐

If yes, provide details of how competing risks are taken into account:

Is at least one seizure outcome (efficacy) defined? Yes ☐ No ☐

If yes specify (and whether the outcome is time-to-event, binary, continuous etc.):

Is at least one adverse event outcome (tolerability) defined?

Yes ☐ No ☐

If yes specify (and whether the outcome is time-to-event, binary, continuous etc.):

Other outcomes reported (e.g. cognitive, quality of life outcomes):

Are outcomes reported consistently in both methods and results sections?

Yes ☐ No ☐

Details of anything unclear in outcome reporting:

STAGE 3: TIME-TO-EVENT OUTCOME REPORTING

Repeat STAGE 3 for each time-to-event outcome reported

Number of time to event outcomes in the study:

Time-to-event outcome (exact definition):

Is the time origin of the outcome defined?

Yes ☐ No ☐ Unclear ☐

Specify details (e.g. date of randomisation, date of first treatment dose etc.):

Is the number of participants contributing to the outcome reported?

Yes ☐ No ☐

Is the number of events for the outcome reported? Yes ☐ No ☐

Is the definition of an event clear for the outcome? Yes ☐ No ☐

If yes, specify definition of event:

Number of participants censored: Yes ☐ No ☐

Number of participants lost to follow-up: Yes ☐ No ☐

Is the definition of a censored observation clear for the outcome?

Yes ☐ No ☐

If yes, specify definition of censored observations:

Are participants censored and participants lost to follow-up reported separately?

Yes ☐ No ☐

Are events, censoring and losses to follow-up reported by treatment group?

Yes ☐ No ☐

Statistical analysis: Are the following clearly reported?

Survival probability / event rate: Yes ☐ No ☐

Median Survival time: Yes ☐ No ☐

Log-rank p-value: Yes ☐ No ☐

Other p-value (specify):

Precision of p-value (e.g. 'not significant,' one decimal place etc.):

Hazard ratio: Yes ☐ No ☐

Other effect size (specify):

Standard deviation / standard error of effect size: Yes ☐ No ☐

Confidence interval of effect size: Yes ☐ No ☐

Observed and Expected number of events: Yes ☐ No ☐

Statistical tests / models / methods used (describe as much detail as possible reported):

Is it stated that all assumptions of statistical methods have been assessed?

Yes ☐ No ☐

Specify details:

Are any multivariable methods /analyses used? Yes ☐ No ☐

If yes specify:

Is a method of choosing variables described (specify)?

Are variables measured at baseline? Yes ☐ No ☐

Are multivariable model coefficients reported? Yes ☐ No ☐

Are multivariable model confidence intervals reported? Yes ☐ No ☐

Are multivariable model p-values reported? Yes ☐ No ☐

Has a statistical software package been used for analysis (specify including version)?

Have any subgroup or sensitivity analyses been performed (specify)?

Are methods reported in sufficient detail to replicate results?

Yes ☐ No ☐

Are reported methods and results sections consistent?

Yes ☐ No ☐

Details of problems or inconsistencies in statistical analyses of outcomes:

Survival plots

Is a survival graph presented for the outcome? Yes ☐ No ☐

Which type of graph is presented? Kaplan-Meier ☐ Actuarial ☐

Other (specify) ☐ Unclear ☐

Is a step function used for Kaplan-Meier plots? Yes ☐ No ☐

Is the direction of the graph up or down? Up ☐ Down ☐

Are censored observations clearly marked? Yes ☐ No ☐

Are (effective) numbers at risk reported at regular intervals?

Yes ☐ No ☐

Are different line types clearly used for multiple curves?

Yes ☐ No ☐

Is a clear legend provided for the graph? Yes ☐ No ☐

Are the axes of the graph appropriate (e.g. intervals, scale etc.)?

Yes ☐ No ☐

Is an effect size (e.g. p-value/hazard ratio) displayed on the graph?

Yes ☐ No ☐

Are standard errors/confidence intervals displayed on the graph?

Yes ☐ No ☐

Details of any problems or inconsistencies with graphs:

STAGE 4: OVERALL REPORTING

Are published survival graphs consistent with published text?

Yes ☐ No ☐ No graph(s) ☐

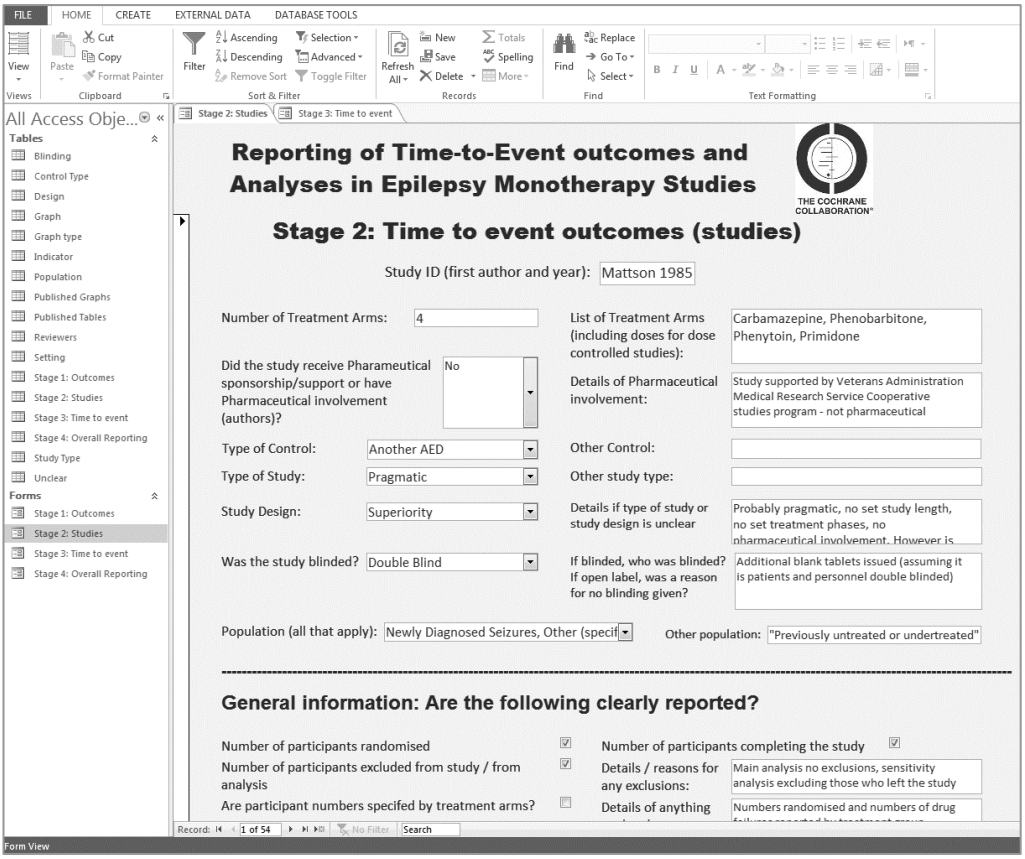
Are published tables consistent with published text?

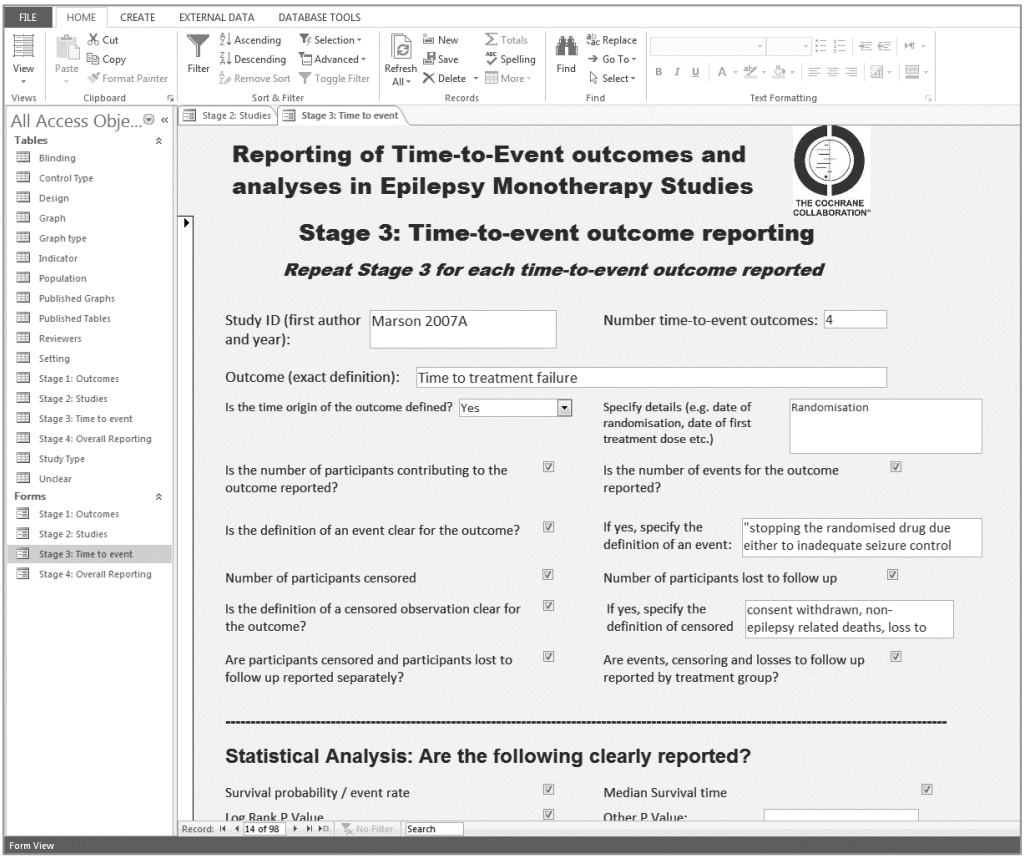
Yes ☐ No ☐ No tables(s) ☐

Details of any problems with consistency of methods, results, tables, graphs, conclusions etc:

Any other issues with any aspect of reporting of outcomes or analyses:

Appendix 4: Screenshot of Microsoft Access Database used in Chapter 3





Appendix 5: References of 108 epilepsy monotherapy trials included in Chapter 3

1. Aikiä M, Kälviäinen R, Sivenius J, et al. Cognitive effects of oxcarbazepine and phenytoin monotherapy in newly diagnosed epilepsy: one year follow-up. Epilepsy Res 1992; 11(3):199-203.
2. Arroyo S, Dodson WE, Privitera MD, et al. Randomized dose-controlled study of topiramate as first-line therapy in epilepsy. Acta Neurol Scand 2005;112(4):214-22.
3. Banu SH, Jahan M, Koli UK, et al. Side effects of phenobarbital and carbamazepine in childhood epilepsy: randomised controlled trial. BMJ (Clinical Res ed) 2007;334(7605):1207.
4. Bast T, Völp A, Wolf C, et al. The influence of sulthiame on EEG in children with benign childhood epilepsy with centrotemporal spikes (BECTS). Epilepsia 2003;44(2):215-20.
5. Baulac M, Brodie MJ, Patten A, et al. Efficacy and tolerability of zonisamide versus controlled-release carbamazepine for newly diagnosed partial epilepsy: a phase 3, randomised, double-blind, non-inferiority trial. Lancet Neurol 2012;11(7):579-88.
6. Bawden HN, Camfield CS, Camfield PR, et al. The cognitive and behavioural effects of clobazam and standard monotherapy are comparable. Canadian Study Group for Childhood Epilepsy. Epilepsy Res 1999;33(2-3):133-43.
7. Beydoun A, Sachdeo RC, Rosenfeld WE, et al. Oxcarbazepine monotherapy for partial-onset seizures: a multicenter, double-blind, clinical trial. Neurology 2000;54(12):2245-51.
8. Beydoun A, Sackellares JC, Shu V. Safety and efficacy of divalproex sodium monotherapy in partial epilepsy: a double-blind, concentration-response design clinical trial. Depakote Monotherapy for Partial Seizures Study Group. Neurology 1997;48(1):182-88.
9. Bill PA, Vigonius U, Pohlmann H, et al. A double-blind controlled clinical trial of oxcarbazepine versus phenytoin in adults with previously untreated epilepsy. Epilepsy Res 1997;27(3):195-204.
10. Biton V, Mirza W, Montouris G, et al. Weight change associated with valproate and lamotrigine monotherapy in patients with epilepsy. Neurology 2001;56(2):172-77.
11. Bittencourt PR, Antoniuk SA, Bigarella MM, et al. Carbamazepine and phenytoin in epilepsies refractory to barbiturates: efficacy, toxicity and mental function [published erratum appears in Epilepsy Res 1994 Mar;17(3):271]. Epilepsy Res 1993;16(2):147-55.
12. Brodie MJ, Chadwick DW, Anhut H, et al. Gabapentin versus lamotrigine monotherapy: a double-blind comparison in newly diagnosed epilepsy. Epilepsia 2002;43(9):993-1000.
13. Brodie MJ, Overstall PW, Giorgi L. Multicentre, double-blind, randomised comparison between lamotrigine and carbamazepine in elderly patients with newly diagnosed epilepsy. The UK Lamotrigine Elderly Study Group. Epilepsy Res 1999;37(1):81-87.
14. Brodie MJ, Perucca E, Ryvlin P, et al. Comparison of levetiracetam and controlled-release carbamazepine in newly diagnosed epilepsy. Neurology 2007;68(6):402-08.
15. Brodie MJ, Richens A, Yuen AWC, et al. Double-blind comparison of lamotrigine and carbamazepine in newly diagnosed epilepsy. Lancet 1995;345(8948):476-79.
16. Brodie MJ, Wroe SJ, Dean ADP, et al. Efficacy and safety of remacemide versus carbamazepine in newly diagnosed epilepsy: comparison by sequential analysis. Epilepsy Behav 2002;3(2):140-46.
17. Callaghan N, Kenny RA, O'Neill B, et al. A prospective study between carbamazepine, phenytoin and sodium valproate as monotherapy in previously untreated and recently diagnosed patients with epilepsy. J Neurol Neurosurg Psychiatry 1985;48(7):639-44.
18. Callaghan N, O'Hare J, O'Driscoll D, et al. Comparative study of ethosuximide and sodium valproate in the treatment of typical absence seizures (petit mal). Dev Med Child Neurol 1982;24(6):830-36.

19. Canadian Study Group for Childhood E. Clobazam has equivalent efficacy to carbamazepine and phenytoin as monotherapy for childhood epilepsy. Epilepsia 1998;39(9):952-59.
20. Chadwick D. Safety and efficacy of vigabatrin and carbamazepine in newly diagnosed epilepsy: a multicentre randomised double-blind study. Vigabatrin European Monotherapy Study Group. Lancet 1999;354(9172):13-19.
21. Chadwick DW, Anhut H, Greiner MJ, et al. A double-blind trial of gabapentin monotherapy for newly diagnosed partial seizures. International Gabapentin Monotherapy Study Group 945-77. Neurology 1998;51(5):1282-88.
22. Chen YJ, Kang WM, So WC. Comparison of antiepileptic drugs on cognitive function in newly diagnosed epileptic children: a psychometric and neurophysiological study. Epilepsia 1996;37(1):81-86.
23. Cho YW, Kim DH, Motamedi GK. The effect of levetiracetam monotherapy on subjective sleep quality and objective sleep parameters in patients with epilepsy: Compared with the effect of carbamazepine-CR monotherapy. Seizure 2011;20(4):336-39.
24. Christe W, Krämer G, Vigonius U, et al. A double-blind controlled clinical trial: oxcarbazepine versus sodium valproate in adults with newly diagnosed epilepsy. Epilepsy Res 1997;26(3):451-60.
25. Coppola G, Auricchio G, Federico R, et al. Lamotrigine versus valproic acid as first-line monotherapy in newly diagnosed typical absence seizures: an open-label, randomized, parallel-group study. Epilepsia 2004;45(9):1049-53.
26. Coppola G, Franzoni E, Verrotti A, et al. Levetiracetam or oxcarbazepine as monotherapy in newly diagnosed benign epilepsy of childhood with centrotemporal spikes (BECTS): an open-label, parallel group trial. Brain Dev 2007;29(5):281-84.
27. Craig I, Tallis R. Impact of valproate and phenytoin on cognitive function in elderly patients: results of a single-blind randomized comparative study. Epilepsia 1994;35(2):381-90.
28. Dam M, Ekberg R, Loyning Y, et al. A double-blind study comparing oxcarbazepine and carbamazepine in patients with newly diagnosed, previously untreated epilepsy. Epilepsy Res 1989;3(1):70-76.
29. de Silva M, MacArdle B, McGowan M, et al. Randomised comparative monotherapy trial of phenobarbitone, phenytoin, carbamazepine, or sodium valproate for newly diagnosed childhood epilepsy. Lancet 1996;347(9003):709-13.
30. Dodrill CB, Arnett JL, Shu V, et al. Effects of tiagabine monotherapy on abilities, adjustment, and mood. Epilepsia 1998;39(1):33-42.
31. Donati F, Gobbi G, Campistol J, et al. The cognitive effects of oxcarbazepine versus carbamazepine or valproate in newly diagnosed children with partial seizures. Seizure 2007;16(8):670-79.
32. Easter D, O'Bryan-Tear CG, Verity C. Weight gain with valproate or carbamazepine--a reappraisal. Seizure 1997;6(2):121-25.
33. Edwards KR, Sackellares JC, Vuong A, et al. Lamotrigine monotherapy improves depressive symptoms in epilepsy: a double-blind comparison with valproate. Epilepsy Behav 2001;2(1):28-36.
34. Eun SH, Kim HD, Eun BL, et al. Comparative trial of low- and high-dose zonisamide as monotherapy for childhood epilepsy. Seizure 2011;20(7):558-63.
35. Eun S-H, Eun B-L, Lee JS, et al. Effects of lamotrigine on cognition and behavior compared to carbamazepine as monotherapy for children with partial epilepsy. Brain Dev 2012;34(10):818-23.
36. Fattore C, Boniver C, Capovilla G, et al. A multicenter, randomized, placebo-controlled trial of levetiracetam in children and adolescents with newly diagnosed absence epilepsy. Epilepsia 2011;52(4):802-09.

37. Faught E, Sachdeo RC, Remler MP, et al. Felbamate monotherapy for partial-onset seizures: an active control trial. Neurology 1993;43(4):688-92.
38. Feksi AT, Kaamugisha J, Sander JW, et al. Comprehensive primary health care antiepileptic drug treatment programme in rural and semi-urban Kenya. ICBERG (International Community-based Epilepsy Research Group). Lancet 991;337(16)(8738):406-09.
39. Forsythe I, Butler R, Berg I, et al. Cognitive impairment in new cases of epilepsy randomly assigned to carbamazepine, phenytoin and sodium valproate. Dev Med Child Neurol 1991;33(6):524-34.
40. Gilad R, Sadeh M, Rapoport A, et al. Monotherapy of lamotrigine versus carbamazepine in patients with post stroke seizure. Clin Neuropharmacol 2007;30(4):189-95.
41. Gillham R, Kane K, Bryant-Comstock L, et al. A double-blind comparison of lamotrigine and carbamazepine in newly diagnosed epilepsy with health-related quality of life as an outcome measure. Seizure 2000;9(6):375-79.
42. Gilliam F, Vazquez B, Sackellares JC, et al. An active-control trial of lamotrigine monotherapy for partial seizures [see comments]. Neurology 1998;51(4):1018-25.
43. Gilliam FG, Veloso F, Bomhof MA, et al. A dose-comparison trial of topiramate as monotherapy in recently diagnosed partial epilepsy. Neurology 2003;60(2):196-202.
44. Glauser TA, Cnaan A, Shinnar S, et al. Ethosuximide, valproic acid, and lamotrigine in childhood absence epilepsy. N Engl J Med 2010;362(9):790-99.
45. Glauser TA, Dlugos DJ, Dodson WE, et al. Topiramate monotherapy in newly diagnosed epilepsy in children and adolescents. J Child Neurol 2007;22(6):693-99.
46. Guerreiro MM, Vigonius U, Pohlmann H, et al. A double-blind controlled clinical trial of oxcarbazepine versus phenytoin in children and adolescents with epilepsy. Epilepsy Res 1997;27(3):205-13.
47. Hakami T, Todaro M, Petrovski S, Macgregor L, Velakoulis D, Tan M, et al. Substitution Monotherapy With Levetiracetam vs Older Antiepileptic Drugs: A Randomized Comparative Trial. Arch Neurol 2012;69(12):1563-71.
48. Heller AJ, Chesterman P, Elwes RD, et al. Phenobarbitone, phenytoin, carbamazepine, or sodium valproate for newly diagnosed adult epilepsy: a randomised comparative monotherapy trial. J Neurol Neurosurg Psychiatry 1995;58(1):44-50.
49. Homan RW, Miller B. Causes of treatment failure with antiepileptic drugs vary over time. Neurology 1987;37(10):1620-23.
50. Kaminow L, Schimschock JR, Hammer AE, et al. Lamotrigine monotherapy compared with carbamazepine, phenytoin, or valproate monotherapy in patients with epilepsy. Epilepsy Behav 2003;4(6):659-66.
51. Kang HC, Eun BL, Wu Lee C, et al. The effects on cognitive function and behavioural problems of topiramate compared to carbamazepine as monotherapy for children with benign rolandic epilepsy. Epilepsia 2007;48(9):1716-23.
52. Kaushal S, Rani A, Chopra SC, et al. Safety and efficacy of clobazam versus phenytoin-sodium in the antiepileptic drug treatment of solitary cysticercus granulomas. Neurol India 2006;54(2):157-60.
53. Ketter TA, Malow BA, Flamini R, et al. Felbamate monotherapy has stimulant-like effects in patients with epilepsy. Epilepsy Res 1996;23(2):129-37.
54. Kwan P, Brodie MJ, Kälviäinen R, et al. Efficacy and safety of pregabalin versus lamotrigine in patients with newly diagnosed partial seizures: a phase 3, double-blind, randomised, parallel-group trial. Lancet Neurol 2011;10(10):881-90.
55. Kwan P, Yip FP, Hui ACF, et al. Effects of valproate or lamotrigine monotherapy on the reproductive endocrine and insulin-related metabolic profile in Chinese adults with epilepsy: a prospective randomized study. Epilepsy Behav 2009;14(4):610-16.

56. Lee SA, Lee HW, Heo K, et al. Cognitive and behavioral effects of lamotrigine and carbamazepine monotherapy in patients with newly diagnosed or untreated partial epilepsy. Seizure 2011;20(1):49-54.
57. Levisohn PM, Holland KD. Topiramate or valproate in patients with juvenile myoclonic epilepsy: a randomized open-label comparison. Epilepsy Behav 2007;10(4):547-52.
58. Marson AG, Al-Kharusi AM, Alwaidh M, et al. The SANAD study of effectiveness of carbamazepine, gabapentin, lamotrigine, oxcarbazepine, or topiramate for treatment of partial epilepsy: an unblinded randomised controlled trial. Lancet 2007;369(9566):1000-15.
59. Marson AG, Al-Kharusi AM, Alwaidh M, et al. The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial. Lancet 2007;369(9566):1016-26.
60. Mattson RH, Cramer JA, Collins JF, et al. Comparison of carbamazepine, phenobarbital, phenytoin, and primidone in partial and secondarily generalized tonic-clonic seizures. N Engl J Med 1985;313(3):145-51.
61. Mattson RH, Cramer JA, Collins JF, et al. A comparison of valproate with carbamazepine for the treatment of complex partial seizures and secondarily generalized tonic clonic seizures in adults. N Engl J Med 1992;327(11):765-71.
62. Mikkelsen B, Berggreen P, Joensen P, et al. Clonazepam (Rivotril) and carbamazepine (Tegretol) in psychomotor epilepsy: a randomized multicenter trial. Epilepsia 1981;22(4):415-20.
63. Mitchell WG, Chavez JM. Carbamazepine versus phenobarbital for partial-onset seizures in children. Epilepsia 1987;28(1):56-60.
64. Mitsudome A, Ohfu M, Yasumoto S, et al. The effectiveness of clonazepam on the Rolandic discharges. Brain Dev 1997;19(4):274-78.
65. Nasreddine W, Beydoun A. Valproate-induced thrombocytopenia: a prospective monotherapy study. Epilepsia 2008;49(3):438-45.
66. Nejad SEM, Nikpour MRA, Rahim F, et al. A randomized open-label comparison of lamotrigine and valproate in patients with juvenile myoclonic epilepsy. Int J Pharm 2009;5(5):313-18.
67. Nieto-Barrera M, Brozmanova M, Capovilla G, et al. A comparison of monotherapy with lamotrigine or carbamazepine in patients with newly diagnosed partial epilepsy. Epilepsy Res 2001;46(2):145-55.
68. Ogunrin O, Adamolekun B, Ogunniyi A. Cognitive effects of antiepileptic drugs in Nigerians with epilepsy. Afr J Neuro Sci 2005;24(1):18-24.
69. Pal DK, Das T, Chaudhury G, et al. Randomised controlled trial to assess acceptability of phenobarbital for childhood epilepsy in rural India. Lancet 1998;351(9095):19-23.
70. Park SP, Hwang YH, Lee HW, et al. Long-term cognitive and mood effects of zonisamide monotherapy in epilepsy patients. Epilepsy Behav 2008;12(1):102-08.
71. Placencia M, Sander JW, Shorvon SD, et al. Antiepileptic drug treatment in a community health care setting in northern Ecuador: a prospective 12-month assessment. Epilepsy Res 1993;14(3):237-44.
72. Privitera MD, Brodie MJ, Mattson RH, et al. Topiramate, carbamazepine and valproate monotherapy: double-blind comparison in newly diagnosed epilepsy. Acta Neurol Scand 2003;107(3):165-75.
73. Pulliainen V, Jokelainen M. Comparing the cognitive effects of phenytoin and carbamazepine in long-term monotherapy: a two-year follow-up. Epilepsia 1995;36(12):1195-202.
74. Ramsay E, Faught E, Krumholz A, et al. Efficacy, tolerability, and safety of rapid initiation of topiramate versus phenytoin in patients with new-onset epilepsy: A randomized double-blind clinical trial. Epilepsia 2010;51(10):1970-77.
75. Ramsay RE, Wilder BJ, Berger JR, et al. A double-blind study comparing carbamazepine with phenytoin as initial seizure therapy in adults. Neurology 1983;33(7):904-10.

76. Ramsay RE, Wilder BJ, Murphy JV, et al. Efficacy and safety of valproic acid versus phenytoin as sole therapy for newly diagnosed primary generalized tonic-clonic seizures. J Epilepsy 1992;5(1):55-60.
77. Rastogi P, Mehrotra TN, Agarwala RK, et al. Comparison of sodium valproate and phenytoin as single drug treatment in generalised and partial epilepsy. J Assoc Physicians India 1991;39(8):606-08.
78. Rating D, Wolf C, Bast T. Sulthiame as monotherapy in children with benign childhood epilepsy with centrotemporal spikes: a 6-month randomized, double-blind, placebo-controlled study. Sulthiame Study Group. Epilepsia 2000;41(10):1284-88.
79. Ravi Sudhir RV, Sawhney IMS, Prabhakar S, et al. Comparative cognitive effects of phenytoin and carbamazepine in adult epileptics. Neurol India 1995;43(4):186-92.
80. Reinikainen KJ, Keränen T, Halonen T, et al. Comparison of oxcarbazepine and carbamazepine: a double-blind study. Epilepsy Res 1987;1(5):284-89.
81. Reunanen M, Dam M, Yuen AW. A randomised open multicentre comparative trial of lamotrigine and carbamazepine as monotherapy in patients with newly diagnosed or recurrent epilepsy. Epilepsy Res 1996;23(2):149-55.
82. Richens A, Davidson DL, Cartlidge NE, et al. A multicentre comparative trial of sodium valproate and carbamazepine in adult onset epilepsy. Adult EPITEG Collaborative Group. J Neurol Neurosurg Psychiatry 1994;57(6):682-87.
83. Rosenow F, Schade-Brittinger C, Burchardi N, et al. The LaLiMo Trial: lamotrigine compared with levetiracetam in the initial 26 weeks of monotherapy for focal and generalised epilepsy—an open-label, prospective, randomised controlled multicenter study. J Neurol Neurosurg Psychiatry 2012.
84. Rowan AJ, Ramsay RE, Collins JF, et al. New onset geriatric epilepsy: a randomized study of gabapentin, lamotrigine, and carbamazepine. Neurology 2005;64(11):1868-73.
85. Sachdeo R, Beydoun A, Schachter S, et al. Oxcarbazepine (Trileptal) as monotherapy in patients with partial seizures. Neurology 2001;57(5):864-71.
86. Sachdeo R, Kramer LD, Rosenberg A, et al. Felbamate monotherapy: controlled trial in patients with partial-onset seizures. Ann Neurol 1992;32(3):386-92.
87. Sachdeo RC, Reife RA, Lim P, et al. Topiramate monotherapy for partial-onset seizures. Epilepsia 1997;38(3):294-300.
88. Sackellares JC, Kwong WJ, Vuong A, et al. Lamotrigine monotherapy improves health-related quality of life in epilepsy: A double-blind comparison with valproate. Epilepsy Behav 2002;3(4):376-82.
89. Saetre E, Abdelnoor M, Amlie JP, et al. Cardiac function and antiepileptic drug treatment in the elderly: A comparison between lamotrigine and sustained-release carbamazepine. Epilepsia 2009;50(8):1841-49.
90. Saetre E, Abdelnoor M, Perucca E, et al. Antiepileptic drugs and quality of life in the elderly: results from a randomized double-blind trial of carbamazepine and lamotrigine in patients with onset of epilepsy in old age. Epilepsy Behav 2010;17(3):395-401.
91. Saetre E, Perucca E, Isojärvi J, et al. An international multicenter randomized double-blind controlled trial of lamotrigine and sustained-release carbamazepine in the treatment of newly diagnosed epilepsy in the elderly. Epilepsia 2007;48(7):1292-302.
92. Shakir RA, Johnson RH, Lambie DG, et al. Comparison of sodium valproate and phenytoin as single drug treatment in epilepsy. Epilepsia 1981;22(1):27-33.
93. So EL, Lai CW, Pellock J, et al. Safety and efficacy of valproate and carbamazepine in the treatment of complex partial seizures. J Epilepsy 1992;5(3):149-52.

94. Sobaniec W, Kułak W, Strzelecka J, et al. A comparative study of vigabatrin vs. carbamazepine in monotherapy of newly diagnosed partial seizures in children. Pharmacol Rep 2005;57(5):646-53.
95. Steiner TJ, Dellaportas CI, Findley LJ, et al. Lamotrigine monotherapy in newly diagnosed untreated epilepsy: a double-blind comparison with phenytoin. Epilepsia 1999;40(5):601-07.
96. Steinhoff BJ, Ueberall MA, Siemes H, et al. The LAM-SAFE Study: Lamotrigine versus carbamazepine or valproic acid in newly diagnosed focal and generalised epilepsies in adolescents and adults. Seizure 2005;14(8):597-605.
97. Stephen LJ, Sills GJ, Leach JP, et al. Sodium valproate versus lamotrigine: A randomised comparison of efficacy, tolerability and effects on circulating androgenic hormones in newly diagnosed epilepsy. Epilepsy Res 2007;75(2-3):122-29.
98. Theodore WH, Albert P, Stertz B, et al. Felbamate Monotherapy: Implications for Antiepileptic Drug Development. Epilepsia 1995;36(11):1105-10.
99. Thilothammal N, Banu K, Ratnam RS. Comparison of phenobarbitone, phenytoin with sodium valproate: randomized, double-blind study. Indian Pediatr 1996;33(7):549-55.
100. Trudeau V, Myers S, LaMoreaux L, et al. Gabapentin in naive childhood absence epilepsy: results from two double-blind, placebo-controlled, multicenter studies. J Child Neurol 1996;11(6):470-75.
101. Turnbull DM, Howel D, Rawlins MD, et al. Which drug for the adult epileptic patient: phenytoin or valproate? BMJ (Clin Res Ed) 1985;290(6471):815-19.
102. Turnbull DM, Rawlins MD, Weightman D, et al. A comparison of phenytoin and valproate in previously untreated adult epileptic patients. J Neurol Neurosurg Psychiatry 1982;45(1):55-59.
103. Verity CM, Hosking G, Easter DJ. A multicentre comparative trial of sodium valproate and carbamazepine in paediatric epilepsy. The Paediatric EPITEG Collaborative Group. Dev Med Child Neurol 1995;37(2):97-108.
104. Wesnes KA, Edgar C, Dean AD, et al. The cognitive and psychomotor effects of remacemide and carbamazepine in newly diagnosed epilepsy. Epilepsy Behav 2009;14(3):522-28.
105. Wheless JW, Neto W, Wang S, et al. Topiramate, carbamazepine, and valproate monotherapy: double-blind comparison in children with newly diagnosed epilepsy. J Child Neurol 2004;19(2):135-41.
106. Wilder BJ, Ramsay RE, Murphy JV, et al. Comparison of valproic acid and phenytoin in newly diagnosed tonic clonic seizures. Neurology 1983;33(11):1474-76.
107. Wilson JT, Höjer B, Tomson G, et al. High incidence of a concentration-dependent skin reaction in children treated with phenytoin. BMJ 1978;1(6127):1583-86.
108. Zamponi N, Cardinali C. Open comparative long-term study of vigabatrin vs carbamazepine in newly diagnosed partial seizures in children. Arch Neurol 1999;56(5):605-07.

Appendix 6: Additional results from Chapter 3

Table 31: Outcomes reported in 54 epilepsy monotherapy studies without a time-to-event outcome defined

Study ID[1]	Journal	Outcomes reported
Aikia 1992	Epilepsy Research	1. Neuropsychological assessment: general intelligence, verbal memory, sustained attention, simple psychomotor speed.
Bawden 1999	Epilepsy Research	1. Cognitive and behavioural effects (intelligence, memory, attention, psychomotor speed, impulsivity) - difference between baseline, 6 weeks and 12 months 2. Emergence of moderate to severe behavioural side effects
Beydoun 1997b	Neurology	1. The change in frequency of complex partial seizures per 8 weeks between the baseline period and the monotherapy portion of the experimental period. 2. The change in seizure frequency of simple partial seizures and secondarily generalized tonic-clonic seizures. 3. The percent change in seizure frequency from baseline (responder rate) 4. Frequency of mild, moderate or severe adverse events
Bittencourt 1993	Epilepsy Research	1. Seizure control 2. Seizure frequency 3. Serum drug concentration 4. EEG results 5. Mental function (cognitive tests and mood scales) 6. Side-effects.
Callaghan 1982	Developmental Medicine and Child Neurology	1. Response to treatment: complete remission, partial remission or no remission 2. Serum levels of anticonvulsant drugs in relation to seizure control 3. Changes in patterns of EEGs 4. Side-effects
Callaghan 1985	Journal of Neurology, Neurosurgery and Psychiatry	1. Seizure control: excellent, good, poor 2. Number and duration of seizures 3. EEG findings 4. Side-effects.
Chen 1996	Epilepsia	1. Seizure frequency (low/moderate/high) and duration; 2. Side-effects 3. Psychometric tests (WISC-R and Bender-Gestalt), 4. Neurophysiological data (P_300 recorded by EEG) 5. AED levels 6. Blood count; liver function test.
Cho 2011	Seizure	1. Change in overnight polysomnography (PSG) scores (sleep latency, REM sleep latency, total sleep time, sleep efficiency, percentage of each sleep stage, arousal index, and Wake time After Sleep Onset (WASO)) from baseline after 4-6 weeks of treatment 2. Change in Sleep questionnaires (sleep diaries, the Pittsburg Sleep Quality Index (PSQI), the Korean version of the Epworth Sleepiness Scale (KESS), Beck's depression inventory-2 (BDI-2) and the Hospital Anxiety Scale (HAS)) and National Hospital Seizure severity Scale (NHS3) from baseline after 4-6 weeks of treatment

Copolla 2004	Epilepsia	1. Proportion of patients remaining seizure-free by treatment group: lack of clinically observed seizures since the previous visit and lack of electroclinical seizures during ambulatory 24-h EEG testing and a video-EEG session with hyperventilation. 2. Incidence of adverse events
Copolla 2007	Brain and Development	1. Seizure-freedom 2. Incidence of adverse events
Craig 1994	Epilepsia	1. Psychological tests: verbal memory, digit symbol substitution, letter cancellation, digit recall, anxiety and depression, recall, visual reproduction, reaction time, motor speed and coordination 2. Health profile questionnaire (neuropsychiatric and other adverse drug effects) 3. Drug concentration in blood 4. Adverse event frequency 5. Seizure control.
Dam 1989	Epilepsy Research	1. Changes in seizure frequency between baseline and the end of each maintenance period 2. Changes in EEG tracings between baseline and the end of each maintenance period 3. Global evaluation of therapeutic efficacy and tolerability by the investigator at the end of each maintenance period 4. Side effects observed by patients and investigators each visit 5. Laboratory tests (WBC counts and liver function tests, Blood pressure and pulse, drug trough serum levels)
Dodrill 1998	Epilepsia	1. Neuropsychological testing difference from baseline: Wechsler Adult Intelligence Scale revised (WAIS-R), Brief Psychiatric Rating Scale (BPRS), Profile of mood states (POMS), Washington Psychosocial seizure Inventory (WPSI), Lafayette Grooved Pegboard, Stroop, Benton Visual Retention, Controlled Oral Word Association, Mood Rating, Symbol Digit Modalities, Rey Auditory Verbal Learning, Wonderlic Personnel, and Digit Cancellation. 2. Relief from seizures (>= 50% reduction in seizures, < 50% reduction in seizures).
Donati 2007	Seizure	1. Cognitive testing: Computerized Visual Searching Task (CVST), assessing mental information processing speed and attention. Rey Auditory Verbal Learning Test (AVLT) and Raven's Standard 2. Progressive Matrices for children: Psychomotor speed, alertness, memory and learning, and non-verbal intelligence. 3. Percentage of patients remaining seizure-free throughout treatment 4. Most common adverse events 5. Treatment satisfaction on a 4 point scale: poor to very good
Easter 1997	Seizure	1. Mean reported percentage weight gain by treatment group 2. Incidence of excessive body weight velocity (weight velocity defined as change in body weight over one year and excessive defined as body weight exceeding the 97th centile)
Eun 2011	Seizure	1. Seizure-free rate over 6 months (maintenance period) by treatment group 2. Change in cognition (neuropsychological), behavior and quality of life from screening to the end of the maintenance phase by treatment group 3. Incidence of adverse events

Eun 2012	Brain and Development	1. Seizure-freedom (seizure-free during 24-wk maintenance period after titration); 2. Greater than 50% seizure reduction after treatment; 3. Cognitive changes (FSIQ, VIQ, PIQ, verbal comprehension, perceptual organisation, attention, concentration); 4. Behavioral changes (social competence, academic competence, total social competence, withdrawn, somatic complaints, depression/anxiety, social problem, thought problem, attention problem, delinquent behaviour, aggressive behaviour, internalizing problem, eternalizing problem, total behaviour problem, Conners [parent and teacher]); 5. Adverse events (rashes, tiredness, somnolence)
Fattore 2011	Epilepsia	1. Responder status: percentage of patients free from clinical seizures on days 13 and 14 2. Freedom from EEG seizures during standard EEG recording with hyperventilation and intermittent photic stimulation on day 14 3. Percentage of patients free from clinical and EEG seizures lasting > 4 seconds on days 4-7 and days 11-14 Percentage of patients free from clinical seizures days 1-14 and free from EEG seizures lasting > 4 seconds on days 7 and 14 4. Percentage change (vs. baseline) in number and total duration of EEG seizures and spikewave discharges lasting >4 s during the 24-h EEG on day 14 5. Percentage of patients with at least 50% reduction (vs. baseline) in the total duration of EEG seizures lasting >4 seconds during the 24-h EEG on day 14 6. Safety and tolerability data (adverse events) Long-term open label follow-up
Faught 1993	Neurology	1. Number of patients meeting escape criteria (doubling of 2-day/monthly seizure frequency, worsening of seizure subtype, prolongation of seizure duration) 2. Adverse events 3. Lab variables (blood counts and urinalyses) 4. Plasma drug concentration 5. Vital sign data 6. Body weights 7. ECG findings.
Feksi 1991	Lancet	1. Seizure activity during therapy (freedom from seizures/ seizure frequency) 2. Difference in drop-out rate 3. Serum drug concentration 4. Adverse events.
Forsythe 1991	Developmental Medicine and Child Neurology	1. Seizure time, duration and severity 2. Side-effects 3. Cognitive assessments (visual recall, auditory recall, visual scanning, concentration, speed of information processing, intellectual functioning and reading)
Homan 1987	Neurology	1. Reasons for failure of the initial drug in 223 patients during 24 months following the onset of therapy Composite score (summation of seizure activity, systemic toxicity and neurotoxicity) 2. Drug serum concentrations present at failure times by drug

Kang 2007	Epilepsia	1. Change on a neuropsychological test battery after 28 weeks of treatment (Bender Gestalt Test (BGT) (correct copy and recall) and KEDI-WISC (Korean Educational 2. Developmental Institute-Wechsler Intelligence Scale for Children) Full Scale Intelligence Quotient (FSIQ), Verbal Intelligence Quotient (VIQ), and Performance Intelligence Quotient (PIQ), and consists of 12 subtests: information, similarities, arithmetic, vocabulary, comprehension, picture completion, picture arrangement, block design, object assembly, coding, digit span, and maze. 3. Evaluation for treatment-emergent adverse events 4. The percentage of patients that were seizure-free
Kaushal 2006	Neurology India	1. Total number of treatment failures (discontinuation or modification to AEDs) in by group due to breakthrough seizures or adverse events during the first 6 months of follow-up
Ketter 1996	Epilepsy Research	1. Seizure type/frequency 2. Psychopathology (anxiety, depression, mania, Bunney-Hamburg, Clinical Global Impression and BPRS scales) 3. Drop-outs.
Kwan 2009	Epilepsy and Behaviour	1. Difference in mean fasting serum insulin concentration at 12 months between the two treatment groups 2. Difference in mean changes from baseline at various time points in metabolic and endocrine measurements and body mass index (BMI) between the two treatment groups and by gender. 3. Frequency of common adverse events experienced by at least %% of subjects by treatment group
Lee 2011	Seizure	1. Change of neuropsychological and cognitive scores from baseline: general intellectual ability, learning and memory, attention and executive function (group-by-time interaction) 2. Frequency of psychological and health related quality of life symptoms 3. Proportion with seizure-freedom during the maintenance period
Levisohn 2007	Epilepsy and Behaviour	1. Seizure reduction from baseline (response rates) 2. Systemic toxicity and neurotoxicity scores (via questionnaires) 3. Physician and patient evaluation of improvement 4. Proportion of patients experiencing most common adverse events during randomised treatment 5. Weight change from baseline
Mitchell 1987	Epilepsia	1. Change in cognitive, intelligence (IQ), behavioural and psychometric scores between baseline, 6 and 12 months. 2. Compliance, drug changes and withdrawal rates 3. Seizure control at 6 and 12 months (excellent / good / fair / poor)
Mitsudome 1997	Brain and Development	1. Effects of anticonvulsants on Rolandic discharges (disappearance of Rolandic discharges)
Nasreddine 2008	Epilepsia	1. The relationship between trough valproate plasma levels and platelet counts 2. Incidence of thromobocytopenia (and relationship with platelet counts) 3. Incidence of adverse events (related to thrombocytopenia) and premature withdrawals

Nejad 2009	International Journal of Pharmacology	1. Incidence of most common adverse events (occuring in more than two patients) 2. Efficacy of anticonvulsants: changes in frequency of myoclonic jerks compared to myoclonic seizure frequency in the six months prior to commencement of treatment 3. Seizure reduction by categories (all seizure types): Seizure-freedom, more than 50% reduction, less than 50% reduction, worsening of seizures
Ogunrin 2005	African Journal of Neurological Sciences	1. Cognitive outcomes: visual and auditory reaction times, psychomotor speed, attention and memory
Park 2008	Epilepsy and Behaviour	1. Change in neuropsychological test scores and cognitive performance from baseline 2. Relationship between neuropsychological test scores and cognitive performance and ZNS dosage 3. Frequency of cognitive and mood complaints between treatment groups 4. Change in seizure frequency at one year from baseline 5. Change in EEG abnormalities at one year from baseline
Placencia 1993	Epilepsy Research	1. Proportion seizure-free at 3, 6 and 12 month follow-ups 2. Proportion seizure-free, with more than 50% seizure reduction and no change in seizure frequency in 6-12 month follow-up period. 3. Incidence of adverse effects
Pulliainen 1995	Epilepsia	1. Change in neuropsychological scores (motor speed, coordination, attention, concentration, memory, learning and reasoning) and Profile of Mood States (POMS) between baseline, 6 and 12 months. 2. Recurrence of seizures in study period 3. Serum drug levels at 6 and 24 months 4. Incidence of adverse events
Ramsay 1983	Neurology	1. Laboratory measures (WBC, platelet, blood drug levels, hematologic indexes etc.) 2. Incidence of side effects (major and minor) 3. Seizure control and treatment failures/discontinuations
Rasgoti 1991	Journal of the Association of Physicians of India	1. Reduction in frequency of seizures by seizure type (response)) 2. Relationship between drug serum levels and response rate
Ravi Sudhir 1995	Neurology India	1. Cognitive outcomes pre and post treatment: verbal and performance intelligence, visual organisation and visuomotor function, memory and dysfunction 2. Relationship between cognitive outcomes and serum drug levels.
Reinikainen 1987	Epilepsy Research	1. Seizure control (antiepileptic efficacy): seizure frequency during therapy (including exacerbation of seizures leading to treatment withdrawal) 2. Incidence of side effects during the study (titration and maintenance phases). 3. Mean dosages and concentrations of drug

Sackellares 2002	Epilepsy and Behaviour	1. Improvements in health related quality of life during the study by treatment groups (mean change) via the Quality of Life in Epilepsy 89 (QOLIE 89) questionnaire 2. Change in body weight during the study (pooled groups) 3. Seizure frequency during the study (pooled groups) 4. Psychological status (at the end of the study, pooled groups) measured by Profile of Mood States, Beck Depression Inventory and Cornell Dysthymia Rating Scale (self-report). 5. Relationships between QOLIE 89 scores and clinical / mood variables
Saetre 2009	Epilepsia	1. Change from baseline to 40 weeks of therapy (mean values by treatment group) of electrocardiography (ECG) parameters via resting 12 lead ECG recording: QRS interval time, heart rate, PQ interval time and QTc interval. 2. Incidence of abnormalities in ECG recordings at baseline and at 40 weeks (by treatment group)
Saetre 2010	Epilepsy and Behaviour	1. Change from baseline health related quality of life scores at 12, 28 and 40 weeks (median by treatment group) assessed by the Side Effects and Life satisfaction (SEALS) inventory and Liverpool Adverse Event Profile (AEP) questionnaires 2. Differences in SEALS and AEP scores for those who did not complete the study 3. Correlation between scores from SEALS and AEP questionnaires
Shakir 1981	Epilepsia	1. Seizure Recurrence during treatment 2. Serum drug levels 3. Incidence of side effects
So 1992	Journal of Epilepsy	1. Proportion of patients free of complex partial seizures (CPS) during the maintenance period 2. Proportion of patients reporting specific adverse events
Sobaniec 2005	Pharmacological Reports	1. Seizure control based on percentage reduction of seizures: very good, good, mild, no effect (efficacy) 2. Most frequently reported treatment adverse events (safety) 3. Changes in EEG activity from baseline
Stephen 2007	Epilepsy Research	1. Percentage of randomised patients achieving a minimum period of 12 months seizure-freedom. 2. Percentage of randomised patients withdrawing due to adverse events 3. Percentage of randomised patients with lack of efficacy at maximum tolerated dose 4. Changes in levels of androgenic hormone levels (testosterone, androstenedione and SHBG levels) 5. Changes in weight and BMI from baseline
Thilothammal 1998	Indian Pediatrics	1. Seizure recurrence 2. Incidence of side effects 3. Relationship between serum AED levels and side effects
Trudeau 1996	Journal of Child Neurology	1. Absence seizure frequency change from baseline to end of double-blind treatment from ambulatory EEGs - response ratio and responder rate (efficacy) 2. Frequency of treatment emergent adverse events (safety) 3. Relationship between gabapentin dosage and plasma concentration 4. Important changes in laboratory assessments from baseline

Turnbull 1982	Journal of Neurology, Neurosurgery, and Psychiatry	1. Outcome of therapy (proportion seizure-free, recurrence of seizures, drug failure, drug withdrawal) by seizure type 2. Compliance 3. Haematological and biochemical parameters
Wesnes 2009	Epilepsy and Behaviour	1. Change of cognitive scores (power of attention, continuity of attention, memory and recognition) from baseline 2. Change of neuropsychological scores from baseline 3. Seizure recurrence / total number of seizures 4. Relationship between computerised and conventional cognitive and neuropsychological outcomes
Wilder 1983	Neurology	1. Seizure control / recurrence of seizures (excellent / good / poor) 2. Incidence of side effects 3. Plasma levels of study drug 4. Laboratory abnormalities
Wilson 1978	British Medical Journal	1. Development of rash after start of phenytoin treatment 2. Relationship between development of rash and plasma concentration
Zamponi 1999	Archives of Neurology	1. Proportion of patients with relapse of seizures 2. Frequency of adverse events (tolerability)

1. Study ID corresponds to first author and year of publication. See Appendix 5 for references of the studies.

Table 32: Outcomes reported in 54 epilepsy monotherapy studies with at least one time-to-event outcome clearly or possibly defined

Study ID[1]	Journal	Outcomes reported[2,4]
Arroyo 2005	Acta Neurologica Scandinavica	*1. Time-to-first partial-onset seizure or generalized-onset tonic clonic seizure during the double-blind phase (P)* 2. Seizure-free rate at 6 months and 1 year 3. Treatment emergent adverse events
Banu 2007	British Medical Journal	1. Behavioural side effects at 1 year (compared to baseline) (P) 2. Seizure control (freedom of seizures during the last quarter of the 12 month follow-up) *3. Time-to-first seizure* *4. Time-to-withdrawal due to adverse events*
Bast 2003[3]	Epilepsia	*1. Rate of treatment failures per group (P)* 2. Changes in sleep EEGs at 4 weeks, 3 months and 6 months
Baulac 2012	The Lancet Neurology	1. Proportion of patients who achieved seizure-freedom for 26 weeks or more (maintenance period) in the per protocol population (P) 2. Incidence of treatment emergent results *3. Time to 26 week (6 months) remission* *4. Time to 52 week (12 month) remission* 5. Proportion of patients with no seizures for at least 52 weeks *6. Time-to-withdrawal because of absence of efficacy or adverse events*
Beydoun 2000	Neurology	1. The percentage of patients meeting one of the exit criteria (P) *2. Time to meeting one of the exit criteria* 3. Incidence of mild, moderate or severe adverse events 4. Clinically relevant laboratory abnormalities

Bill 1997	Epilepsy Research	1. The proportion of seizure-free patients who had at least one seizure during the maintenance period (P) *2. Time to premature discontinuation due to adverse experiences (P)* *3. Rate of premature discontinuations for any reason (P)* 4. Overall assessments of efficacy and tolerability and therapeutic effect 5. Individual adverse experiences 6. Laboratory values 7. Seizure frequency during maintenance
Biton 2001[3]	Neurology	1. Weight change (P) 2. The proportion of patients seizure-free during the entire study 3. Incidence of the most common drug related adverse events *4. Time-to-withdrawal from the study*
Brodie 1995	The Lancet	*1. Time-to-first seizure after 6 weeks of treatment* *2. Time-to-withdrawal* 3. Proportion of randomised patients remaining seizure-free during the last 40 and 24 weeks of trial 4. Percentages of patients who reported adverse events
Brodie 1999	Epilepsy Research	*1. Time-to-first seizure after 6 weeks of treatment* *2. Time-to-withdrawal* 3. Percentage of patients reporting an adverse event 4. Proportion of patients who were both seizure-free in the last 16 weeks of the study and did not discontinue treatment
Brodie 2002a	Epilepsia	*1. Time-to-exit (P)* *2. Percentage of completers / time-to-withdrawal for any reason* *3. Time-to-first seizure* 4. Percentage who remained seizure-free during the final 12 weeks of the 30 week evaluation period *5. Withdrawal rate due to adverse events*
Brodie 2002b	Epilepsy and Behaviour	*1. Time-to-first seizure after 6 weeks dose titration (P)* *2. Time-to-treatment failure* 3. *Time to second, third and fourth seizures after randomisation* 4. Proportion of patients remaining seizure-free after 6 and 12 months of treatment 5. Most frequently reported adverse events
Brodie 2007	Neurology	1. Proportion of per protocol (PP) patients achieving at least 6 months of seizure-freedom at the last evaluated dose (P) 2. 1 year seizure-freedom rate 3. 6 month and 1 year seizure-freedom rate by dose level *4. Time to study withdrawal* 5. Incidence of adverse events
Canadian 1998	Epilepsia	*1. Retention on the study medication for 12 months (P)* 2. Seizure control 3. Incidence of side effects 4. Compliance and tolerance of study medication
Chadwick 1999	The Lancet	*1. Time-to-treatment failure (withdrawal because of lack of therapeutic effects or adverse events) (P)* *2. Time to 6-month remission of seizures* *3. Time-to-first seizure after initial dose stabilisation and* *4. Time-to-withdrawal due to adverse events* 5. Incidence and severity of adverse events

Chadwick 1998	Neurology	*1. Time-to-exit (P)* *2. Time-to-exit plus withdrawals because of adverse events* 3. Completion rate (percentage of patients attending end-of-phase visit), 4. Exit event rate (percentage of patients who experienced an exit event during the evaluation phase), 5. Adverse event withdrawal rate (percentage of patients who withdrew because of adverse events during either titration or evaluation phases), 6. Exit plus adverse event withdrawal rate (the sum of the exit rate plus the adverse event withdrawal rate). 7. Incidence of adverse events
Christe 1997	Epilepsy Research	1. The proportion of seizure-free patients who had at least one seizure during the maintenance period (P) *2. Time to premature discontinuation due to adverse experiences (P)* *3. Rate of premature discontinuations for any reason (P)* 4. Overall assessments of efficacy and tolerability and therapeutic effect 5. Individual adverse experiences 6. Seizure frequency during maintenance
De Silva 1996	The Lancet	*1. Time-to-first seizure after the start of treatment (P)* *2. Time to enter one year remission (P)* 3. Incidence of side effects leading to treatment withdrawal
Edwards 2001	Epilepsy and Behaviour	1. Change in depression scores (Beck Depression Inventory, the Cornell Dysthymia Rating Scale, and the Profile of Mood States) from screening to weeks 10 and 32 (P) 2. Weight change *3. Time-to-withdrawal from the study* 4. Incidence of adverse events 5. Percentage of patients seizure-free
Gilad 2007[3]	Clinical Neuropharmacy	*1. The appearance of a second seizure under treatment or by finishing the 12 month follow-up without seizures (P)* 2. Tolerability: Incidence of adverse events *3. Withdrawals due to adverse events*
Gillham 2000	Seizure	1. Change in Health Related Quality of Life (HRQOL) from baseline according to the modified Side Effect and Life Satisfaction (SEALS) Inventory with subscales Worry, Temper, Cognition, Dysphoria, and Tiredness and the relationship between SEALS scores *2. Withdrawal from the study*
Gilliam 1998	Neurology	1. Proportion of patients in each treatment group meeting escape criteria any time after initiation of concomitant AED withdrawal (P) *2. Time to escape* 3. Incidence of adverse events 4. Changes in plasma concentrations by treatment group
Gilliam 2003	Neurology	*1. Time-to-exit (i.e., time to meeting a seizure-related exit criterion) (P)* 2. Seizure frequency distribution (i.e., the proportion of patients completing the trial seizure-free or experiencing one or two seizures) *3. Time-to-first-seizure* 4. Association of plasma topiramate concentration with time-to-first-seizure. *5. Incidence of adverse events*

Glauser 2007	Journal of Child Neurology	*1. Time-to-first partial-onset or generalized-onset tonic-clonic seizure during the double-blind phase (P)* 2. Probability of being seizure-free for patients remaining in the study at 6 months and at 1 year 3. Incidence of treatment emergent adverse events
Glauser 2010[3]	The New England Journal of Medicine	*1. Freedom from treatment failure after 16 weeks (P)* 2. Attention dysfunction (Confidence Index of 0.60 or higher on the Conners' Continuous Performance Test) 3. Incidence of adverse events by treatment group 4. Drug concentrations between treatment failures and successes within treatment groups
Guerreiro 1997	Epilepsy Research	1. The proportion of seizure-free patients who had at least one seizure during the maintenance period (P) *2. Time to premature discontinuation due to adverse experiences (P)* *3. Rate of premature discontinuations for any reason (P)* 4. Overall assessments of efficacy and tolerability and therapeutic effect 5. Individual adverse experiences 6. Laboratory values 7. Seizure frequency during maintenance
Hakami 2012	Archives of Neurology	1. Proportions of patients who showed improvement in depression symptoms and QOL at 3 months following randomization (P) 2. Proportions of patients who showed improvement in depression symptoms and QOL at 12 months following randomization. 3. Changes in HADS depression and anxiety scores and QOL scores at 3 months and 12 months 4. Liverpool Adverse Events Profile (LAEP) scores at 3 months and 12 months 5. Mean IntegNeuro cognitive function scores at 3 months *6. Freedom from seizures (excluding those occurring during the drug titration period)* *7. Adherence to treatment (time on allocated drug)* 8. *Treatment failure.*
Heller 1995	Journal of Neurology, Neurosurgery, and Psychiatry	*1. Time-to-first seizure after the start of treatment (P)* *2. Time to enter one year remission (P)* 3. Incidence of side effects leading to treatment withdrawal
Kwan 2011	Lancet Neurology	1. Proportion of patients who remained seizure-free for 6 or more continuous months during the efficacy phase (P) *2. Time-to-exit because of lack of efficacy, adverse events, or any reason* *3. Time to 6-month seizure-freedom* *4. Time-to-first seizure after the dose-escalation phase* 5. Number of seizures during the dose escalation phase 6. Monthly seizure frequency for all patients 7. Incidence of adverse events (severity and causal relation) 8. Anxiety and depression on the Hospital Anxiety and Depression Scale (HADS) *4.* Sleep on the Medical Outcomes Sleep Study Scale

Kaminow 2003[3]	Epilepsy and Behaviour	1. Percentage of patients with treatment success: completion of the 24-week Maintenance Phase (MP) without prematurely discontinuing study medication because of inadequate seizure control or unacceptable side effects (P) *2. Mean time (days) to withdrawal from the study.* 3. Percentage of patients seizure-free during the last 8 weeks of the MP and during the 24-week MP. 4. Mean percentage reduction in seizure frequency during MP. 5. Percentage of patients with marked improvement on investigator ratings of clinical status, assessed at Monotherapy Week 24 (on a 7 point scale) 6. Percentage of patients rating current study medication as better than previous treatment in a self-assessment administered at Monotherapy Week 24 (on a 5 point scale) 7. Health-related quality of life assessed at screening and Monotherapy Week 24 with the QOLIE-31 8. Percentage of patients with adverse events during the Escalation/Taper and Maintenance phases
Marson 2007A	The Lancet	*1. Time-to-treatment failure (P)* *2. Time to 1 year (12 month) remission (P)* *3. Time to 2 year remission* *4. Time-to-first seizure* 5. Health related quality of life via the NEWQOL (Newly Diagnosed Epilepsy Quality of Life Battery) 6. Health economic assessment and cost effectiveness of the drugs (cost per QALY gained and cost per seizure avoided) 7. Frequency of clinically important adverse events
Marson 2007B	The Lancet	*1. Time-to-treatment failure (P)* *2. Time to 1 year (12 month) remission (P)* *3. Time to 2 year remission* *4. Time-to-first seizure* 5. Health related quality of life via the NEWQOL (Newly Diagnosed Epilepsy Quality of Life Battery) 6. Health economic assessment and cost effectiveness of the drugs (cost per QALY gained and cost per seizure avoided) 7. Frequency of clinically important adverse events
Mattson 1985	The New England Journal of Medicine	*1. Patient retention (length of time that patient continued to take the randomly assigned drug) (P)* 2. Composite score (combined score for the control of seizures and incidence of adverse events) 3. Total seizure control /Seizure rate 4. Incidence of side effects
Mattson 1992	The New England Journal of Medicine	1. Total number of seizures (of each type) during 12 months 2. Number of seizures per month 3. Percentage of patients with seizures completely controlled *4. Time-to-first seizure* 5. Severity of seizures at 12 and 24 months 6. Composite score (combined score for the control of seizures and incidence of adverse events) 7. Incidence and severity of systemic and neurologic adverse events *8. Time-to-treatment failure*

Mikkelsen 1981[3]	Epilepsia	*1. Percentage (of patients) free of seizures* 2. Number of treatment withdrawals 3. Withdrawal time 4. Number of seizures until time of treatment withdrawal 5. Incidence of side effects
Nieto-Barrera 2001	Epilepsy Research	1. Proportion of patients seizure-free during the last 16 weeks of treatment 2. Efficacy success: proportion of patients who did not withdraw before the end of week 18 and were seizure-free in the last 16 weeks of the study *3. Time-to-withdrawal from the study (proportion of patients completing the study)* 4. Proportion of patients experiencing adverse events 5. Withdrawals due to adverse events
Pal 1998	The Lancet	1. Frequency of behavioural side effects (assessed by Conner's parent rating scale or pre-school behaviour screening questionnaire) at 12 months or at withdrawal (P) 2. Incidence of side effects *3. Time-to-first seizure after randomisation* 4. Actuarial proportion entering each quarter seizure-free
Privitera 2003	Acta Neurologica Scandinavica	*1. Time-to-exit (P)* *2. Time-to-first seizure* 3. Proportion of seizure-free patients during the last 6 months of double-blind treatment. 4. Safety : most commonly occurring adverse events
Ramsay 1992	Journal of Epilepsy	*1. Time-to-first generalised tonic-clonic seizure after initiation of therapy* 2. Six month seizure recurrence rate 3. Frequency of adverse reactions (events) 4. Serum drug levels
Ramsay 2010	Epilepsia	*1. Time-to-first complex partial seizure or generalised tonic clonic seizure (P)* *2. Patient retention (time to discontinuation of treatment)* 3. Incidence and summary of adverse events
Rating 2000[3]	Epilepsia	*1. Rate of treatment failure events per group (P)* 2. Individual change in EEG recordings over time 3. Occurrences of Adverse Events (by exposure days, by group)
Reunanen 1996	Epilepsy Research	1. Proportion completing seizure-free after 6 weeks treatment *2. Time-to-first seizure* *3. Time-to-withdrawal* 4. Frequency of Adverse Events with at least 5% incidence in any treatment group
Richens 1994	Journal of Neurology, Neurosurgery and Psychiatry	*1. Remission analysis (time to 6, 12 and 24 month remission)* *2. Retention analysis (time-to-treatment failure)* 3. Adverse event incidence 4. Incidence of treatment failures due to poor seizure control and adverse events
Rowan 2005	Neurology	*1. Retention in the trial for 12 months (P)* 2. Seizure-freedom at 12 months *3. Time-to-first, second, fifth and tenth seizure (Time to seizures)* 4. Incidence of systemic and neurologic toxicities 5. Serum drug levels and compliance *5.* Seizure-free retention rates

Rosenow 2012	Journal of Neurology, Neurosurgery and Psychiatry	1. Rate of seizure-free patients by treatment group in the first 6 weeks after randomisation (P) 2. Rate of seizure-free patients by treatment group in the last 16 weeks of the trial (categorical) *3. Retention time in the trial (randomisation to 26 weeks)* 4. Rate of seizure-free patients by treatment group in the whole treatment period (26 weeks) *5. Seizure-free time (time-to-first seizure)* 6. Incidence of adverse events 7. Quality of life according to the Quality of Life in Epilepsy 10 (QOLIE-10) questionnaire
Sachdeo 1992[3]	Annals of Neurology	*1. The number of patients in each treatment group who met escape criteria (P)* 2. Seizure frequency data (percentage improvement since baseline) 3. Incidence of adverse experiences 4. Laboratory test values, vital signs and body weight
Sachdeo 1997	Epilepsia	*1. Time until exit (P)* 2. Clinical response to study medication (investigator and patient's global assessments) 3. Reduction from baseline in average monthly seizure rate for patients completing the 16 week double-blind treatment phase (>50%, >75% or 100% reduction) 4. Tolerability of study medication (investigator and patient's global assessments)
Sachdeo 2001	Neurology	*1. Efficacy analysis: Time to meeting one of the exit criteria (P)* 2. Percentage of patients meeting one of the exit criteria 3. Safety analysis: Incidence of adverse events rated as mild, moderate or severe 4. Pharmacokinetic analysis: population-pharmacokinetic modelling of MHD plasma levels
Saetre 2007	Epilepsia	*1. Retention in the trial (time-to-withdrawal of allocated treatment for any cause) (P)* 2. Seizure-freedom after week 4 3. Seizure-freedom after week 20 *4. Time-to-first seizure* 5. Adverse event reports 6. Tolerability according to the Liverpool Adverse Event profile
Steiner 1999	Epilepsia	1. Percentage of patients remaining on treatment 2. Percentage of patients remaining seizure-free in the last 24 and last 16 weeks of treatment 3. Number of seizures (percentage change from baseline) in the last 24 weeks and 16 weeks of treatment *4. Time-to-first seizure after the first 6 weeks of treatment (dose-titration period)* *5. Time to discontinuation* 6. Incidence of adverse events and adverse events leading to discontinuation 7. Quality of Life according to the Side Effects and Life Satisfaction (SEALs) inventory
Steinhoff 2005	Seizure	1. Number of seizure-free patients during weeks 17-24 (P) *2. "Leaving the study" (retention rates)* 3. Adverse event rates

Theodore 1995	Epilepsia	1. Average frequency of all seizure types among study (P) completers during the 14-day evaluation period 2. Average frequency of complex partial seizures, simple partial seizures and secondarily generalised seizures among study completers during the 14-day evaluation period *3. Time to dropout* 4. Incidence of side effects and discontinuations
Turnbull 1985	British Medical Journal	*1. Time to 24-month remission of seizures* *2. Time-to-first seizure* 3. Incidence of adverse events 4. Haematological and biochemical variables
Verity 1995	Developmental Medicine and Child Neurology	*1. Remission analysis (time to 6, 12 and 24 month remission)* *2. Retention analysis (time-to-treatment failure)* 3. Adverse event incidence 4. Rate of withdrawals and treatment failures
Wheless 2004	Journal of Child Neurology	*1. Time-to-exit from the study for any reason (P)* *2. Time-to-first seizure* 3. Proportion of patients seizure-free in the last 6 months of treatment 4. Frequency of most common adverse events

1. Study ID corresponds to first author and year of publication. See Appendix 5 for references of the studies.
2. Outcomes in Italics were considered to be TTE outcomes for further data extraction and outcomes marked with (P) were the study defined primary outcome.
3. Unclear if the study reported at least one TTE outcome

Appendix 7: Search strategies used in Chapter 4

1 Cochrane Central Register of Controlled Trials (CENTRAL) via CRSO *(search dates: 12/06/2014 and 17/08/2015)*

#1 ("individual patient*" ADJ6 (data OR report* OR outcome* OR level*)):TI,AB

#2 ("individual participant data" OR ipd):TI,AB

#3 ("individual subject*" ADJ6 (data OR report* OR outcome* OR level*)):TI,AB

#4 ("raw patient*" ADJ6 (data OR report* OR outcome* OR level*)):TI,AB

#5 ("raw subject*" ADJ6 (data OR report* OR outcome* OR level*)):TI,AB

#6 (idiopathic OR "immediate pigment darkening" OR "intermittent peritoneal dialysis" OR "invasive pneumococcal disease" OR "indirect photometric detection" OR "interaural phase disparity"):TI,AB

#7 #1 OR #2 OR #3 OR #4 OR #5

#8 #7 NOT #6

#9 2004 TO 2014:YR NOT INMEDLINE

#10 #8 AND #9

2 MEDLINE (Ovid) (search dates: 10/06/2014 and 17/08/2015)

1. (individual patient$ adj6 data).ti,ab.

2. (individual patient$ adj6 report$).ti,ab.

3. (individual patient$ adj6 outcome$).ti,ab.

4. (individual patient$ adj6 level$).ti,ab.

5. individual participant data.ti,ab.

6. ipd.ti,ab.

7. (individual subject$ adj6 data).ti,ab.

8. (individual subject$ adj6 report$).ti,ab.

9. (individual subject$ adj6 outcome$).ti,ab.

10. (individual subject$ adj6 level$).ti,ab.

11. (raw patient$ adj6 data).ti,ab.

12. (raw patient$ adj6 report$).ti,ab.

13. (raw patient$ adj6 outcome$).ti,ab.

14. (raw patient$ adj6 level$).ti,ab.

15. (raw subject$ adj6 data).ti,ab.

16. (raw subject$ adj6 report$).ti,ab.

17. (raw subject$ adj6 outcome$).ti,ab.

18. (raw subject$ adj6 level$).ti,ab.

19. idiopathic.ti,ab.

20. immediate pigment darkening.ti,ab.

21. intermittent peritoneal dialysis.ti,ab.

22. invasive pneumococcal disease.ti,ab.

23. indirect photometric detection.ti,ab.

24. interaural phase disparity.ti,ab.

25. or/1-18

26. or/19-24

27. 25 not 26

28. limit 27 to ed=20050601-20140610

3. SCOPUS (search dates: 10/06/2014 and 18/08/2015)

(((((TITLE-ABS-KEY("individual patient*" PRE/6 data)) OR (TITLE-ABS-KEY("individual patient*" PRE/6 report*)) OR (TITLE-ABS-KEY("individual patient*" PRE/6 outcome*)) OR (TITLE-ABS-KEY("individual patient*" PRE/6 level*)) OR (TITLE-ABS-KEY("individual participant data")) OR (TITLE-ABS-KEY(ipd)) OR (TITLE-ABS-KEY("individual subject*" PRE/6 data)) OR (TITLE-ABS-KEY("individual subject*" PRE/6 report*)) OR (TITLE-ABS-KEY("individual subject*" PRE/6 outcome*))) OR ((TITLE-ABS-KEY("individual subject*" PRE/6 level*)) OR (TITLE-ABS-KEY("raw patient*" PRE/6 data)) OR (TITLE-ABS-KEY("raw patient*" PRE/6 report*)) OR (TITLE-ABS-KEY("raw patient*" PRE/6 outcome*)) OR (TITLE-ABS-KEY("raw patient*" PRE/6 level*)) OR (TITLE-ABS-KEY("raw subject*" PRE/6 data))) OR ((TITLE-ABS-KEY("raw subject*" PRE/6 report*)) OR (TITLE-ABS-KEY("raw subject*" PRE/6 outcome*)) OR (TITLE-ABS-KEY("raw subject*" PRE/6 level*)))) AND NOT ((TITLE-ABS-KEY(idiopathic)) OR (TITLE-ABS-KEY("immediate pigment darkening")) OR (TITLE-ABS-KEY("intermittent peritoneal dialysis")) OR (TITLE-ABS-KEY("invasive pneumococcal disease")) OR (TITLE-ABS-KEY("indirect photometric detection")) OR (TITLE-ABS-KEY("interaural phase disparity")))) AND PUBYEAR > 2004) AND NOT (INDEX(medline)) AND (LIMIT-TO(SUBJAREA,"MEDI") OR LIMIT-TO(SUBJAREA,"BIOC") OR LIMIT-TO(SUBJAREA,"NEUR") OR LIMIT-TO(SUBJAREA,"PHAR") OR LIMIT-TO(SUBJAREA,"IMMU") OR LIMIT-TO(SUBJAREA,"NURS") OR LIMIT-TO(SUBJAREA,"HEAL") OR LIMIT-TO(SUBJAREA,"PSYC") OR LIMIT-TO(SUBJAREA,"DENT"))

4 Epilepsy Specialized Register (CRS) (search dates: 12/06/2014 and 17/08/2015)

#1 ("individual patient*" ADJ6 (data OR report* OR outcome* OR level*)):TI,AB

#2 ("individual participant data" OR ipd):TI,AB

#3 ("individual subject*" ADJ6 (data OR report* OR outcome* OR level*)):TI,AB

#4 ("raw patient*" ADJ6 (data OR report* OR outcome* OR level*)):TI,AB

#5 ("raw subject*" ADJ6 (data OR report* OR outcome* OR level*)):TI,AB

#6 (idiopathic OR "immediate pigment darkening" OR "intermittent peritoneal dialysis" OR "invasive pneumococcal disease" OR "indirect photometric detection" OR "interaural phase disparity"):TI,AB

#7 #1 OR #2 OR #3 OR #4 OR #5

#8 (#7 NOT #6) AND >2003:YR

5. CINAHL Plus and PsycINFO (EBSCOhost) (search dates: 10/06/2014 and 18/08/2015)

S9	S7 NOT S6 Published: 20040601-
S8	S7 NOT S6
S7	S1 OR S2 OR S3 OR S4 OR S5
S6	TX (idiopathic OR "immediate pigment darkening") OR TX "intermittent peritoneal dialysis" OR TX "invasive pneumococcal disease" OR TX "indirect photometric detection" OR TX "interaural phase disparity"
S5	"raw subject*" W6 (data OR report* OR outcome* OR level*)
S4	"raw patient*" W6 (data OR report* OR outcome* OR level*)
S3	"individual subject*" W6 (data OR report* OR outcome* OR level*)
S2	"individual participant data" OR TX ipd
S1	"individual patient*" W6 (data OR report* OR outcome* OR level*)

6. Web of Science: Core Collection 1900- and BIOSIS Previews (Biological Abstracts)

(search dates: 10/06/2014 and 18/08/2015)

#12	#5 OR #4 OR #3 OR #2 OR #1 Refined by: **PUBLICATION YEARS:** (2013 OR 2004 OR 2012 OR 2011 OR 2010 OR 2009 OR 2008 OR 2007 OR 2005 OR 2006 OR 2014) AND **Databases:** (WOS OR BIOSIS) AND **[excluding]:Databases:** (MEDLINE) AND **RESEARCH DOMAINS:** (SCIENCE TECHNOLOGY) AND **RESEARCH AREAS:** (BEHAVIORAL SCIENCES OR OTORHINOLARYNGOLOGY OR PHARMACOLOGY PHARMACY OR PSYCHIATRY OR TRANSPLANTATION OR ONCOLOGY OR RHEUMATOLOGY OR ALLERGY OR CARDIOVASCULAR SYSTEM CARDIOLOGY OR ANESTHESIOLOGY OR PSYCHOLOGY OR DERMATOLOGY OR GENERAL INTERNAL MEDICINE OR REHABILITATION OR HEMATOLOGY OR ORTHOPEDICS OR REPRODUCTIVE BIOLOGY OR IMMUNOLOGY OR CRITICAL CARE MEDICINE OR NEUROSCIENCES NEUROLOGY OR

	OPHTHALMOLOGY OR NUTRITION DIETETICS OR TROPICAL MEDICINE OR PUBLIC ENVIRONMENTAL OCCUPATIONAL HEALTH OR GASTROENTEROLOGY HEPATOLOGY OR DENTISTRY ORAL SURGERY MEDICINE OR OBSTETRICS GYNECOLOGY OR GERIATRICS GERONTOLOGY OR PATHOLOGY OR RADIOLOGY NUCLEAR MEDICINE MEDICAL IMAGING OR PARASITOLOGY OR SUBSTANCE ABUSE OR RESEARCH EXPERIMENTAL MEDICINE OR SURGERY OR RESPIRATORY SYSTEM OR DEVELOPMENTAL BIOLOGY OR HEALTH CARE SCIENCES SERVICES OR PHYSIOLOGY OR UROLOGY NEPHROLOGY OR VIROLOGY OR ENDOCRINOLOGY METABOLISM OR PEDIATRICS OR INFECTIOUS DISEASES OR NURSING OR MICROBIOLOGY OR TOXICOLOGY) AND **[excluding]:DOCUMENT TYPES:** (PATENT OR EDITORIAL) *DocType=All document types; Language=All languages;*
#11	#5 OR #4 OR #3 OR #2 OR #1 Refined by: **PUBLICATION YEARS:** (2013 OR 2004 OR 2012 OR 2011 OR 2010 OR 2009 OR 2008 OR 2007 OR 2005 OR 2006 OR 2014) AND **Databases:** (WOS OR BIOSIS) AND **[excluding]:Databases:** (MEDLINE) AND **RESEARCH DOMAINS:** (SCIENCE TECHNOLOGY) AND **RESEARCH AREAS:** (BEHAVIORAL SCIENCES OR OTORHINOLARYNGOLOGY OR PHARMACOLOGY PHARMACY OR PSYCHIATRY OR TRANSPLANTATION OR ONCOLOGY OR RHEUMATOLOGY OR ALLERGY OR CARDIOVASCULAR SYSTEM CARDIOLOGY OR ANESTHESIOLOGY OR PSYCHOLOGY OR DERMATOLOGY OR GENERAL INTERNAL MEDICINE OR REHABILITATION OR HEMATOLOGY OR ORTHOPEDICS OR REPRODUCTIVE BIOLOGY OR IMMUNOLOGY OR CRITICAL CARE MEDICINE OR NEUROSCIENCES NEUROLOGY OR OPHTHALMOLOGY OR NUTRITION DIETETICS OR TROPICAL MEDICINE OR PUBLIC ENVIRONMENTAL OCCUPATIONAL HEALTH OR GASTROENTEROLOGY HEPATOLOGY OR DENTISTRY ORAL SURGERY MEDICINE OR OBSTETRICS GYNECOLOGY OR GERIATRICS GERONTOLOGY OR PATHOLOGY OR RADIOLOGY NUCLEAR MEDICINE MEDICAL IMAGING OR PARASITOLOGY OR SUBSTANCE ABUSE OR RESEARCH EXPERIMENTAL MEDICINE OR SURGERY OR RESPIRATORY SYSTEM OR DEVELOPMENTAL BIOLOGY OR HEALTH CARE SCIENCES SERVICES OR PHYSIOLOGY OR UROLOGY NEPHROLOGY OR VIROLOGY OR ENDOCRINOLOGY METABOLISM OR PEDIATRICS OR INFECTIOUS DISEASES OR NURSING OR MICROBIOLOGY OR TOXICOLOGY) *DocType=All document types; Language=All languages;*
#10	#5 OR #4 OR #3 OR #2 OR #1 Refined by: **PUBLICATION YEARS:** (2013 OR 2004 OR 2012 OR 2011 OR 2010 OR 2009 OR 2008 OR 2007 OR 2005 OR 2006 OR 2014) AND **Databases:** (WOS OR BIOSIS) AND **[excluding]:Databases:** (MEDLINE) AND **RESEARCH DOMAINS:** (SCIENCE TECHNOLOGY) *DocType=All document types; Language=All languages;*
#9	#5 OR #4 OR #3 OR #2 OR #1 Refined by: **PUBLICATION YEARS:** (2013 OR 2004 OR 2012 OR 2011 OR 2010 OR 2009 OR 2008 OR 2007 OR 2005 OR 2006 OR 2014) AND **Databases:** (WOS OR BIOSIS) AND **[excluding]:Databases:** (MEDLINE)

	DocType=All document types; Language=All languages;
#8	#5 OR #4 OR #3 OR #2 OR #1 Refined by: **PUBLICATION YEARS:** (2013 OR 2004 OR 2012 OR 2011 OR 2010 OR 2009 OR 2008 OR 2007 OR 2005 OR 2006 OR 2014) AND **Databases:** (WOS OR BIOSIS) *DocType=All document types; Language=All languages;*
#7	#5 OR #4 OR #3 OR #2 OR #1 Refined by: **PUBLICATION YEARS:** (2013 OR 2004 OR 2012 OR 2011 OR 2010 OR 2009 OR 2008 OR 2007 OR 2005 OR 2006 OR 2014) *DocType=All document types; Language=All languages;*
#6	#5 OR #4 OR #3 OR #2 OR #1 *DocType=All document types; Language=All languages;*
#5	**TOPIC:** ("raw subject*" NEAR/6 (data OR report* OR outcome* OR level*)) *NOT* **TOPIC:** (idiopathic OR "immediate pigment darkening" OR "intermittent peritoneal dialysis" OR "invasive pneumococcal disease" OR "indirect photometric detection" OR "interaural phase disparity") *DocType=All document types; Language=All languages;*
#4	**TOPIC:** ("raw patient*" NEAR/6 (data OR report* OR outcome* OR level*)) *NOT* **TOPIC:** (idiopathic OR "immediate pigment darkening" OR "intermittent peritoneal dialysis" OR "invasive pneumococcal disease" OR "indirect photometric detection" OR "interaural phase disparity") *DocType=All document types; Language=All languages;*
#3	**TOPIC:** ("individual subject*" NEAR/6 (data OR report* OR outcome* OR level*)) *NOT* **TOPIC:** (idiopathic OR "immediate pigment darkening" OR "intermittent peritoneal dialysis" OR "invasive pneumococcal disease" OR "indirect photometric detection" OR "interaural phase disparity") *DocType=All document types; Language=All languages;*
#2	**TOPIC:** ("individual participant data" OR ipd) *NOT* **TOPIC:** (idiopathic OR "immediate pigment darkening" OR "intermittent peritoneal dialysis" OR "invasive pneumococcal disease" OR "indirect photometric detection" OR "interaural phase disparity") *DocType=All document types; Language=All languages;*
#1	**TOPIC:** ("individual patient*" NEAR/6 (data OR report* OR outcome* OR level*)) *NOT* **TOPIC:** (idiopathic OR "immediate pigment darkening" OR "intermittent peritoneal dialysis" OR "invasive pneumococcal disease" OR "indirect photometric detection" OR "interaural phase disparity") *DocType=All document types; Language=All languages;*

Appendix 8: Data extraction form used in Chapter 4

Individual Participant Data Meta-Analyses: Data Extraction Form

Date of Extraction:

Name of data extractor:

Meta-analysis First Author:

Meta-analysis Year:

Meta-analysis Title:

Journal or Source:

Authorship policy (individual authorship, collaborative group, none):

Source of funding:

Clinical area (lung cancer, breast cancer, epilepsy, diabetes etc.):

Design of studies included (Randomised / Non-randomised/ both / Other (diagnostic test accuracy etc.):

Type of studies included (Drug / Device / Observational / Other (diagnostic test accuracy etc.):

Type of pooled analysis: Systematic search performed or existing database of studies pooled/ collaboration?

Number of studies eligible for meta-analysis:

Number of participants in all eligible studies:

Year range of eligible studies:

Number of studies providing IPD:

Number of participants IPD is provided for:

Number of studies providing aggregate data (AD):

Number of participants AD is provided for:

Number of studies excluded due to no IPD or AD available:

Number of patients excluded due to no IPD or AD available:

Year range of studies IPD is not available for:

Any reported reasons that IPD was not provided (data no longer available, authors unwilling to collaborate)?

Were any adjustments/ sensitivity analyses performed to account for missing IPD? Or do meta-analysis authors note the limitation of missing IPD?

Additional notes:

Footnotes:

1. Reasons for IPD not being provided and sensitivity analyses recorded as free text and later classified into broad categories.
2. Source of funding recorded as free text and later classified as Commercial, Non-Commercial, Mixed (Commercial and Non-Commercial), No funding, Not stated.
3. Clinical area was also recorded as free text and later classified in broad categories based on the clinical areas covered by the review groups of the Cochrane Collaboration.

Appendix 9: Additional results from Chapter 4

Table 33 Characteristics of 85 reviews in which IPD was extracted from included studies

IPD-MA Characteristic[1]	Number of IPD-MA[2]
Year of publication of IPD-MA	
1987 – 1995	2
1996 – 2000	3
2001 – 2005	9
2006 – 2010	22
2010 – 2015	49
Clinical area of IPD-MA	
Breast Cancer	1
Cancer (other)	11
Cardiology	9
Central Nervous System, Neurology and Brain Injury	11
Cervical Cancer and Ovarian Cancer	1
Diabetes and Endocrinology	0
Gastroenterology, Colorectal and Gastric Cancer	13
Gynaecology, Pregnancy and Neonatology	3
Haematology, Leukaemia and Blood Cancer	2
Head and Neck Cancer	6
Hepatitis and Liver Disease	1
HIV	1
Infection and Infectious Diseases	3
Injuries and Wounds	1
Lung Cancer	0
Mental and Psychiatric Disorders	5
Musculoskeletal and Pain	4
Other	2
Otolaryngology, Ophthalmology and Periodontology	5
Renal and Urology	3
Respiratory and Pulmonary	0
Stroke, Thrombosis and Hypertension	3
Design of included studies	
Randomised	8
Non-Randomised	59
Diagnostic Test Accuracy	8
Both Randomised and Non-Randomised	10
Type of included studies	
Diagnostic Test Accuracy	8
Drug or device	20
Epidemiology / Risk Factor	28
Non-drug (interventional)	29

Type of IPD-MA	
Cochrane Review	2
Non Cochrane Review	83
Authorship Policy	
Individual authorship	0
Collaborative Group	2
None	83
Source of Funding	
Non-commercial	24
Commercial	2
Mixed	1
No funding	13
Not stated	45
Number of eligible studies	
2 to 5	2
6 to 10	14
11 to 15	7
16 to 20	6
21 to 30	10
31 to 40	12
41 to 50	8
over 50	22
Not stated	4
Number of eligible participants	
under 100	10
101 to 200	12
201 to 500	17
501 to 1000	12
1001 to 5000	10
5001 to 10000	2
over 10000	0
Not stated	22

1. See Table 5 for full definitions of characteristics.
2. All IPD-MA were systematic

Results of additional and sensitivity analyses

See **Table 7** in Chapter 4.3.3.1 for definitions of characteristics and footnotes of the tables

Table 34: Results of univariate logistic regression analysis

IPD-MA Characteristic	100% of IPD retrieved compared to less than 100% of IPD[1]			At least 80% of IPD retrieved compared to less than 80% of IPD[1]		
	OR	95% CI	P-value	OR	95% CI	P-value
Age of publication[2]	0.918	0.748 to 1.126	0.411	1.102	0.897 to 1.355	0.355
Number of eligible participants[2]	0.701	0.625 to 0.787	<0.001	0.872	0.789 to 0.965	0.008
Includes randomised studies only	1.113	0.768 to 1.613	0.571	2.517	1.729 to 3.666	<0.001
Cochrane IPD-MA	0.365	0.179 to 0.746	0.006	0.471	0.264 to 0.839	0.011
Authorship Policy[3]	1.225	0.840 to 1.786	0.292	3.123	2.132 to 4.577	<0.001
Commercial source of funding[4]	1.360	0.834 to 2.218	0.218	1.606	0.938 to 2.750	0.084

Table 35: Multivariable logistic regression: further examination of authorship policy

IPD-MA Characteristic	100% of IPD retrieved compared to less than 100% of IPD[1]			At least 80% of IPD retrieved compared to less than 80% of IPD[1]		
	OR	95% CI	P-value	OR	95% CI	P-value
Age of publication[2]	1.137	0.923 to 1.402	0.228	1.155	0.939 to 1.420	0.173
Number of eligible participants[2]	0.843	0.791 to 0.898	<0.001	0.889	0.838 to 0.943	<0.001
Includes randomised studies only	1.491	0.957 to 2.322	0.078	2.748	1.761 to 4.288	<0.001
Cochrane IPD-MA	0.450	0.208 to 0.973	0.042	0.432	0.221 to 0.844	0.014
No Authorship Policy	*Ref.*	*Ref.*	*Ref.*	*Ref.*	*Ref.*	*Ref.*
Individual authorship	2.795	1.729 to 4.519	<0.001	3.583	2.173 to 5.908	<0.001
Collaborative Group[3]	0.761	0.442 to 1.309	0.324	3.130	1.867 to 5.249	<0.001
Commercial source of funding[4]	1.623	0.933 to 2.824	0.086	1.061	0.576 to 1.953	0.849

Table 36: Multivariable logistic regression: inclusion of type of study

IPD-MA Characteristic	100% of IPD retrieved compared to less than 100% of IPD[1]			At least 80% of IPD retrieved compared to less than 80% of IPD[1]		
	OR	95% CI	P-value	OR	95% CI	P-value
Age of publication[2]	1.068	0.874 to 1.306	0.520	1.138	0.924 to 1.402	0.224
Number of eligible participants[2]	0.845	0.794 to 0.898	<0.001	0.881	0.830 to 0.936	<0.001
Includes randomised studies only	1.220	0.756 to 1.971	0.415	2.253	1.372 to 3.670	0.001
Drug or device	1.374	0.887 to 2.130	0.155	1.492	0.936 to 2.379	0.093
Cochrane IPD-MA	0.403	0.189 to 0.862	0.019	0.429	0.219 to 0.841	0.014
Authorship Policy[3]	1.710	1.101 to 2.658	0.017	3.491	2.252 to 5.413	<0.001
Commercial source of funding[4]	1.207	0.707 to 2.059	0.491	0.948	0.511 to 1.757	0.865

Table 37: Multivariable logistic regression results: excluding studies with no information regarding funding

IPD-MA Characteristic	100% of IPD retrieved compared to less than 100% of IPD[1]			At least 80% of IPD retrieved compared to less than 80% of IPD[1]		
	OR	95% CI	P-value	OR	95% CI	P-value
Age of publication[2]	1.029	0.804 to 1.318	0.818	1.094	0.848 to 1.411	0.490
Number of eligible participants[2]	0.847	0.789 to 0.909	<0.001	0.891	0.833 to 0.952	0.001
Includes randomised studies only	1.318	0.783 to 2.217	0.298	3.013	1.779 to 5.103	<0.001
Cochrane IPD-MA	0.419	0.180 to 0.977	0.044	0.392	0.188 to 0.818	0.013
Authorship Policy[3]	1.726	1.014 to 2.936	0.044	3.583	2.154 to 5.961	<0.001
Commercial source of funding[4]	1.544	0.885 to 2.694	0.126	1.003	0.529 to 1.902	0.992

A sensitivity analysis was conducted including all 760 IPD-MAs, assuming the following scenarios for the 257 IPD-MAs for which the proportion of IPD retrieved was unknown:

a. Less than 80% of IPD was retrieved
b. 80% or more IPD was retrieved
c. 100% of IPD was retrieved

Table 38: Multivariable logistic regression results: a. less than 80% of IPD was retrieved

IPD-MA Characteristic	100% of IPD retrieved compared to less than 100% of IPD[1]			At least 80% of IPD retrieved compared to less than 80% of IPD[1]		
	OR	95% CI	P-value	OR	95% CI	P-value
Age of publication[2]	0.641	0.554 to 0.739	<0.001	0.713	0.624 to 0.816	<0.001
Includes randomised studies only	1.667	1.160 to 2.393	0.006	2.906	2.076 to 4.067	<0.001
Cochrane IPD-MA	0.334	0.161 to 0.701	0.004	0.489	0.278 to 0.861	0.013
Authorship Policy[3]	0.505	0.372 to 0.685	<0.001	0.842	0.635 to 1.115	0.230
Commercial source of funding[4]	1.102	0.687 to 1.767	0.687	0.994	0.642 to 1.541	0.980

Table 39: Multivariable logistic regression results: b. 80% or more IPD was retrieved

IPD-MA Characteristic	100% of IPD retrieved compared to less than 100% of IPD[1]			At least 80% of IPD retrieved compared to less than 80% of IPD[1]		
	OR	95% CI	P-value	OR	95% CI	P-value
Age of publication[2]	0.641	0.554 to 0.739	<0.001	1.206	1.046 to 1.392	0.010
Includes randomised studies only	1.667	1.160 to 2.393	0.006	1.479	0.999 to 2.190	0.051
Cochrane IPD-MA	0.334	0.161 to 0.701	0.004	0.428	0.233 to 0.784	0.006
Authorship Policy[3]	0.505	0.372 to 0.685	<0.001	3.222	2.340 to 4.439	<0.001
Commercial source of funding[4]	1.102	0.687 to 1.767	0.687	1.002	0.577 to 1.743	0.993

Table 40: Multivariable logistic regression results: C. 100% of IPD was retrieved

IPD-MA Characteristic	100% of IPD retrieved compared to less than 100% of IPD[1]			At least 80% of IPD retrieved compared to less than 80% of IPD[1]		
	OR	95% CI	P-value	OR	95% CI	P-value
Age of publication[2]	1.109	0.975 to 1.392	0.114	1.206	1.046 to 1.392	0.010
Includes randomised studies only	0.682	0.492 to 0.947	0.022	1.479	0.999 to 2.190	0.051
Cochrane IPD-MA	0.477	0.269 to 0.846	0.011	0.428	0.233 to 0.784	0.006
Authorship Policy[3]	1.762	1.331 to 2.332	<0.001	3.222	2.340 to 4.439	<0.001
Commercial source of funding[4]	1.092	0.704 to 1.693	0.993	1.002	0.577 to 1.743	0.993

Table 41: Multivariable logistic regression results: proportion of IPD known compared to proportion of IPD unknown

IPD MA Characteristic	Proportion of IPD retrieved unknown compared to proportion of IPD retrieved known		
	OR	95% CI	P-value
Age of publication[2]	0.869	0.761 to 0.992	0.039
Includes randomised studies only	0.361	0.256 to 0.508	<0.001
Cochrane IPD-MA	0.658	0.332 to 1.303	0.231
Authorship Policy[3]	1.397	1.043 to 1.869	0.025
Commercial source of funding[4]	0.849	0.523 to 1.379	0.509

Table 42: Results of fractional logistic regression

IPD-MA Characteristic	OR	95% CI*	P-value
Age of publication[2]	1.339	1.152 to 1.555	<0.001
Number of eligible participants[2]	0.998	0.946 to 1.032	0.591
Includes randomised studies only	2.432	1.775 to 3.333	<0.001
Cochrane IPD-MA	0.446	0.288 to 0.691	<0.001
Authorship Policy[3]	2.511	1.835 to 3.436	<0.001
Commercial source of funding[4]	0.871	0.544 to 1.394	0.565

*Calculated with robust standard errors.

Table 43 Multivariate logistic regression by proportion of study data retrieved

IPD MA Characteristic	100% of study data retrieved compared to less than 100% of study data			At least 80% of study data retrieved compared to less than 80% of study data		
	OR	95% CI	P-value	OR	95% CI	P-value
Age of publication[2]	1.172	0.961 to 1.431	0.116	1.235	1.050 to 1.454	0.011
Number of eligible participants[2]	0.498	0.428 to 0.576	<0.001	0.681	0.610 to 0.759	<0.001
Includes randomised studies only	1.555	1.050 to 2.304	0.028	1.301	0.936 to 1.807	0.117
Cochrane IPD-MA	0.441	0.207 to 0.937	0.033	0.664	0.373 to 1.181	0.163
Authorship Policy[3]	1.078	0.739 to 1.573	0.695	1.851	1.355 to 2.529	<0.001
Commercial source of funding[4]	1.339	0.819 to 2.187	0.244	1.227	0.781 to 1.927	0.375

Appendix 10: References of studies included in Cochrane Epilepsy IPD-NMA

1. Aikia M, Kalviainen R, Sivenius J, Halonen T, Riekkinen PJ. Cognitive effects of oxcarbazepine and phenytoin monotherapy in newly diagnosed epilepsy: one year follow-up. Epilepsy Res 1992;11(3):199-203.
2. Banu SH, Jahan M, Koli UK, Ferdousi S, Khan NZ, Neville B. Side effects of phenobarbital and carbamazepine in childhood epilepsy: randomised controlled trial. BMJ 2007;334(7605):1207.
3. Baulac M, Brodie MJ, Patten A, Segieth J, Giorgi L. Efficacy and tolerability of zonisamide versus controlled-release carbamazepine for newly diagnosed partial epilepsy: a phase 3, randomised, double-blind, non-inferiority trial. Lancet Neurol 2012;11(7):579-88.
4. Bidabadi E. Comparison Of The Effects Of Phenobarbital Versus Carbamazepine As Single Drug Therapy In Partial Seizure With Secondary Generalization In Children. Epilepsia 2009;50:167.
5. Bill PA, Vigonius U, Pohlmann H, Guerreiro CA, Kochen S, Saffer D, et al. A double-blind controlled clinical trial of oxcarbazepine versus phenytoin in adults with previously untreated epilepsy. Epilepsy Res 1997;27(3):195-204.
6. Biton V, Mirza W, Montouris G, Vuong A, Hammer AE, Barrett PS. Weight change associated with valproate and lamotrigine monotherapy in patients with epilepsy. Neurology 2001;56(2):172-7.
7. Brodie MJ, Chadwick DW, Anhut H, Otte A, Messmer S-L, Maton S, et al. Gabapentin versus lamotrigine monotherapy: a double-blind comparison in newly diagnosed epilepsy. Epilepsia 2002;43(9):993-1000.
8. Brodie MJ, Overstall PW, Giorgi L. Multicentre, double-blind, randomised comparison between lamotrigine and carbamazepine in elderly patients with newly diagnosed epilepsy. The UK Lamotrigine Elderly Study Group. Epilepsy Res 1999;37(1):81-7.
9. Brodie MJ, Perucca E, Ryvlin P, Ben-Menachem E, Meencke HJ, Levetiracetam Monotherapy Study G. Comparison of levetiracetam and controlled-release carbamazepine in newly diagnosed epilepsy. Neurology 2007;68(6):402-8.
10. Brodie MJ, Richens A, Yuen AW. Double-blind comparison of lamotrigine and carbamazepine in newly diagnosed epilepsy. UK Lamotrigine/Carbamazepine Monotherapy Trial Group.[Erratum appears in Lancet 1995 Mar 11;345(8950):662]. Lancet 1995;345(8948):476-9.
11. Callaghan N, Kenny RA, O'Neill B, Crowley M, Goggin T. A prospective study between carbamazepine, phenytoin and sodium valproate as monotherapy in previously untreated and recently diagnosed patients with epilepsy. J Neurol Neurosurg Psychiatry 1985;48(7):639-44.
12. Capone L, Carrieri L, Centrone G, Olivieri G, Ragno G, Roca M, et al. Randomised open trial of 35 subjects with epileptic seizures after stroke, treated with levetiracetam or carbamazepine CR. Boll Lega Ital Epilessia 2008(138):203-6.
13. Castriota O, Guido M, Goffredo R, Di Claudio T, Specchio LM. Event-related potentials (ERPS) in the evaluation of the effect of levetiracetam and carbamazepine on cognitive functions in adult newly diagnosed epileptic patients. Preliminary results of an opened randomised trial. Boll Lega Ital Epilessia 2008(138):235-7.
14. Chadwick DW, Anhut H, Greiner MJ, Alexander J, Murray GH, Garofalo EA, et al. A double-blind trial of gabapentin monotherapy for newly diagnosed partial seizures. International Gabapentin Monotherapy Study Group 945-77. Neurology 1998;51(5):1282-8.
15. Chen YJ, Kang WM, So WC. Comparison of antiepileptic drugs on cognitive function in newly diagnosed epileptic children: a psychometric and neurophysiological study. Epilepsia 1996;37(1):81-6.
16. Cho YW, Kim DH, Motamedi GK. The effect of levetiracetam monotherapy on subjective sleep quality and objective sleep parameters in patients with epilepsy: compared with the effect of carbamazepine-CR monotherapy. Seizure 2011;20(4):336-9.
17. Christe W, Kramer G, Vigonius U, Pohlmann H, Steinhoff BJ, Brodie MJ, et al. A double-blind controlled clinical trial: oxcarbazepine versus sodium valproate in adults with newly diagnosed epilepsy. Epilepsy Res 1997;26(3):451-60.

18. Consoli D, Bosco D, Postorino P, Galati F, Plastino M, Perticoni GF, et al. Levetiracetam versus carbamazepine in patients with late poststroke seizures: a multicenter prospective randomized open-label study (EpIC Project). Cerebrovasc Dis 2012;34(4):282-9.
19. Craig I, Tallis R. Impact of valproate and phenytoin on cognitive function in elderly patients: results of a single-blind randomized comparative study. Epilepsia 1994;35(2):381-90.
20. Czapinski P, Terczynski A, Czapinska E. Randomised 36-month comparative study of valproic acid, phenytoin, phenobarbital and carbamazepine efficacy in patients with newly diagnosed epilepsy with partial complex seizures. Epilepsia 1997;38 Suppl 3:42.
21. Dam M, Ekberg R, Loyning Y, Waltimo O, Jakobsen K. A double-blind study comparing oxcarbazepine and carbamazepine in patients with newly diagnosed, previously untreated epilepsy. Epilepsy Res 1989;3(1):70-6.
22. de Silva M, MacArdle B, McGowan M, Hughes E, Stewart J, Neville BG, et al. Randomised comparative monotherapy trial of phenobarbitone, phenytoin, carbamazepine, or sodium valproate for newly diagnosed childhood epilepsy. Lancet 1996;347(9003):709-13.
23. Dizdarer G, Kangin M, Sutcuoglu S, Ozmen Y, Yaprak I. A comparison of carbamazepine and oxcarbazepine in partial epilepsies. Epilepsia 2000;41 Suppl Florence:67-8.
24. Donati F, Gobbi G, Campistol J, Rapatz G, Daehler M, Sturm Y, et al. The cognitive effects of oxcarbazepine versus carbamazepine or valproate in newly diagnosed children with partial seizures. Seizure 2007;16(8):670-9.
25. Eun S-H, Eun B-L, Lee JS, Hwang YS, Kim KJ, Lee Y-M, et al. Effects of lamotrigine on cognition and behavior compared to carbamazepine as monotherapy for children with partial epilepsy. Brain Dev 2012;34(10):818-23.
26. Feksi AT, Kaamugisha J, Sander JW, Gatiti S, Shorvon SD. Comprehensive primary health care antiepileptic drug treatment programme in rural and semi-urban Kenya. ICBERG (International Community-based Epilepsy Research Group). Lancet 1991;337(8738):406-9.
27. Forsythe I, Butler R, Berg I, McGuire R. Cognitive impairment in new cases of epilepsy randomly assigned to carbamazepine, phenytoin and sodium valproate. Dev Med Child Neurol 1991;33(6):524-34.
28. Fritz N, Elger C, Helmstaedter C. Effects of lamotrigine and oxcarbazepine on seizures, cognition, mood and health-related quality of life in patients with untreated epilepsy. Epilepsia 2006;47 Suppl 3:157.
29. Gilad R, Sadeh M, Rapoport A, Dabby R, Boaz M, Lampl Y. Monotherapy of lamotrigine versus carbamazepine in patients with poststroke seizure. Clin Neuropharmacol 2007;30(4):189-95.
30. Guerreiro MM, Vigonius U, Pohlmann H, de Manreza ML, Fejerman N, Antoniuk SA, et al. A double-blind controlled clinical trial of oxcarbazepine versus phenytoin in children and adolescents with epilepsy. Epilepsy Res 1997;27(3):205-13.
31. Heller AJ, Chesterman P, Elwes RD, Crawford P, Chadwick D, Johnson AL, et al. Phenobarbitone, phenytoin, carbamazepine, or sodium valproate for newly diagnosed adult epilepsy: a randomised comparative monotherapy trial. J Neurol Neurosurg Psychiatry 1995;58(1):44-50.
32. Jung DE, Yu R, Yoon J-R, Eun B-L, Kwon SH, Lee YJ, et al. Neuropsychological effects of levetiracetam and carbamazepine in children with focal epilepsy. Neurology 2015;84(23):2312-9.
33. Kalviainen R, Tigaran S, Keranen T, Peltola J, Aikia M. Open, multicenter, randomised comparison of cognitive effects of lamotrigine and slow-release carbamazepine monotherapy in newly diagnosed patients with epilepsy. Epilepsia 2002;43 Suppl 8:47.
34. Kopp UA, Gaus V, Wandschneider B, Schmitz B. Cognitive side effects of levetiracetam in monotherapy: comparison with carbamazepine and valproate. Epilepsia 2007;48 Suppl 3:39-40, Abstract No: P178.
35. Korean Lamotrigine Study Group. An Open, Randomized, Multicenter Comparative Clinical Trial of Lamotrigine and Carbamazepine as Initial Monotherapy in Previously Untreated Epilepsies. J Korean Epilepsy Soc 2008;12(1):27-34.
36. Kwan P, Yip FP, Hui ACF, Leung H, Ng PW, Hui KF, et al. Effects of valproate or lamotrigine monotherapy on the reproductive endocrine and insulin-related metabolic profile in Chinese adults with epilepsy: a prospective randomized study. Epilepsy Behav 2009;14(4):610-6.

37. Lee S-A, Lee H-W, Heo K, Shin D-J, Song H-K, Kim O-J, et al. Cognitive and behavioral effects of lamotrigine and carbamazepine monotherapy in patients with newly diagnosed or untreated partial epilepsy. Seizure 2011;20(1):49-54.
38. Lukic SR, Spasic MJ, Lukic NR. Comparison of valproate and lamotrigine for the treatment of newly diagnosed epilepsy: interim report. Epilepsia 2005;46 Suppl 6:278.
39. Marson AG, Al-Kharusi AM, Alwaidh M, Appleton R, Baker GA, Chadwick DW, et al. The SANAD study of effectiveness of carbamazepine, gabapentin, lamotrigine, oxcarbazepine, or topiramate for treatment of partial epilepsy: an unblinded randomised controlled trial. Lancet 2007;369(9566):1000-15.
40. Marson AG, Al-Kharusi AM, Alwaidh M, Appleton R, Baker GA, Chadwick DW, et al. The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial. Lancet 2007;369(9566):1016-26.
41. Mattson RH, Cramer JA, Collins JF. A comparison of valproate with carbamazepine for the treatment of complex partial seizures and secondarily generalized tonic-clonic seizures in adults. The Department of Veterans Affairs Epilepsy Cooperative Study No. 264 Group. N Engl J Med 1992;327(11):765-71.
42. Mattson RH, Cramer JA, Collins JF, Smith DB, Delgado-Escueta AV, Browne TR, et al. Comparison of carbamazepine, phenobarbital, phenytoin, and primidone in partial and secondarily generalized tonic-clonic seizures. N Engl J Med 1985;313(3):145-51.
43. Mitchell WG, Chavez JM. Carbamazepine versus phenobarbital for partial-onset seizures in children. Epilepsia 1987;28(1):56-60.
44. Miura H. [Aspects of antiepileptic drug therapy in children]. No to Hattatsu 1990;22(2):154-9.
45. Motamedi M, Ghini MR, Etemadi P, Ramim T. Levetiracetam and lamotrigine in old aged onset of epilepsy: A randomized double-blind clinical trial. Tehran Univ Med J 2013;71(9):568-76.
46. NCT01498822. Levetiracetam Versus Oxcarbazepine as Monotherapy to Evaluate Efficacy and Safety in Subject With Partial Epilepsy. June 2011.
47. NCT01954121. Open-label, Randomized, Active-controlled Study of LEV Used as Monotherapy in Patients With Partial-Onset Seizures. 2013.
48. NCT02208492. The Effects on Cognitive Function of Levetiracetam (Keppra®) Compared to Carbamazepine (Tegretol®, Carmazepine®) as Monotherapy for Children With Partial Seizure; A Multicentric Randomized Controlled Study. 2011.
49. Nieto-Barrera M, Brozmanova M, Capovilla G, Christe W, Pedersen B, Kane K, et al. A comparison of monotherapy with lamotrigine or carbamazepine in patients with newly diagnosed partial epilepsy. Epilepsy Res 2001;46(2):145-55.
50. Ogunrin O, Adamolekun B, Ogunniyi A. Cognitive effects of antiepileptic drugs in Nigerians with epilepsy. Afr J Neuro Sci 2005;24(1):18-24.
51. Pal DK, Das T, Chaudhury G, Johnson AL, Neville BG. Randomised controlled trial to assess acceptability of phenobarbital for childhood epilepsy in rural India. Lancet 1998;351(9095):19-23.
52. Placencia M, Sander JW, Shorvon SD, Roman M, Alarcon F, Bimos C, et al. Antiepileptic drug treatment in a community health care setting in northern Ecuador: a prospective 12-month assessment. Epilepsy Res 1993;14(3):237-44.
53. Privitera MD, Brodie MJ, Mattson RH, Chadwick DW, Neto W, Wang S, et al. Topiramate, carbamazepine and valproate monotherapy: double-blind comparison in newly diagnosed epilepsy. Acta Neurol Scand 2003;107(3):165-75.
54. Pulliainen V, Jokelainen M. Effects of phenytoin and carbamazepine on cognitive functions in newly diagnosed epileptic patients. Acta Neurol Scand 1994;89(2):81-6.
55. Ramsay E, Faught E, Krumholz A, Naritoku D, Privitera M, Schwarzman L, et al. Efficacy, tolerability, and safety of rapid initiation of topiramate versus phenytoin in patients with new-onset epilepsy: a randomized double-blind clinical trial. Epilepsia 2010;51(10):1970-7.
56. Ramsay RE, Wilder BJ, Berger JR, Bruni J. A double-blind study comparing carbamazepine with phenytoin as initial seizure therapy in adults. Neurology 1983;33(7):904-10.
57. Ramsay RE, Wilder BJ, Murphy JV, Holmes GL, Uthman B, Slater J, et al. Efficacy and safety of valproic acid versus phenytoin as sole therapy for newly diagnosed primary generalized tonic-clonic seizures. J Epilepsy 1992;5(1):55-60.

58. Ramsay T, Bainbridge J, Fredericks T, Slater J, Nemire R, Ramsay RE. Results of a randomized double-blind comparison of levetiracetam and carbamazepine in new onset seizures in a geriatric population. Epilepsia 2007;48 Suppl 6:36-7.
59. Rastogi P, Mehrotra TN, Agarwala RK, Singh VS. Comparison of sodium valproate and phenytoin as single drug treatment in generalised and partial epilepsy. J Assoc Physicians India 1991;39(8):606-8.
60. Ravi Sudhir RV, Sawhney IMS, Prabhakar S, Pershad D, Nain CK. Comparative cognitive effects of phenytoin and carbamazepine in adult epileptics. Neurol India 1995;43(4):186-92.
61. Resendiz-Aparicio JC, Rodriguez-Rodriguez E, Contreras-Bernal J, Ceja-Moreno H, Barragan-Perez E, Garza-Morales S, et al. [A randomised open trial comparing monotherapy with topiramate versus carbamazepine in the treatment of paediatric patients with recently diagnosed epilepsy]. Rev Neurol 2004;39(3):201-4.
62. Reunanen M, Dam M, Yuen AW. A randomised open multicentre comparative trial of lamotrigine and carbamazepine as monotherapy in patients with newly diagnosed or recurrent epilepsy. Epilepsy Res 1996;23(2):149-55.
63. Richens A, Davidson DL, Cartlidge NE, Easter DJ. A multicentre comparative trial of sodium valproate and carbamazepine in adult onset epilepsy. Adult EPITEG Collaborative Group. J Neurol Neurosurg Psychiatry 1994;57(6):682-7.
64. Rowan AJ, Ramsay RE, Collins JF, Pryor F, Boardman KD, Uthman BM, et al. New onset geriatric epilepsy: a randomized study of gabapentin, lamotrigine, and carbamazepine. Neurology 2005;64(11):1868-73.
65. Saetre E, Perucca E, Isojarvi J, Gjerstad L, Group LAMS. An international multicenter randomized double-blind controlled trial of lamotrigine and sustained-release carbamazepine in the treatment of newly diagnosed epilepsy in the elderly. Epilepsia 2007;48(7):1292-302.
66. Shakir RA, Johnson RH, Lambie DG, Melville ID, Nanda RN. Comparison of sodium valproate and phenytoin as single drug treatment in epilepsy. Epilepsia 1981;22(1):27-33.
67. So EL, Lai CW, Pellock J, Mashman J, Brugger A. Safety and efficacy of valproate and carbamazepine in the treatment of complex partial seizures. J Epilepsy 1992;5(3):149-52.
68. Steiner TJ, Dellaportas CI, Findley LJ, Gross M, Gibberd FB, Perkin GD, et al. Lamotrigine monotherapy in newly diagnosed untreated epilepsy: a double-blind comparison with phenytoin. Epilepsia 1999;40(5):601-7.
69. Steinhoff BJ, Ueberall MA, Siemes H, Kurlemann G, Schmitz B, Bergmann L, et al. The LAM-SAFE Study: lamotrigine versus carbamazepine or valproic acid in newly diagnosed focal and generalised epilepsies in adolescents and adults. Seizure 2005;14(8):597-605.
70. Stephen LJ, Sills GJ, Leach JP, Butler E, Parker P, Hitiris N, et al. Sodium valproate versus lamotrigine: a randomised comparison of efficacy, tolerability and effects on circulating androgenic hormones in newly diagnosed epilepsy. Epilepsy Res 2007;75(2-3):122-9.
71. Suresh SH, Chakraborty A, Virupakshaiah A, Kumar N. Efficacy and safety of levetiracetam and carbamazepine as monotherapy in partial seizures. Epilepsy Res Treat 2015;2015:Article ID 415082.
72. Thilothammal N, Banu K, Ratnam RS. Comparison of phenobarbitone, phenytoin with sodium valproate: randomized, double-blind study. Indian Pediatr 1996;33(7):549-55.
73. Trinka E, Marson AG, Van Paesschen W, Kälviäinen R, Marovac J, Duncan B, et al. KOMET: An unblinded, randomised, two parallel-group, stratified trial comparing the effectiveness of levetiracetam with controlled-release carbamazepine and extended-release sodium valproate as monotherapy in patients with newly diagnosed epilepsy. J Neurol Neurosurg Psychiatry 2012.
74. Turnbull DM, Howel D, Rawlins MD, Chadwick DW. Which drug for the adult epileptic patient: phenytoin or valproate? BMJ (Clin Res Ed) 1985;290(6471):815-9.
75. Verity CM, Hosking G, Easter DJ. A multicentre comparative trial of sodium valproate and carbamazepine in paediatric epilepsy. The Paediatric EPITEG Collaborative Group. Dev Med Child Neurol 1995;37(2):97-108.
76. Werhahn KJ, Trinka E, Dobesberger J, Unterberger I, Baum P, Deckert-Schmitz M, et al. A randomized, double-blind comparison of antiepileptic drug treatment in the elderly with new-onset focal epilepsy. Epilepsia 2015;56(3):450-9.

Appendix 11: Template IPD request letter and data request form

Cochrane Epilepsy Group
Department of Molecular and Clinical Pharmacology
Institute of Translational Medicine
Room 2.28, Clinical Sciences Centre for Research and Education
Lower Lane, Liverpool, L9 7LJ
UK

<Date / month / year>

Dear *<Author>*

We are contacting you on behalf of the Cochrane Epilepsy Group. The Cochrane Collaboration is an international organisation dedicated to providing summaries of the best evidence about medical treatments in the form of systematic reviews. You can find more about the Collaboration and the Cochrane Epilepsy Group at www.epilepsy.cochrane.org

We plan to carry out a Cochrane overview and network meta-analysis of studies comparing monotherapy for patients with epilepsy. The current overview (published 2007) which we wish to update as a Cochrane overview is enclosed:

i) Time-to-withdrawal from treatment (retention time)
ii) Time to 6 and 12 month remission (seizure-free period)
iii) Time-to-first seizure post randomisation.
iv) Adverse events

Due to the complexity in definition of these outcomes, the review is an individual patient data review which allows the most reliable methods to be used for meta-analysis. As you will see, this approach has been used in a number of epilepsy monotherapy Cochrane Reviews.

Following a recent search for reports for the review update, we believe your study may meet the criteria for inclusion in the review update:

<Insert full citation of study>

We would be grateful if you could provide some further information about your study so that we can confirm that it meets our inclusion criteria.

If your study does meet the inclusion criteria we would like to ask if you would be willing to provide individual patient data from your trial. It would also be helpful if you could complete the attached form to indicate the data that you have for your study.

If you would like any further information please do not hesitate to contact us.

Yours sincerely

Sarah Nolan, Statistician (sarah.nolan@liv.ac.uk)
Dr Catrin Tudur Smith, Statistical Editor (cat1@liv.ac.uk)
Professor Tony Marson, Co-ordinating Editor (a.g.marson@liv.ac.uk)

Cochrane Epilepsy Group

Enclosed:

Current Overview: *Multiple treatment comparisons in epilepsy monotherapy trials*
Data request form

Data request form

Cochrane Epilepsy Network Meta-Analysis

Lead Trial Author:

Address:

Telephone:

E-mail:

Study Reference:

Would you be able to supply the following data for each individual patient in the study (delete as appropriate)?

If data is unavailable, please indicate reasons (where possible) in comments section:

Patient Characteristics

- Patient Identifier **YES/NO**
- Age **YES/NO**
- Gender **YES/NO**
- Presence of neurological signs **YES/NO**
- EEG results (prior to randomization) **YES/NO**
- MRI / CT results (prior to randomisation) **YES/NO**
- Aetiology (known/unknown origin of seizures) **YES/NO**
- Number of seizures before randomisation **YES/NO**
- **Dates of seizures before randomisation** **YES/NO**
- Seizure type at randomisation **YES/NO**
    - Partial: simple/complex partial or secondary generalised tonic-clonic
    - Generalised: generalised tonic-clonic
    - Other seizure type (myoclonic, absence etc.)
- Patient type **YES/NO**
    - Newly onset epilepsy
    - Active Seizures
    - Drug resistant seizures etc.

Comments:

Follow-up data

- Date of randomisation **YES/NO**
- Name of drug randomised **YES/NO**
- Dose of drug randomised **YES/NO**
- Dates of follow-up **YES/NO**
- Dates of dose changes **YES/NO**
- Number of seizures after randomisation (at each follow-up) **YES/NO**
- Dates of seizures after randomisation (at each follow-up) **YES/NO**
- Date of withdrawal of randomised treatment **YES/NO**
- Reason for withdrawal of randomised treatment **YES/NO**
- **Adverse events (at each follow-up)** **YES/NO**

Comments:

Trial Methods

- Method of randomisation (generation of random list) **YES/NO**
- Method of concealment of randomisation **YES/NO**
- Stratification factors **YES/NO**
- Blinding methods (if applicable) **YES/NO**
- Analysis approach **YES/NO**
    - Intention to treat / per protocol / other

Comments:

Please indicate any further information you feel may be relevant:

Please return completed form to:

Professor A.G Marson, Co-ordinating Editor, Cochrane Epilepsy Group, Department of Molecular and Clinical Pharmacology, Institute of Translational Medicine, Room 2.28, Clinical Sciences Centre for Research and Education, Lower Lane, Liverpool, L9 7LJ, UK

Appendix 12: Procedure for preparing a raw IPD dataset into an analysis ready dataset

Context: Data is received via an e-mail attachment, CD, other portable device or via remote SAS data access system (ClinicalStudyDataRequest.Com)

Stage 1: Initial general checks to make on the data in the format in which it is sent:

1. Has all additional documentation required been provided with the data?
    - *Documents: Protocols, Case Report Forms, Variable Labels or Formats etc.*
2. What format is/are the dataset(s) and accompanying documentation in?
    - Do the data and accompanying documents open?
    - Is any reformatting required?
3. Is documentation and narrative text within the data written in clearly in English?
    - *Translation may be required into English from authors or Cochrane translators.*
4. Has a single dataset been provided or multiple datasets?
    - *If multiple datasets, how has the data been split? Can the datasets be linked with a Unique Identifier?*
5. Did the data provider previously complete an individual participant data request form to specify which data could be provided?
    a. *If yes, does the data provider give data for all of the variables specified as available on the IPD request form or provide reasons why data is not available?*
    b. *Has sufficient data been provided to calculate outcomes? (i.e. randomisation date, withdrawal reason and date, last follow-up date, seizure dates following randomisation)*
6. Are variable labels with full and clear descriptions provided?
    - *Does each column of data have a variable name and label?*
    - *Are clear labels given for coded/categorical variables?*
7. Is original or de-identified data provided? If data has been de-identified, are sufficient details provided of the de-identification?
    - *For example, whether actual dates or offset dates are provided, do empty cells correspond to missing or redacted values?*

Action to take following checks

- Contact with data providers will be required if:
    - Data does not open
    - Data is not in an appropriate format
    - Essential accompanying documentation is missing or if insufficient
    - Important data which was reported to be available is missing
    - Data is not clearly labelled or described.
- Data which satisfies all of these checks (in the first instance or following contact with data providers) can now be imported into SAS and proceed with Stage 2 checking.

Stage 2: Checks on the content of the data (general):

1. Has a unique identification number been provided for each participant in the dataset?
    - Is this the original randomisation number (or equivalent) or has a new number been produced for de-identification purposes?
    - If a new de-identified ID number is used, are details provided of how this number was generated? If not, no checks can be performed.
    - If the original number is used or a de-identified number has been produced which takes account of the original sequence:
        - Are any numbers missing from the sequence which may indicate excluded patients?
        - Does the randomisation sequence appear random? (i.e. no trends are present in participants being randomised to any of the drugs, baseline characteristics, prognostic factors are balanced across the allocations)
2. For each variable start with exploratory analyses (range for continuous variables, frequencies in each category for categorical variable)
    - Are there any extreme (e.g. age of 100) or impossible (age of -1) values?
    - How much data is for missing each variable?
    - Will missing data for a given variable affect calculation of outcomes (e.g. missing seizure dates)?
3. In comparison to any publications (journal article / clinical trials.gov entry etc.) or other related documents (protocols / clinical study reports etc.):
    - Is the same number of participants included in the publication / document provided in the dataset?
    - Does all information reported in publication/ document match what is provided in the data? For example
        - *Inclusion and exclusion criteria*
        - *Enrolment / randomisation dates / other specified dates*
        - *Demographics and participant characteristics*
        - *Number of participants contributing to each outcome (exclusions etc.)*
        - *Number of participants randomised to each drug and doses started / achieved.*
    - Can relevant results within the publication be recreated accurately (where possible with data provided)?
        - If not, why not? For example numerical differences between published results and IPD, unclear methods of analysis in publications etc.

Action to take following checks

- If data does not seem to follow a random sequence, enquire with the data provider regarding randomisation methods. Data which is not truly random the data may not be able to be used.
- If large amounts of missing data are present, particularly within variables needed to code outcomes, contact data provider for reasons of missingness:

    - If missingness cannot be explained or data required cannot be provided for an outcome, the data cannot be included in the analysis of the outcome.
- Contact data provider for clarification of numerical inconsistencies between published results and IPD or where published results cannot be recreated due to unclear methodology.
- Make a list of 'general' inconsistencies (prioritising large numerical inconsistencies or those which influence calculation of outcomes) for the attention of the data provider.
    - This list will be combined with any queries or inconsistencies following Stage 3.

Stage 3: Checks on the content of the data and initial coding of outcomes for the analysis dataset (clinically (epilepsy) specific):

1. Checks to make on dates/ relative days (convert dates to relative days before and after randomisation if desired):
    - If provided, is date of birth the earliest date in the dataset?
    - If provided, is date of death the latest date in the dataset?
    - If provided, do all dates of seizures before randomisation occur before date of randomisation?
    - If provided, do follow-up dates / post randomisation visit dates occur after the date of randomisation?
        - Do follow-up dates occur in order? (E.g. follow-up 2 occurs after follow-up 1).
        - Do follow-up dates match (approximately) information reported in publications (e.g. follow-ups every six weeks)?
        - If any follow-up dates are missing, how is this recorded? For example if a participant misses follow-up visit 2 at 2 months but attends the third visit at 3 months, is this visit recorded as visit 2 or visit 3 for this participant?
    - If provided, do all dates of seizures after randomisation occur after the date of randomisation?
    - If provided, does the date of withdrawal of allocated treatment occur after randomisation?
        - Is the date of withdrawal of allocated treatment also the last date in the dataset for all participants? If not, are dates recorded after treatment withdrawal (e.g. death, seizure dates for the participant on an alternative treatment)?
2. Data provided relating to epilepsy diagnoses and seizure types:
    - Where results of pre-randomisation investigations are reported (EEG, CT, MRI, Neurological Signs):
        - Are results reported as normal / abnormal or are specific results given?
        - Where specific results are given are the results clearly described and clinically reasonable (consult AGM for assistance here if needed)?
    - Are all participants defined as 'newly diagnosed' (specified in dataset or publications)?
        - For trials also recruiting participants with relapsed seizures, are any concomitant antiepileptic drug treatments reported for these participants during the trial?

- Where seizures before randomisation are specified, did all participants have at least two seizures before randomisation or satisfy other specified diagnostic criteria of epilepsy?
- Has the seizure type for all participants been classified before randomisation?
  - If applicable, are any participants over the age of 30 are classified as experiencing newly diagnosed generalised-onset seizures?
  - If types of seizures before randomisation are specified, do seizure types recorded before and after randomisation correspond?
  - Where generalised seizures are specified before randomisation, do all participants experience a generalised tonic clonic seizure with or without other generalised seizure types? Are other generalised seizure types only without other than tonic clonic seizures reported for any participants?

3. Is sufficient information provided to calculate primary effectiveness outcome for all participants (time-to-withdrawal of allocated treatment)?
   - Date of withdrawal of allocated treatment / time spent on allocated treatment from randomisation.
   - Reason for withdrawal of allocated treatment in enough detail to allow a judgement whether the withdrawal was related to the allocated drug or not.
4. Is sufficient information provided to calculate secondary efficacy outcomes for all participants (time-to-first seizure and time-to-6-month and 12 month remission)?
   - Dates / relative dates (e.g. study day) of seizure recurrences after randomisation (first seizure recurrence and subsequent seizure recurrences)
   - Numbers of seizures over a specific time period (between follow-up visits)
   - Total number of seizures during the follow-up period
5. Adverse event / side effects (if provided)
   - How are adverse events recorded? Narratively as reported by participants or according to dictionary terms (e.g. MEDRA high level or low level terms)?
   - Are adverse events reported at each follow-up or overall across the whole study?
   - Is it specified when adverse events were 'serious' or led to treatment withdrawal?
   - Is it specified whether adverse events were likely to be directly caused by study drug?

Action to take following checks

- Enquire with data providers regarding any potentially wrong dates (e.g. follow-up data reported before randomisation or apparently outside the scope of the study) with inconsistencies of dates relating to outcomes (e.g. withdrawal or seizure dates) taking priority.
- Enquire with data provider regarding any uncertainties regarding diagnosis of epilepsy and classification of seizure type. Particularly note where potentially:
  - Participants may not have had a certain diagnosis of epilepsy
  - Participants may not be treated with monotherapy
  - Participants were not experiencing either partial-onset or generalised-onset tonic clonic seizures with or without other seizure types

- Participants over the age of 30 were experiencing new onset generalised seizures or seizures of both partial and generalised types.
- Is enough data provided for the calculation of appropriate outcomes? If not, enquire with data providers for additional information.
- Enquire with data providers regarding any uncertainties or inconsistencies regarding adverse events, particularly serious adverse events and/or those resulting in treatment withdrawal.

Make a list of specific inconsistencies (by order of magnitude and importance) for the attention of the data provider to be combined with any enquiries from Stage 2. In order of priority:

1. Enquiries relating to variables used in the calculation of outcomes
2. Large inconsistencies within other variables in the dataset (compared to publications)
3. Any other enquiries

Data which satisfies all checks and is sufficient for analysis (in the first instance or following clarification from data provider) can now be used to calculate outcomes for analysis.

Note: If inconsistencies remain following clarification with data provider which cannot be resolved, discuss the magnitude of inconsistencies with AGM and CTS before proceeding to Stage 4.

Stage 4: Coding of effectiveness and efficacy outcomes

Primary outcome: Time-to-withdrawal of allocated treatment

Two variables are to be calculated:

- Time-to-withdrawal of allocated treatment = 'withtime'
- Censoring indicator for withtime = 'wcens'

Coding 'Time-to-withdrawal of allocated treatment'

1. The following reasons for withdrawal are classed as **'events'** in time-to-event analysis (drug related withdrawals) and the censoring indicator **wcens = 1**;
    a. Recurrent seizures, lack of efficacy etc.
    b. Intolerable adverse events, side effects, poor tolerability etc.
    c. Combination of lack of efficacy and poor tolerability
    d. Non-compliance, poor compliance, patient choice etc.
    e. Any other reason described as related to the allocated drug

If wcens=1 and both date of withdrawal and date of randomisation are not missing:

Withtime = Date of withdrawal – Date of randomisation

2. The following reasons for withdrawal are classed as **'censored'** in time-to-event analysis (withdrawals not related to the study drug) and the censoring indicator **wcens = 0**;
    a. Completed the study without withdrawing from treatment
    b. Loss to follow-up
    c. Remission of seizures
    d. Death (for reasons not related to study drug)
    e. Any other reason not related to the study drug

If wcens=0 and participant completed the study without withdrawing (reason a):

Withtime = Study duration or

Withtime = Date of last follow-up – Date of randomisation

If wcens=0 and participant withdrew for a reasons not related to the study drug (reasons b to e):

Withtime = Date of withdrawal/last follow-up – Date of randomisation

Points to consider when coding this outcome

- If all participants completed the study, time-to-withdrawal of allocated treatment cannot be calculated.
- If reason for withdrawal, date of withdrawal / last follow-up or date of randomisation are missing (following clarification from data providers, see Stage 3), wcens and/or withtime are classed as missing for the participant
- Where any minor inconsistencies exist between data and publication results (reported number of withdrawals or reasons) or unclear reasons for withdrawal are documented which could not be resolved by data provider in Stage 3; if assumptions are made (e.g. whether to class a reason for withdrawal as event or censored), perform sensitivity analysis to test robustness of results.

Efficacy outcomes: Seizure and Remission outcomes

Six new variables are to be calculated:

- Time-to-first seizure after randomisation = 'seztime'
- Censoring indicator for seztime = 'scens'
- Time-to-12-month remission = 'remtime'
- Time-to-6-month remission = 'remtime6'
- Censoring indicator for remtime = 'rcens'
- Censoring indicator for remtime6 = 'rcens6'

The procedure for coding the outcomes described below assumes the data is in "wide" format:

Example (with seizure dates provided):

ID	Seizure1	Seizure2	Seizure3	Seizure4	Seizure5	Seizure6
101						
102	01/01/2010	04/01/2010	22/01/2010	26/01/2010	05/06/2010	01/09/2010
103	06/06/2010	01/01/2011	03/04/2011			

Example (no seizure dates provided):

ID	Visit1	Seizures	Visit2	Seizures	Visit3	Seizures
111	01/02/2010	0	01/03/2010	0	01/06/2010	0
112	22/01/2010	2	25/03/2010	1	02/07/2010	0
113	01/06/2010	5	05/07/2010	1	01/12/2010	1

Therefore data in "long" format (below) must be reshaped to "wide" format (use PROC TRANSPOSE in SAS, reshape in Stata or reshape() in R)

Example (with seizure dates provided)

ID	Seizuredate
102	01/01/2010
102	04/01/2010
102	22/01/2010
102	26/01/2010
102	05/06/2010
102	01/09/2010
103	06/06/2010
103	01/01/2011
103	03/04/2011

Example (no seizure dates provided)

ID	Visitdate	Seizures
111	01/02/2010	0
111	01/03/2010	0
111	01/06/2010	0
112	22/01/2010	2
112	25/03/2010	1
112	02/07/2010	0
113	01/06/2010	5
113	05/07/2010	1
113	01/12/2010	1

Coding time-to-first seizure before randomisation

If no seizure recurrence occurs at any point during follow-up, the participant is **censored** for this outcome and **scens=0**

Seztime = Date of last follow-up – Date of randomisation

If a seizure occurs at any point during follow-up, an **event** has occurred for this outcome and **scens=1.**

1. Calculation of seztime if scens =1 and seizure dates are provided

Sort seizure dates into ascending order and minimum date is the date of first seizure

Seztime = Date of first seizure after randomisation – Date of randomisation

2. Calculation of seztime if scens =1 and seizure dates are not provided

Calculation of the following variables is needed for the following steps

- Total number of seizures in time period X = 'totalsezX'
    - Sum of all seizures in the time period (including if seizures split by seizure type)
- Total number of seizures in the whole study = 'totalsez'
    - Sum of all seizures in the study (including if seizures split by seizure type)
- Follow-up time in days of time period X = 'fuX'
    - fuX = Date of fu(X) – Date of Randomisation
- Time between follow-up visits in days = 'futimeX'
    - futimeX= Date of fu(X) – Date of fu(X-1) where Date of fu(0) is date of randomisation
    - futimeX=Last day of fu(X) – first day of fu(X)

a. Number of seizures over a specified time period X (e.g. between follow-up visits) is given

- Calculate approximate seizure times during the time period X by assuming a uniform distribution of seizure times across time period X (rounded up to the nearest day):

  Days been seizures in time period X (sezrateX) = CEILING (futimeX / (totalsezX +1))

For example: over follow-up period 1 from 01/01/2010 to 17/01/2010 (16 days), 3 seizures occurred

sezrate1 = CEILING (16/ (3+1))) = 4 days – i.e. a seizure occurred every 4 days

- For n seizures occurring during time period X, the dates of the seizure i where i=1…n :

sezdateX(i) = sezdateX(i-1)+sezrateX where sezdateX(0) is the first day of time period X

For example: over follow-up period 1 from 01/01/2010 to 17/01/2010 (16 days), 3 seizures occurred

sezdate1(1) = 01/01/2010 + 4 days = 05/01/2010

sezdate1(2) = 05/01/2010 + 4 days = 09/01/2010

sezdate1(3) =09/01/2010 + 4 days = 13/01/2010

- Date of first seizure (first seizure) is the minimum date across all time periods (1, 2, …, N)

Firstsez = minimum (sezdate1(1), sezdate1(2), …, sezdateX(i))

Seztime = Date of first seizure after randomisation (firstsez) – Date of randomisation

b. Total number of seizures across the whole study duration are given

- Calculate approximate seizure times, assuming uniform distribution of seizure times across the whole study duration (rounded up to the nearest day):

sezrate=CEILING (study duration / (totalsez +1))

For example over a 24 week study (168 days from 01/02/2011 to 19/07/2011), 3 seizures occurred

sezrate=CEILING (168/ (3+1)) = 42 days – i.e. a seizure occurred every 42 days

- For n seizures occurring over the study duration, the dates of the seizure i where i=1…n :

sezdate(i) = sezdate(i-1)+sezrate where sezdateX(0) is the first day of time period X

For example over a 24 week study (168 days from 01/02/2011 to 19/07/2011), 3 seizures occurred

sezdate(1) = 01/02/2011 + 56 days = 15/03/2011

sezdate(2) = 15/03/2011 + 56 days = 26/04/2011

sezdate(2) = 26/04/2011 + 56 days = 07/06/2011

- Date of first seizure (first seizure) is the minimum date across all time periods (1, 2, …, N)

Firstsez = minimum (sezdate(1), sezdate(2), …, sezdate(i))
Seztime = Date of first seizure after randomisation (firstsez) – Date of randomisation

c. If seizure recurrence status is reported over a specified time period X (e.g. seizure-free since last follow-up, yes or no)

- Calculate approximate first seizure time (sezday) during time period X, assuming uniform distribution of seizure times across the time period X (rounded up to the nearest day):

SezdateX= CEILING (fuX + (futimeX/2))

For example seizure-free up to 3 weeks (21 days between 01/09/2012 and 22/09/2012) and seizure recurrence happened in week 4 (7 days between 23/09/2012 and 30/09/2012).

SezdateX=CEILING (21 + (7/2)) = 24days – i.e. on the 24/09/2012

- Date of first seizure (first seizure) is the minimum date across all time periods (1, 2, …, N)

Firstsez = minimum (sezdate1, sezdate2, … , sezdateN)
Seztime = Date of first seizure after randomisation (firstsez) – Date of randomisation

Coding time to remission of seizures after randomisation (6 and 12 month remission).

The following steps calculate 12 month remission but can also be extended to a remission period of 6 months or any other length by replacing 365 days with the required number of days of remission.

Calculation of the following variable is needed:

Follow-up time (maxfu) = Date of last follow-up – date of randomisation

1. **If participant has less than 12 months follow-up (i.e. maxfu < 365 days)**

The participant is **censored** for the outcome (cannot achieve 12 month remission); **rcens=0** and **remtime=maxfu**

2. **If participant has at least 12 months follow-up and no seizure recurrence during the trial**

The participant has an **event** for the outcome (immediate remission); **rcens=1** and **remtime=365**

3. **If participant has at least 12 months follow-up and one-or-more seizures during the trial**

Using exact seizure dates / days provided (or estimates where number of seizures are provided using steps 2a – 2c described in the coding of 'time-to-first seizure') sorted into ascending order across all time periods, calculate the time between each seizure during the trial (i = 0…n, where n is the last seizure) :

Time between seizures (sezdiff(i))=sezdate (i+1)- sezdate(i)

Note that *sezdiff(0) = sezdate(1) – date of randomisation*

Note that *sezdiff(n) = Maxfu – sezdate(n)*

Calculate the maximum difference (maxdiff) between seizure times:

maxdiff = maximum (sezdiff(0), sezdiff(1)… sezdiff(n))

- If maxdiff less than 365 then the participant does not have a remission period of at least 365 days during the trial; **rcens=0** and **remtime=maxfu**

- If maxdiff is at least 365 then the participant does have a remission period of at least 365 days; **rcens=1**
    - For each participant, identify the first time between seizures (sezdiff (i)) greater than 365; by sorting sezdiff (i) variables or by manual inspection. Let sezdiff(j) be the first time between seizures greater than 365:

Remtime=sezdate(j)+365

Points to consider when coding these outcomes (first seizure and reemission)

- If no participants experienced seizure recurrence, time-to-first seizure cannot be calculated
- If the duration of the study was less than 6 or 12 months, time to 6 and 12 month remission respectively cannot be calculated
- If all data on seizure recurrence, all follow-up dates or date of randomisation is missing, the participant is excluded from analyses.
- If relative study days are provided rather than dates (e.g. seizures occurred on day 5, 10, 15 etc. of the study) then 'firstsez' and 'seztime' are equivalent.
- If any minor inconsistencies exist between data and publication results (reported number of seizure recurrences etc.) which cannot be resolved by the data provider, or if seizure data required for the calculation of outcomes is missing, censor outcomes at the time of last follow-up. Consider sensitivity analysis excluding these participants from analysis and/or assuming seizure recurrence occurs during earlier or subsequent time periods. The presence of missing data may require visual inspection to judge whether remission occurred.
- Where exact seizure recurrence dates are not available and seizure dates are estimated (using methods 2 a – 2c outlined above in the coding of 'time-to-first seizure', consider sensitivity analysis (for example, assuming seizure recurrence time over a range of times over the time period X or study duration).

Stage 5: Adding new data into the Epilepsy IPD master dataset

The analysis ready dataset has the following variables, create an analysis dataset for the new data with as many of the following variables as possible:

Note: Missing data is indicated by ` . ' for all variables.

TRIALNO = Unique patient identifier for each participant in each trial (may be original or de-identified ID provided with data, or a new unique ID created)

TRIAL = Trial number assigned to each trial (see assignment list prepared by SJN)

RAND = Date of randomisation

DRUG= Randomised drug (labelled from 1 – 10)

1= Carbamazepine (CBZ)	2= Phenobarbitone (PHB)
3= Phenytoin (PHT)	4= Sodium Valproate (VPS)
5= Lamotrigine (LTG)	6= Oxcarbazepine (OXC)
7= Levetiracetam (LEV)	8= Topiramate (TPM)
9= Gabapentin (GBP)	10 Zonisamide (ZNS)

Outcome variables created in Stage 4

WITHTIME = Time-to-withdrawal of allocated treatment

WCENS = Withdrawal censoring indicator (1 = event, 0= censored)

REMTIME = Time-to-12-month remission

RCENS = 12 month remission censoring indicator (1 = event, 0= censored)

REMTIME6= Time-to-6-month remission

RCENS6 = 6 month remission censoring indicator (1 = event, 0= censored)

SEZTIME = Time-to-first seizure after randomisation

SCENS = First seizure censoring indicator (1 = event, 0= censored)

Demographic variables (if data provided, further coding may be needed from original dataset)

AGE = age at randomisation

SEX = Gender (1 = female, 0 = male)

EPTYPE= Epilepsy Type (1 = partial, 0 = generalised)

EPTYPE2 = Epilepsy Type Reclassified (participants with generalised seizures and age of onset over 30 years reclassified to partial - 1 = partial, 0 = generalised)

EPTYPE3 = Epilepsy Type Reclassified (participants with generalised seizures and age of onset over 30 years and participants with missing seizure type reclassified to unknown seizure type – 2 – unknown, 1 = partial, 0 = generalised)

NEURSIGN=Presence of Neurological Signs (1 = Yes, 0 = No)

EEG = EEG results (1 = abnormal 0 = normal)

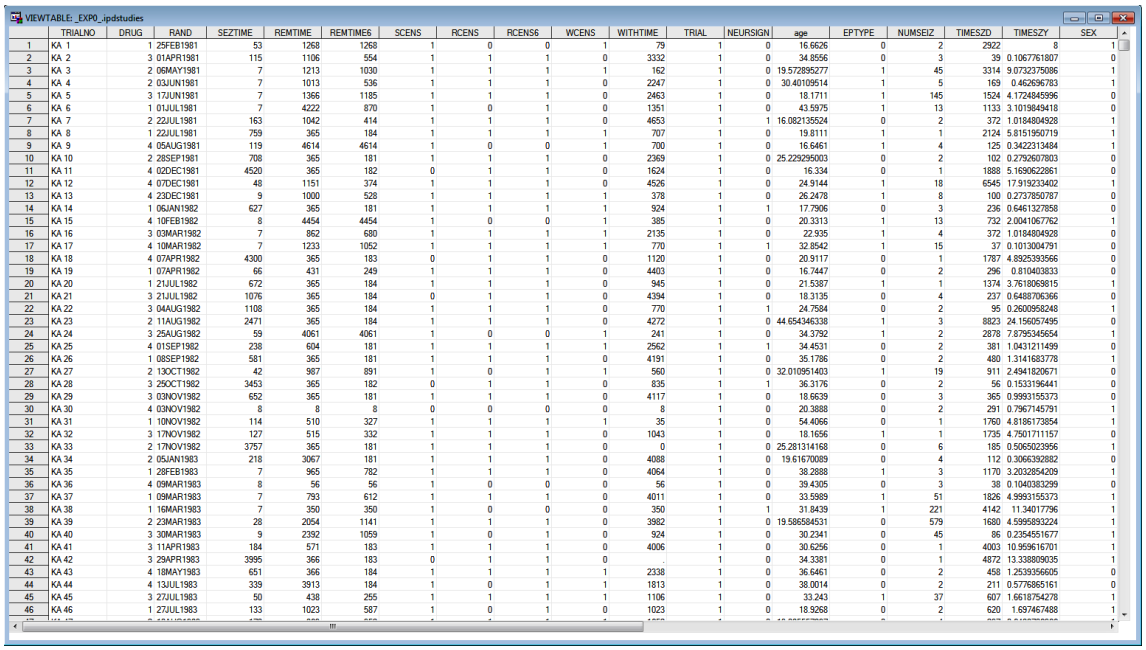
SCAN= MRI / CT results (1 = abnormal 0 = normal)

TIMESZD = Time from first (ever) seizure to nearest day to randomisation (calculated in a similar manner to 'time-to-first seizure, see Stage 4)

TIMESZY = Time from first (ever) seizure to nearest year to randomisation (calculated in a similar manner to 'time-to-first seizure, see Stage 4)

NUMSEIZ = Number of seizures 6 months prior to randomisation

Example of analysis ready data



VIEWTABLE: _EXP0_ipdstudies

	TRIALNO	DRUG	RAND	SEZTIME	REMTIME	REMTIME6	SCENS	RCENS	RCENS6	WCENS	WITHTIME	TRIAL	NEURSIGN	age	EPTYPE	NUMSEIZ	TIMESZD	TIMESZY	SEX
1	KA 1	1	25FEB1981	53	1268	1268	1	0	0	1	79	1	0	16.6626	0	2	2922	8	1
2	KA 2	3	01APR1981	115	1106	554	1	1	1	0	3332	1	0	34.8556	0	3	39	0.1067761807	0
3	KA 3	2	06MAY1981	7	1213	1030	1	1	1	1	162	1	0	19.572895277	1	45	3314	9.0732375086	1
4	KA 4	2	03JUN1981	7	1013	536	1	1	1	0	2247	1	0	30.40109514	1	5	169	0.462696783	1
5	KA 5	3	17JUN1981	7	1366	1185	1	1	1	0	2463	1	0	18.1711	1	145	1524	4.1724845996	0
6	KA 6	1	01JUL1981	7	4222	870	1	0	1	0	1351	1	0	43.5975	1	13	1133	3.1019849418	0
7	KA 7	2	22JUL1981	163	1042	414	1	1	1	0	4653	1	1	16.082135524	0	2	372	1.0184804928	1
8	KA 8	1	22JUL1981	759	365	184	1	1	1	1	707	1	0	19.8111	1	1	2124	5.8151950719	1
9	KA 9	4	05AUG1981	119	4614	4614	1	0	0	1	700	1	0	16.6461	1	4	125	0.3422319484	1
10	KA 10	2	28SEP1981	708	365	181	1	1	1	0	2369	1	0	25.229295003	0	2	102	0.2792607803	0
11	KA 11	4	02DEC1981	4520	365	182	0	1	1	0	1624	1	0	16.334	0	1	1888	5.1690622861	0
12	KA 12	4	07DEC1981	48	1151	374	1	1	1	0	4526	1	0	24.9144	1	18	6545	17.919233402	1
13	KA 13	4	23DEC1981	9	1000	528	1	1	1	1	378	1	0	26.2478	1	8	100	0.2737850787	0
14	KA 14	1	06JAN1982	627	365	181	1	1	1	1	924	1	1	17.7906	0	3	236	0.6461327858	0
15	KA 15	4	10FEB1982	8	4454	4454	1	0	0	1	385	1	0	20.3313	1	13	732	2.0041067762	1
16	KA 16	3	03MAR1982	7	862	680	1	1	1	1	2135	1	0	22.935	1	4	372	1.0184804928	0
17	KA 17	4	10MAR1982	7	1233	1052	1	1	1	1	770	1	1	32.8542	1	15	37	0.1013004791	0
18	KA 18	4	07APR1982	4300	365	183	0	1	1	0	1120	1	0	20.9117	0	1	1787	4.8925393566	0
19	KA 19	1	07APR1982	66	431	249	1	1	1	0	4403	1	0	16.7447	0	2	296	0.810403833	0
20	KA 20	1	21JUL1982	672	365	184	1	1	1	0	945	1	0	21.5387	1	1	1374	3.7618069815	1
21	KA 21	3	21JUL1982	1076	365	184	0	1	1	0	4394	1	0	18.3135	0	4	237	0.6488706366	0
22	KA 22	3	04AUG1982	1108	365	184	1	1	1	0	770	1	1	24.7584	0	2	95	0.2600958248	1
23	KA 23	2	11AUG1982	2471	365	184	1	1	1	0	4272	1	0	44.654346338	1	3	8823	24.156057495	0
24	KA 24	3	25AUG1982	59	4061	4061	1	0	0	1	241	1	0	34.3792	1	2	2878	7.8795345654	1
25	KA 25	4	01SEP1982	238	604	181	1	1	1	1	2562	1	1	34.4531	0	2	381	1.0431211499	0
26	KA 26	1	08SEP1982	581	365	181	1	1	1	0	4191	1	0	35.1786	0	2	480	1.3141683778	1
27	KA 27	2	13OCT1982	42	987	891	1	0	1	1	560	1	0	32.010951403	1	19	911	2.4941820671	0
28	KA 28	3	25OCT1982	3453	365	182	0	1	1	0	835	1	1	36.3176	0	2	56	0.1533196441	0
29	KA 29	3	03NOV1982	652	365	181	1	1	1	0	4117	1	0	18.6639	0	3	365	0.9993155373	0
30	KA 30	4	03NOV1982	8	8	8	0	0	0	0	8	1	0	20.3888	0	2	291	0.7967145791	1
31	KA 31	1	10NOV1982	114	510	327	1	1	1	1	35	1	0	54.4066	0	1	1760	4.8186173854	1
32	KA 32	3	17NOV1982	127	515	332	1	1	1	0	1043	1	0	18.1656	1	1	1735	4.7501711157	0
33	KA 33	2	17NOV1982	3757	365	181	1	1	1	0	0	1	0	25.281314168	0	6	185	0.5065023956	1
34	KA 34	2	05JAN1983	218	3067	181	1	1	1	0	4088	1	0	19.61670089	0	4	112	0.3066392882	0
35	KA 35	1	28FEB1983	7	965	782	1	1	1	0	4064	1	0	38.2888	1	3	1170	3.2032854209	1
36	KA 36	4	09MAR1983	8	56	56	1	0	0	0	56	1	0	39.4305	0	3	38	0.1040383299	0
37	KA 37	1	09MAR1983	7	793	612	1	1	1	0	4011	1	0	33.5989	1	51	1826	4.9993155373	1
38	KA 38	1	16MAR1983	7	350	350	1	0	0	0	350	1	1	31.8439	1	221	4142	11.34017796	1
39	KA 39	2	23MAR1983	28	2054	1141	1	1	1	0	3982	1	0	19.586584531	0	579	1680	4.5995893224	1
40	KA 40	3	30MAR1983	9	2392	1059	1	0	1	0	924	1	0	30.2341	0	45	86	0.2354551677	1
41	KA 41	3	11APR1983	184	571	183	1	1	1	0	4006	1	0	30.6256	0	1	4003	10.959616701	1
42	KA 42	3	29APR1983	3995	366	183	0	1	1	0	.	1	0	34.3381	0	1	4872	13.338809035	1
43	KA 43	4	18MAY1983	651	366	184	1	1	1	1	2338	1	0	36.6461	0	2	458	1.2539356605	0
44	KA 44	4	13JUL1983	339	3913	184	1	0	1	1	1813	1	0	38.0014	0	2	211	0.5776865161	0
45	KA 45	3	27JUL1983	50	438	255	1	1	1	1	1106	1	0	33.243	1	37	607	1.6618754278	1
46	KA 46	1	27JUL1983	133	1023	587	1	0	1	0	1023	1	0	18.9268	0	2	620	1.697467488	1

Screenshot taken from SAS 9.3 of analysis ready data for Cochrane Epilepsy Individual Participant Data Network Meta-Analysis

Appendix 13: Example data structure for NMA in Chapter 6

Method described in Chapter 6 and Chapter 8 reduce IPD to summary statistics, with the overall objective of being able to incorporate aggregate data with IPD in NMA. Summary statistics estimated from IPD from each study were then combined in separate NMAs by epilepsy type as if they were aggregate data.

Table 44: Example structure of a dataset in 'pairs' format in Stata version 14.

Id	loghr	seloghr	drug1	drug2	eptype
trial 1	0.651982	0.467607	CBZ	PHB	partial
trial 1	0.235209	0.449843	CBZ	PHB	generalised
trial 1	0.064524	0.490908	CBZ	PHT	partial
trial 1	-1.21468	0.667371	CBZ	PHT	generalised
trial 1	-0.05105	0.519561	CBZ	VPS	partial
trial 1	0.140863	0.451576	CBZ	VPS	generalised
trial 1	-0.58746	0.450693	PHB	PHT	partial
trial 1	-1.44989	0.651693	PHB	PHT	generalised
trial 1	-0.70303	0.484911	PHB	VPS	partial
trial 1	-0.09435	0.427372	PHB	VPS	generalised
trial 1	-0.11558	0.505825	PHT	VPS	partial
trial 1	1.355542	0.651716	PHT	VPS	generalised
trial 2	1.711753	0.60801	CBZ	PHB	partial
trial 2	0.794814	0.592827	CBZ	PHB	generalised
trial 2	0.241746	0.425658	CBZ	PHT	partial
trial 2	-0.58901	0.548477	CBZ	PHT	generalised
trial 2	0.227631	0.460312	CBZ	VPS	partial
trial 2	-0.02997	0.479234	CBZ	VPS	generalised
trial 2	-1.47001	0.590079	PHB	PHT	partial
trial 2	-1.38382	0.673133	PHB	PHT	generalised
trial 2	-1.48412	0.610886	PHB	VPS	partial
trial 2	-0.82478	0.62	PHB	VPS	generalised
trial 2	-0.01412	0.442092	PHT	VPS	partial
trial 2	0.559036	0.573396	PHT	VPS	generalised
trial 3	0.450873	0.177159	CBZ	PHB	partial
trial 3	0.183851	0.179511	CBZ	PHT	partial
trial 3	-0.26702	0.162971	PHB	PHT	partial
trial 4	-0.03332	0.154534	CBZ	VPS	partial
trial 5	0.190944	0.311932	CBZ	VPS	partial
trial 5	-0.61638	0.353263	CBZ	VPS	generalised
trial 6	-0.02297	0.353657	CBZ	VPS	partial
trial 6	0.306479	0.386035	CBZ	VPS	generalised

This was achieved by producing a dataset of the summary statistics structured as a list of pairwise comparisons and converted from 'pairs' to 'augmented' format via the 'network' command within Stata version 14 [310]. In other words, the following code was applied to the data in Table 44 which produced a dataset of the format in Table 45:

network import, tr(drug1 drug2) eff(loghr) study(id) stderr(seloghr)

network convert augmented

Table 45: Example structure of a dataset in 'augmented' format in Stata version 14.

id	eptype	logvar	_design	loghr_2	loghr_3	loghr_4	_S_2_2	_S_2_3	_S_2_4	_S_3_3	_S_3_4	_S_4_4
trial 1	partial	0.255859	1 2 3 4	0.651982	0.064524	-0.05105	0.218656	0.128262	0.126731	0.240991	0.127538	0.269944
trial 1	generalised	0.424734	1 2 3 4	0.235209	-1.21468	0.140863	0.202359	0.111519	0.111816	0.445384	0.112285	0.203921
trial 2	partial	0.195445	1 2 3 4	1.711753	0.241746	0.22763	0.369676	0.101334	0.104191	0.181185	0.098814	0.211887
trial 2	generalised	0.328783	1 2 3 4	0.794813	-0.58901	-0.02997	0.351444	0.099581	0.098355	0.300827	0.100854	0.229665
trial 3	partial	0.02656	1 2 3	0.450873	0.183851		0.031385	0.018525		0.032224		
trial 4	partial	0.023881	1 4			-0.03332						0.023881
trial 5	partial	0.097302	1 4			0.190944						0.097302
trial 5	generalised	0.124795	1 4			-0.61638						0.124795
trial 6	partial	0.125073	1 4			-0.02297						0.125073
trial 6	generalised	0.149023	1 4			0.306479						0.149023

Subsequently, NMA was performed via the 'network meta' command.

Appendix 14: Additional analysis and sensitivity analyses of IPD-NMA performed to account for inconsistencies in IPD (Chapter 7)

The following sensitivity analyses were performed where minor inconsistencies were identified when preparing IPD for analysis (see Chapter 5.3.3 and Appendix 12). See Chapter 7.1.3 for full details of other additional and sensitivity analyses performed.

- In Stephen 2007, there were minor inconsistencies between rates of seizure recurrence and reasons for withdrawal between the data provided and the publication, which could not be resolved with the original trial authors. Therefore, a sensitivity analysis was performed excluding Stephen 2007 from all analyses.
- In Reunanen 1996, participants were considered to have completed the trial and hence treatment was withdrawn if they experienced a seizure after week six. This does not correspond with the treatment withdrawal definition used in this review and analysis (see Chapter 5.3.3.2). Therefore a sensitivity analysis was performed excluding Reunanen 1996 for the analysis of 'Time-to-withdrawal of allocated treatment.'
- In Banu 2007, there were minor inconsistencies between rates of seizure recurrence between the data provided and the published paper, which could not be resolved. Therefore, a sensitivity analysis was performed excluding Banu 2007 from analysis of 'Time-to-first seizure.' (Data provided was insufficient to contribute to outcomes time-to-6-month remission and time-to-12-month remission, see Chapter 5.4.3).
- In Nieto-Barrera 2001, seizures that occurred during the first four weeks of the trial were not included in efficacy analyses and dates of seizures before week four were not provided. Therefore, 'time-to-first seizure' was calculated after week four rather than after randomisation and a sensitivity analysis was performed excluding Nieto-Barrera 2001 from analysis of 'time-to-first seizure' (this trial was 24 weeks duration so did not contribute to outcomes time-to-6-month remission and time-to-12-month remission).
- In Placencia 1993, there were minor inconsistencies between reasons for withdrawal between the data provided and the published paper. In the primary analysis, withdrawals were classified according to the reasons provided in IPD and a sensitivity analysis was performed for analysis of 'time-to-withdrawal of allocated treatment' with withdrawals reclassified according to definitions from the published paper.

Results of additional and sensitivity analyses

Figure 19: Time-to-withdrawal of allocated treatment and time-to-first seizure adjusted for age (partial seizures)

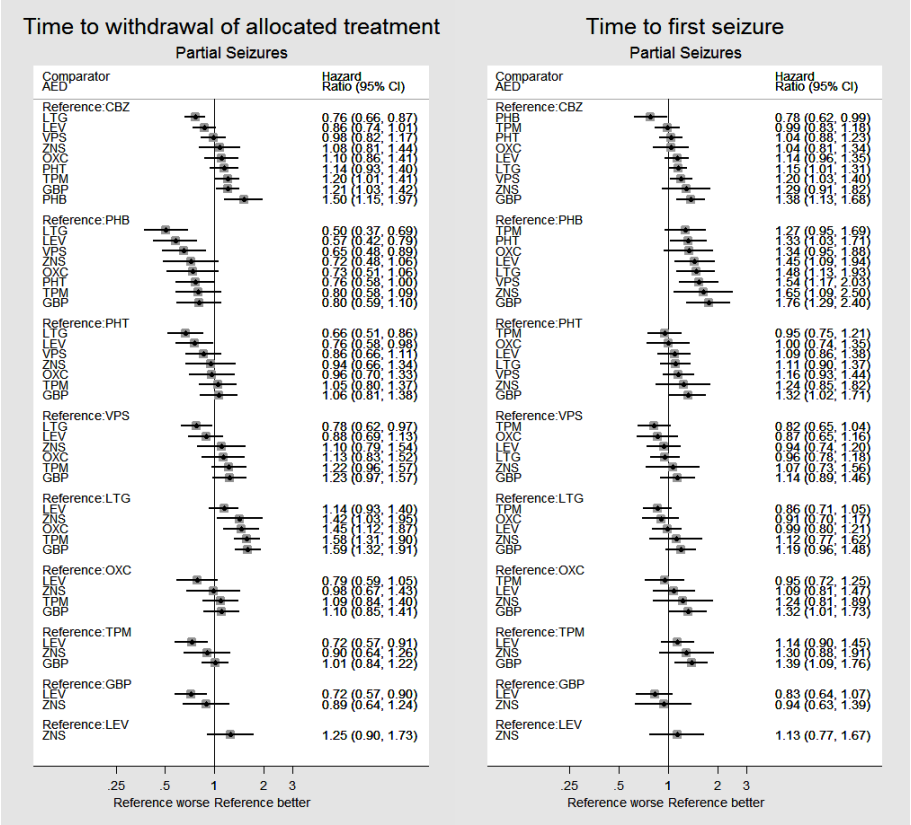


Figure 20: Time-to-withdrawal of allocated treatment and time-to-first seizure adjusted for age (generalised seizures)

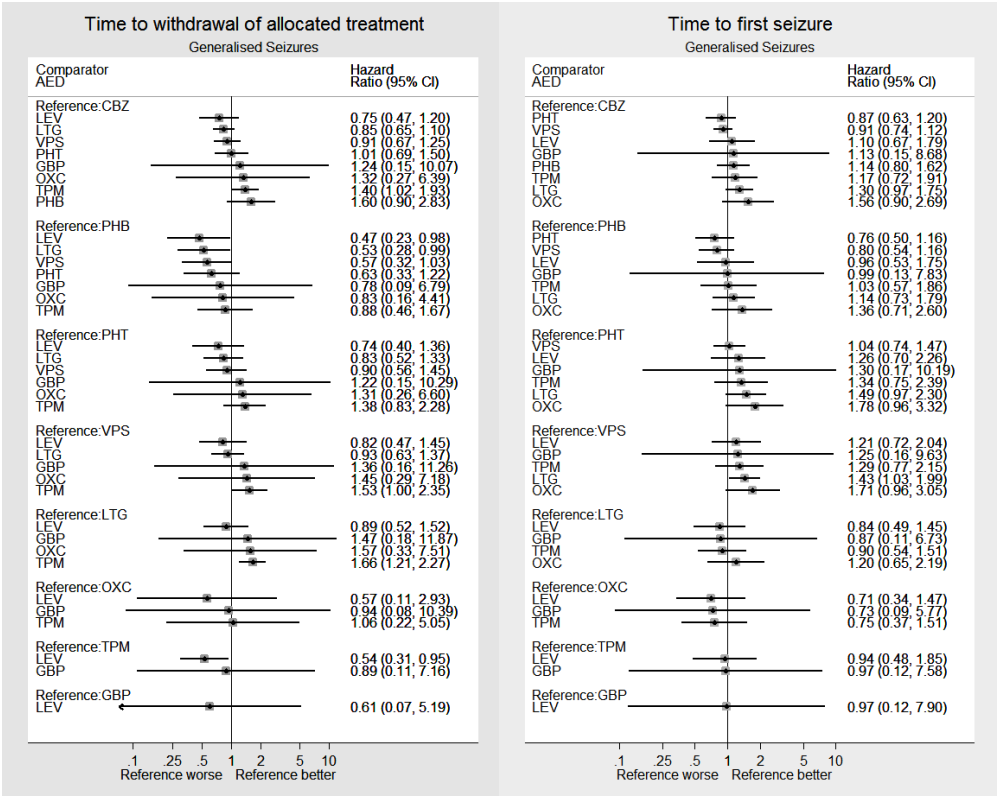


Figure 21: Time-to-withdrawal of allocated treatment using parametric AFT model (both seizure types)

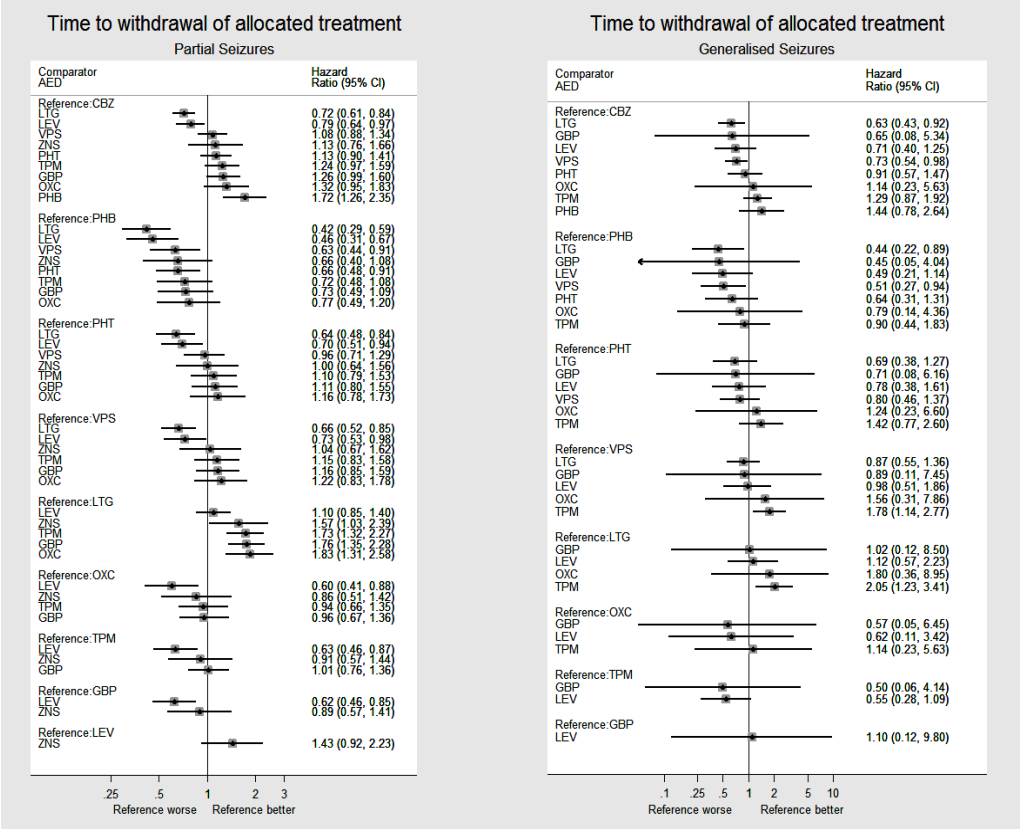


Figure 22: Time-to-first seizure using parametric AFT model (both seizure types)

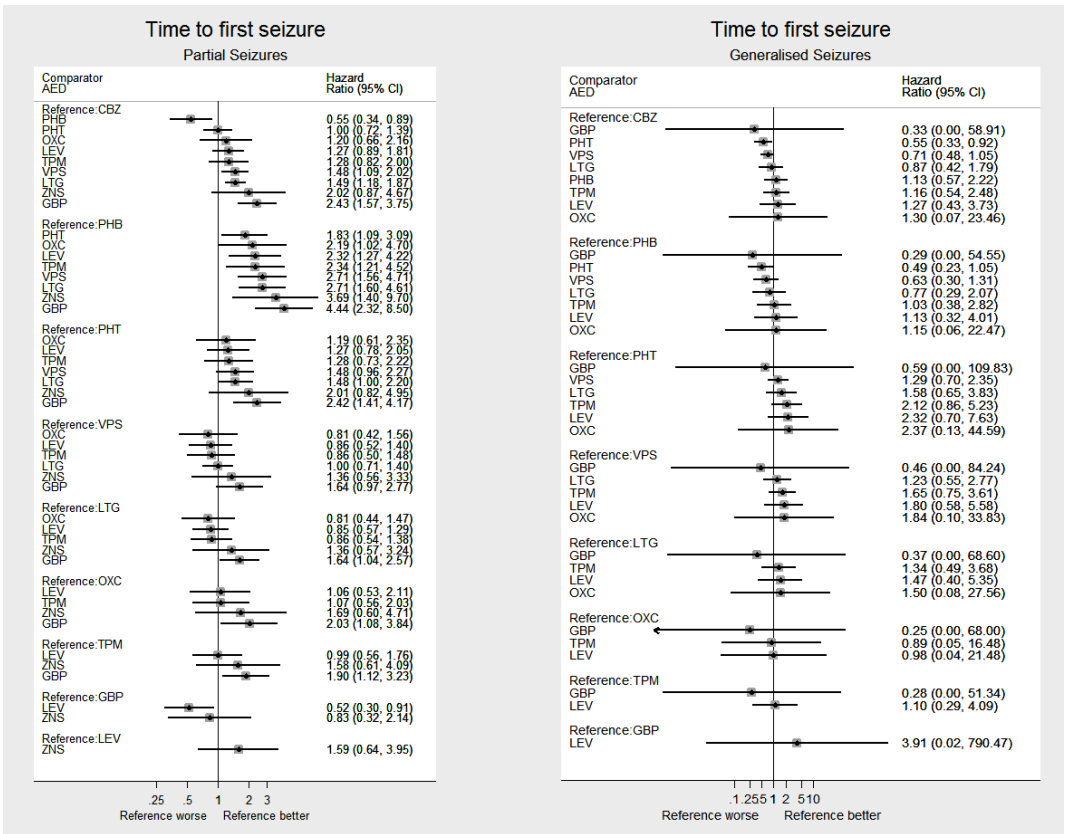


Figure 23: Time-to-withdrawal of allocated treatment and time-to-first seizure, Stephen 2007 excluded (partial seizures)

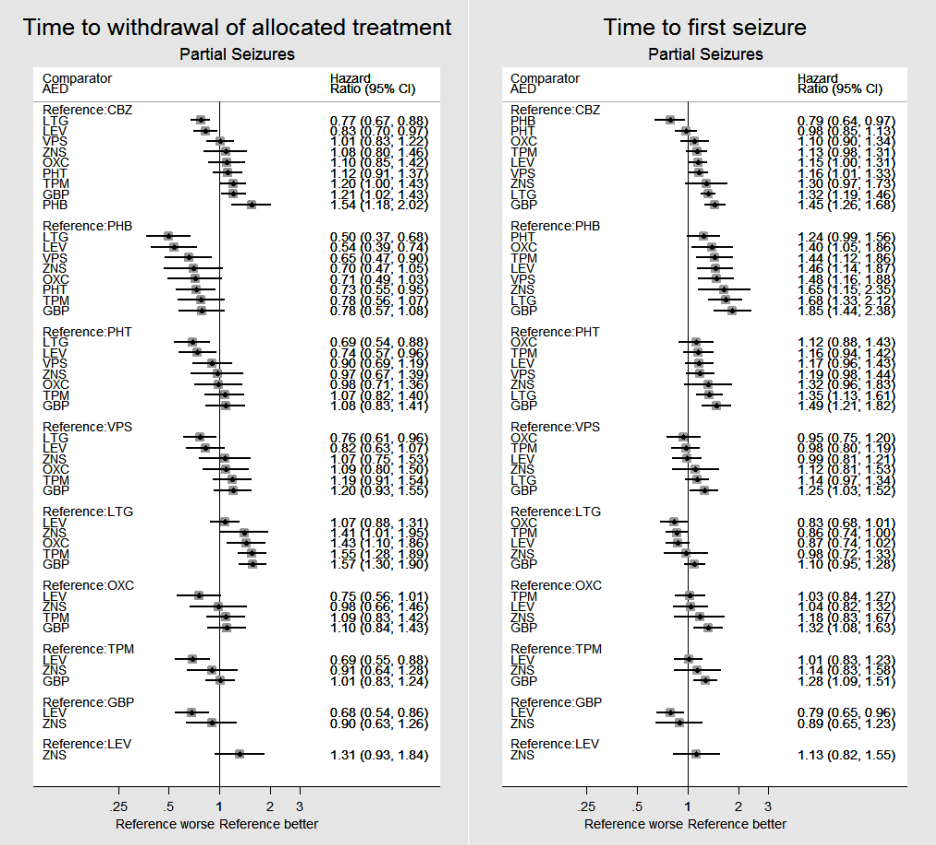


Figure 24: Time-to-withdrawal of allocated treatment and time-to-first seizure, Stephen 2007 excluded (generalised seizures)

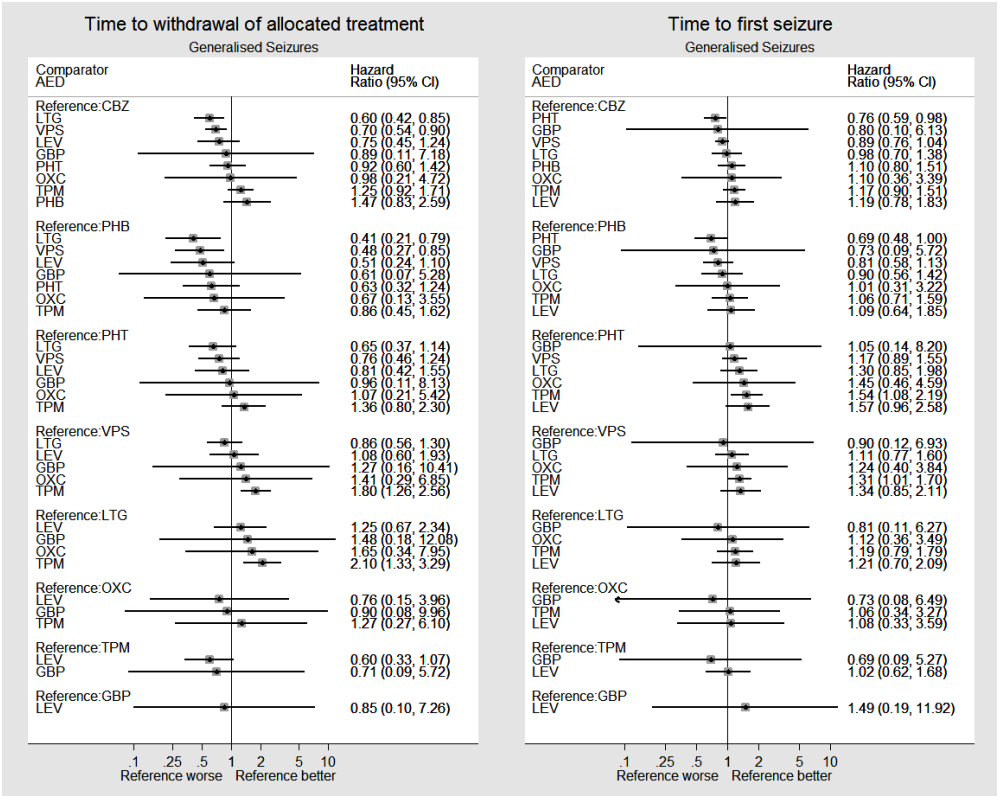


Figure 25: Time-to-withdrawal of allocated treatment, Reunanen 1996 excluded (both seizure types)

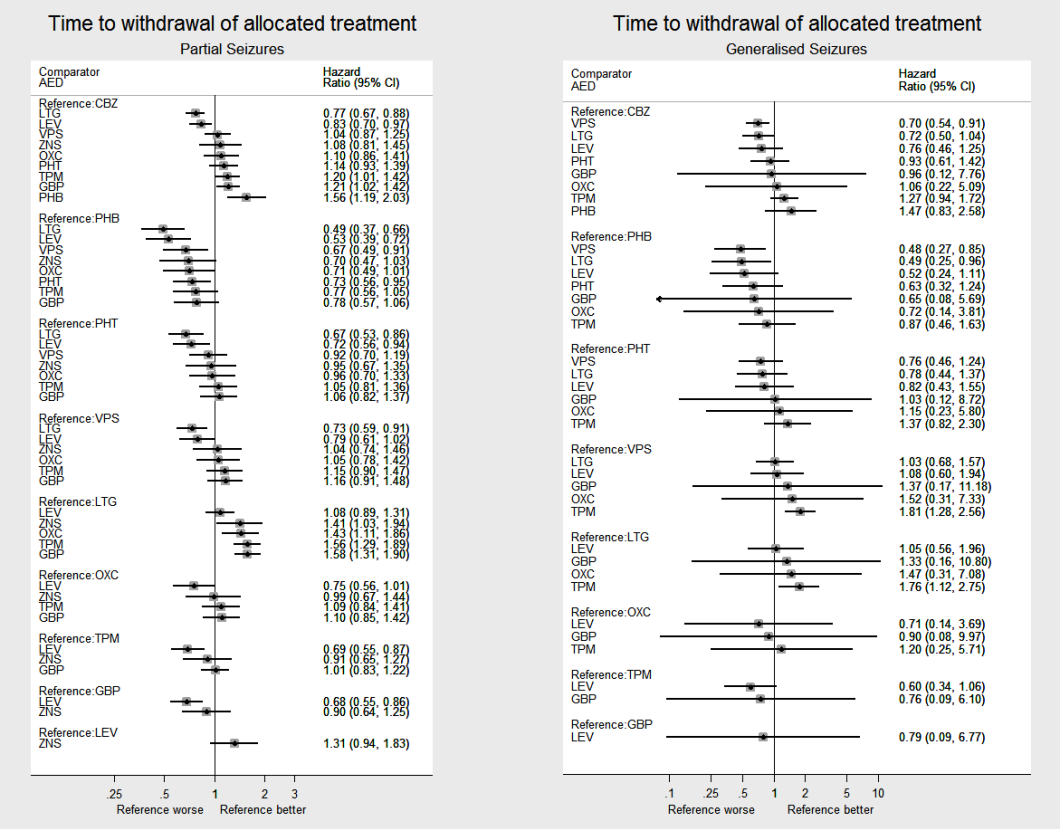


Figure 26: Time-to-withdrawal of allocated treatment, Placencia 1996 excluded (both seizure types)

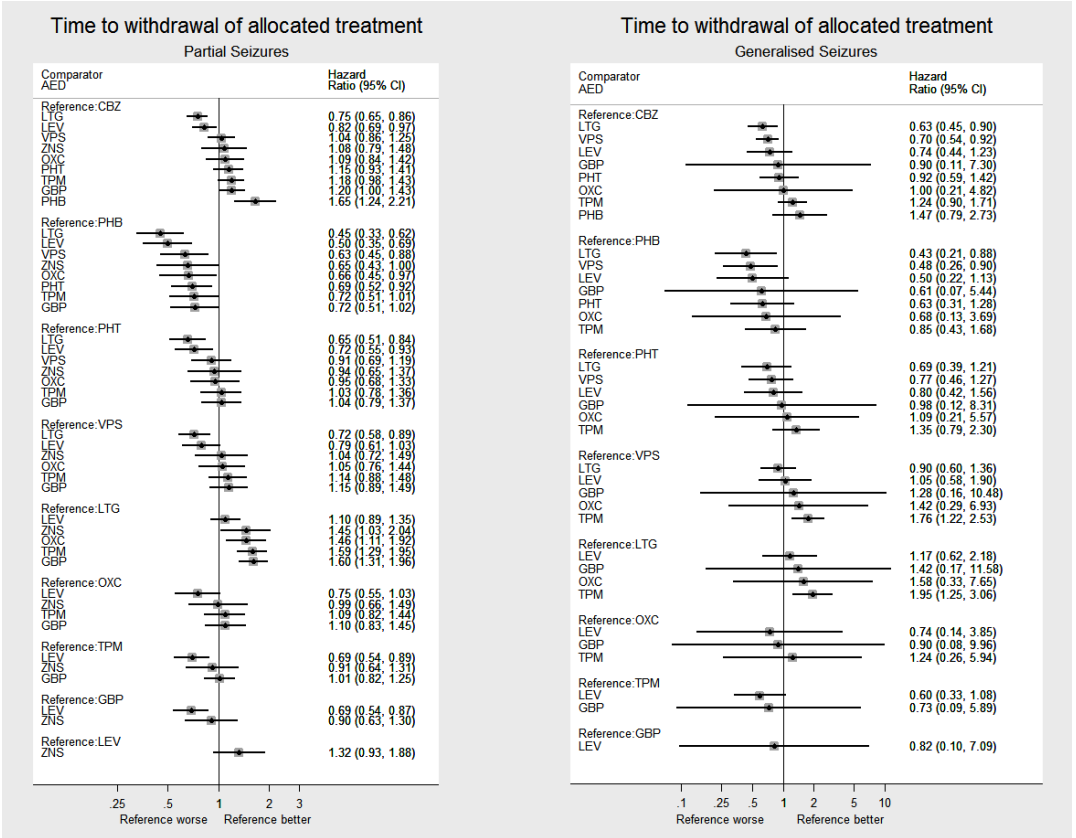


Figure 27: Time-to-first seizure, Banu 2007 excluded (both seizure types)

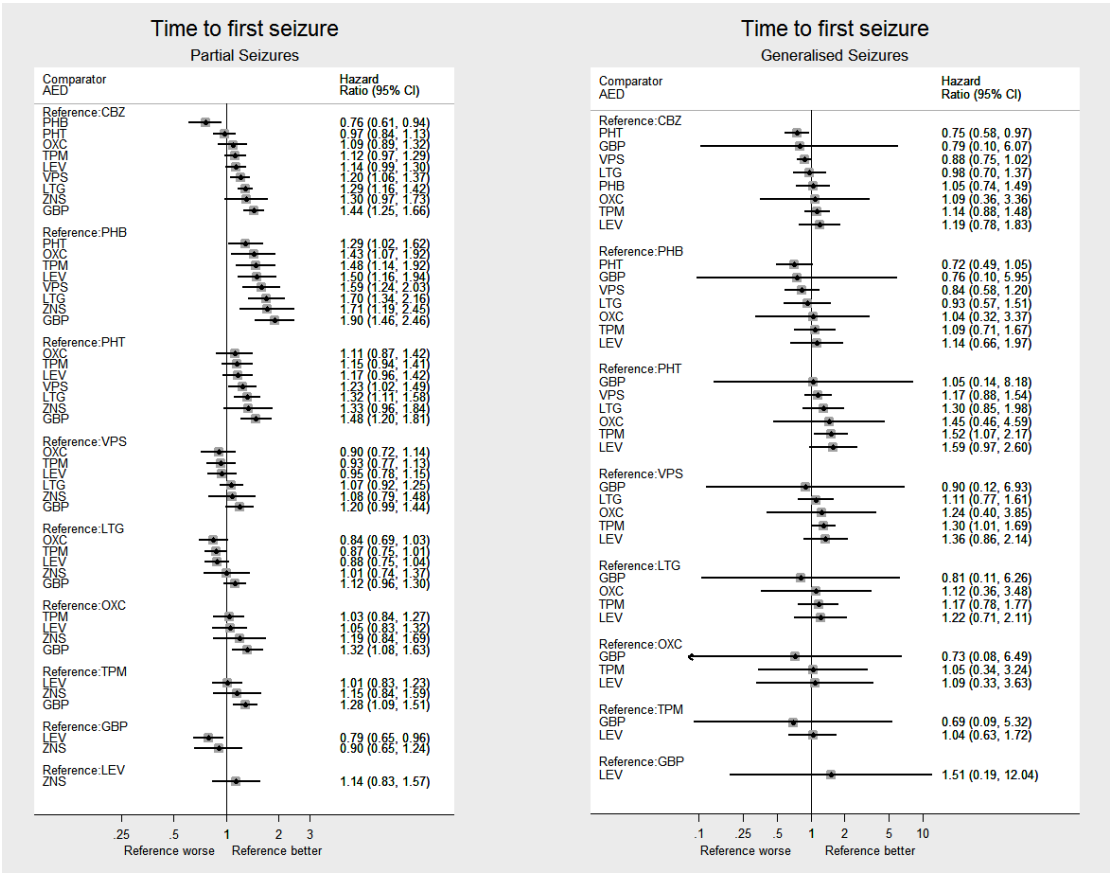


Figure 28: Time-to-first seizure, Nieto-Barrera 2001 excluded (both seizure types)

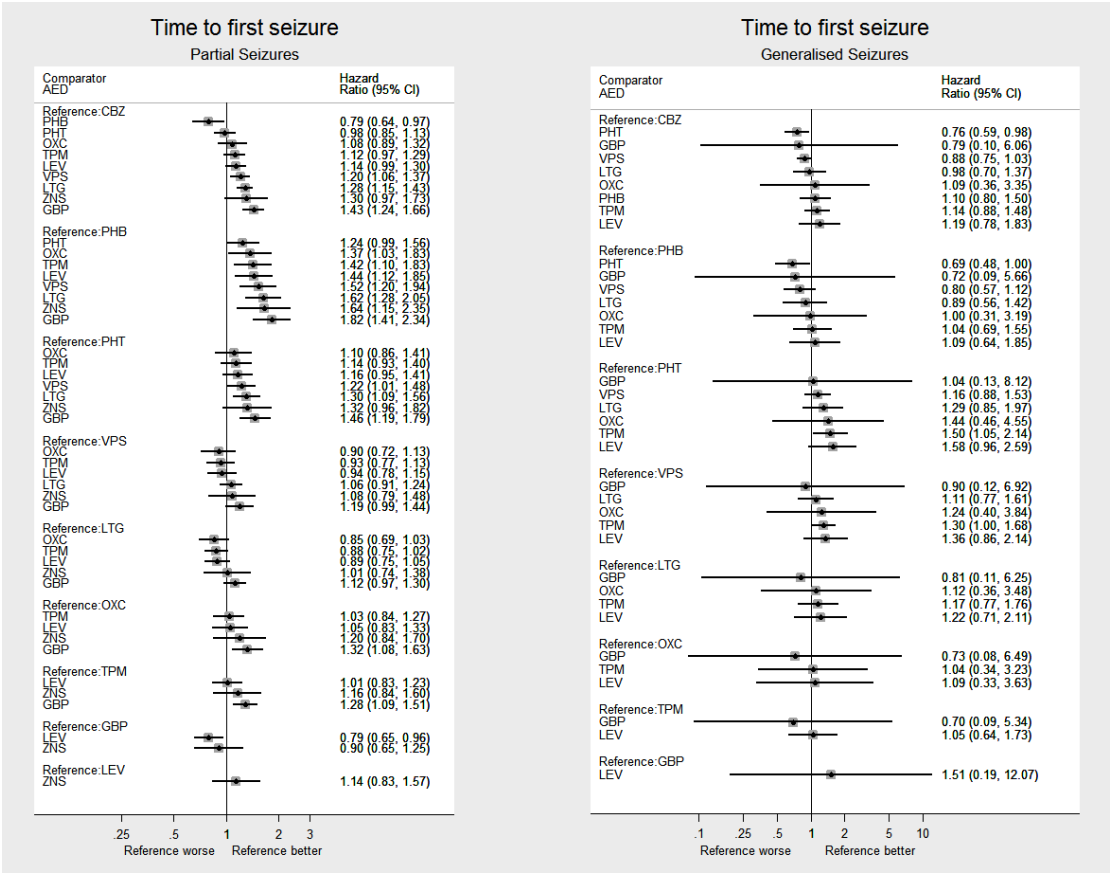


Figure 29: Time-to-withdrawal of allocated treatment and time-to-first seizure, seizure type reclassification 1 (see Notes below, results for individuals with partial seizures)

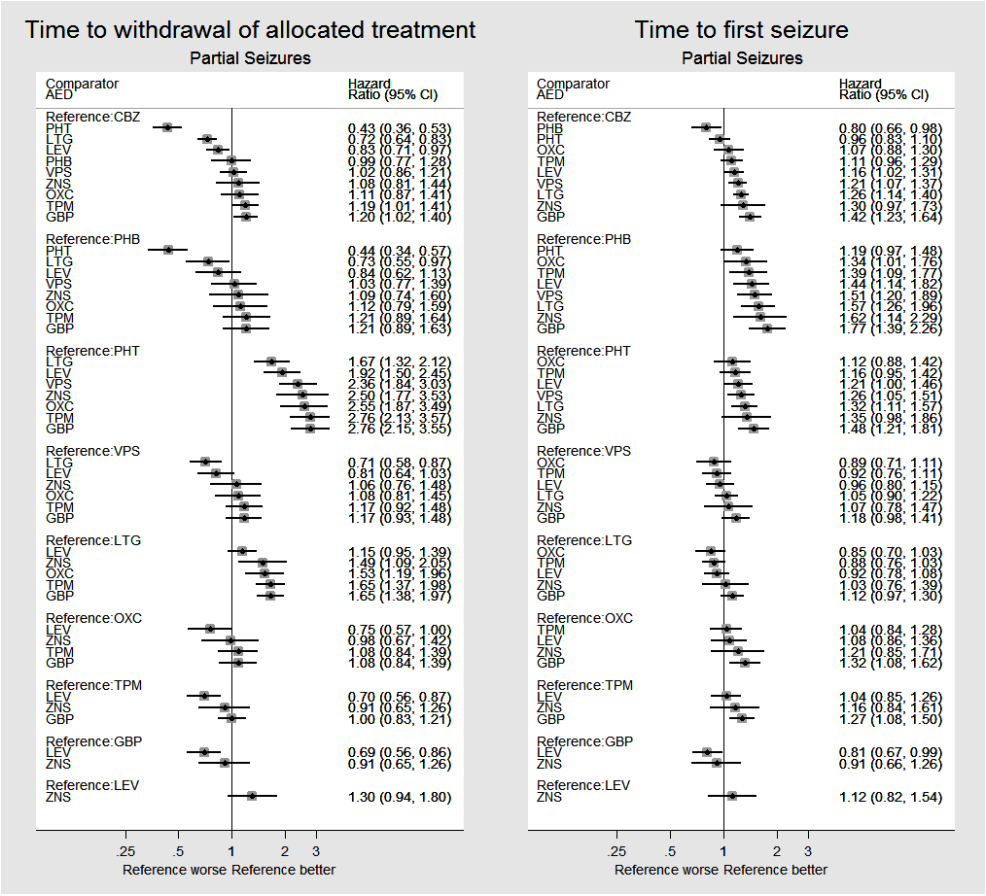


Figure 30: Time-to-withdrawal of allocated treatment and time-to-first seizure, seizure type reclassification 1 (see Notes below, results for individuals with generalised seizures)

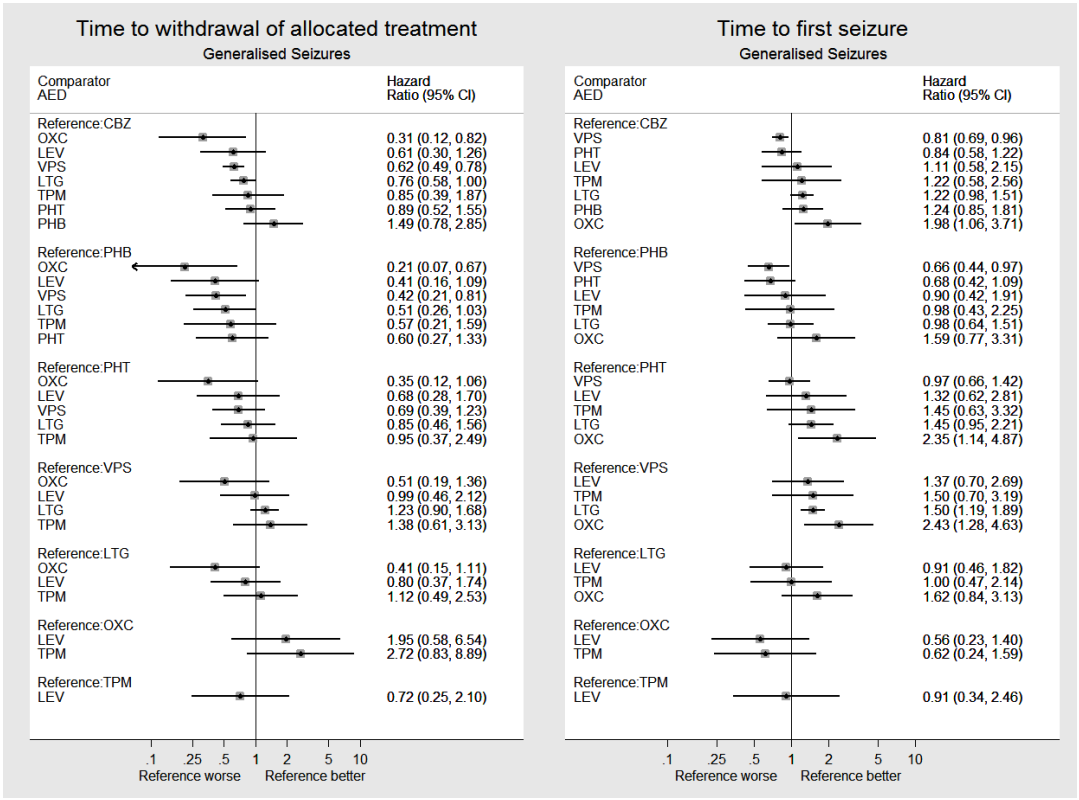


Figure 31: Time-to-withdrawal of allocated treatment and time-to-first seizure, seizure type reclassification 2 (see Notes below, results for individuals with partial seizures)

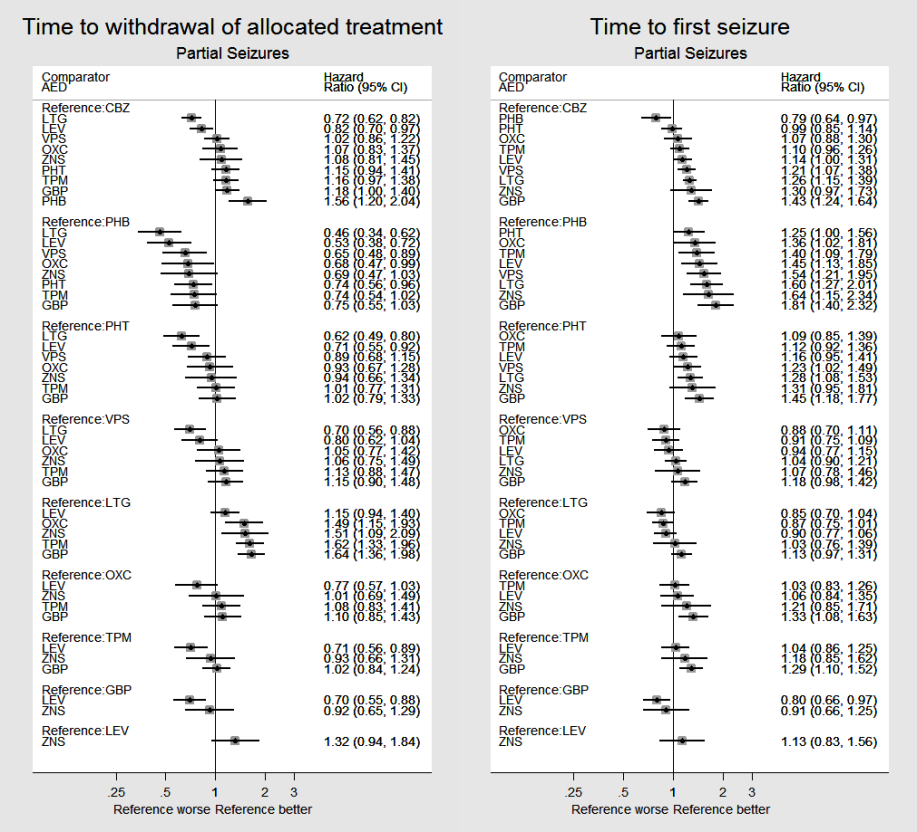


Figure 32: Time-to-withdrawal of allocated treatment and time-to-first seizure, seizure type reclassification 2 (see Notes below, results for individuals with generalised seizures)

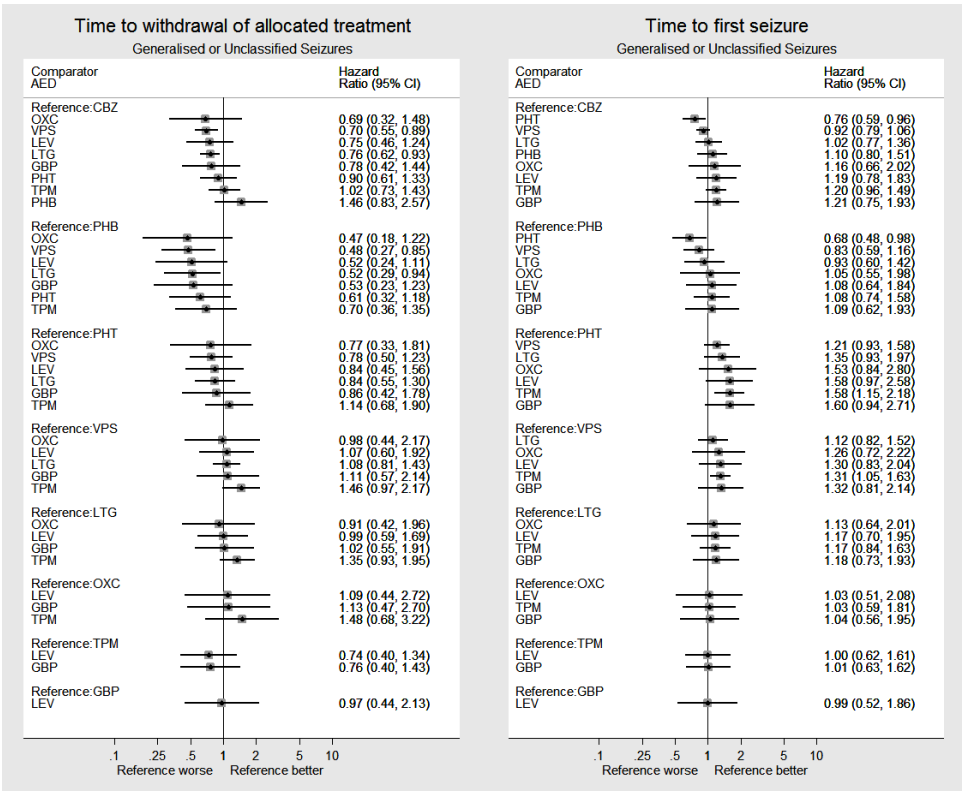
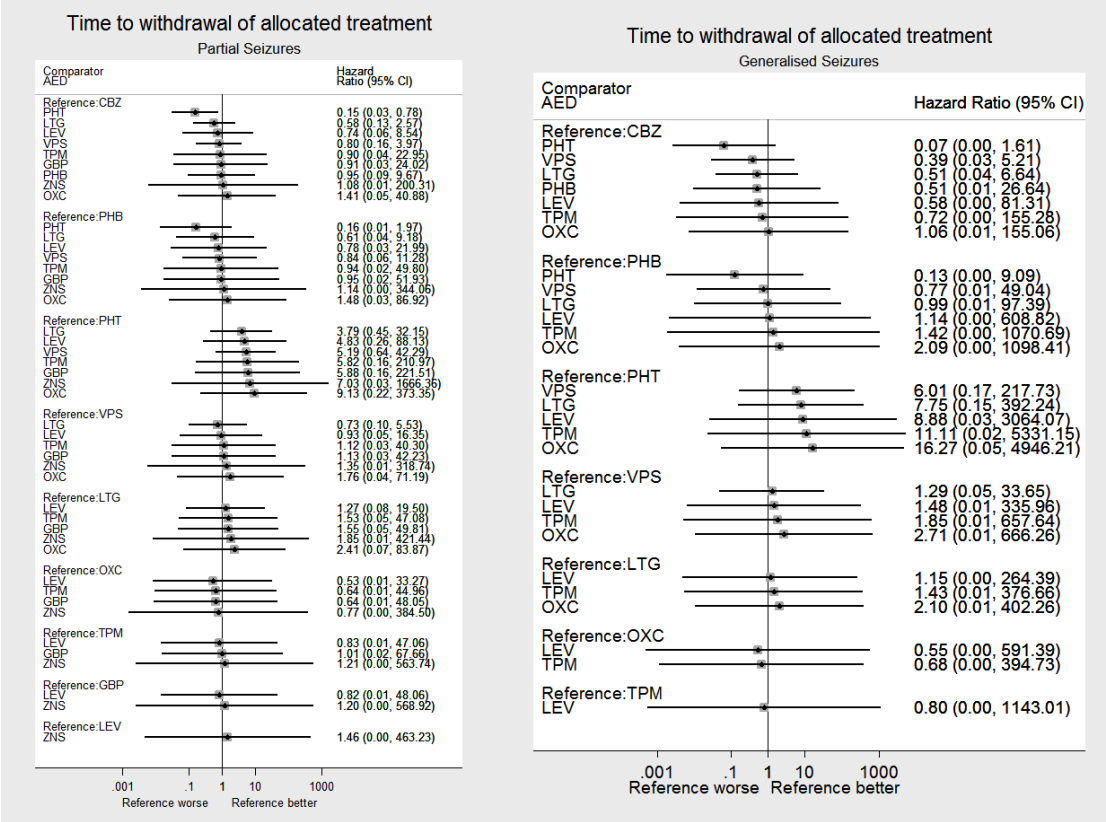


Figure 33: Time-to-withdrawal of allocated treatment, seizure type reclassification 1 with random-effects (see Notes below, both seizure types)



Notes

See Appendix 10 for references of the trials included in the Cochrane IPD-NMA and Chapter 5.2.1.3 for abbreviations of drugs.

Results shown on all figures are those from NMA (direct and indirect evidence combined).

Generalised tonic-clonic seizures with or without other seizure types is shortened to 'Generalised seizures' for brevity

Reclassification 1: Re-classification of 1,164 individuals with generalised seizures and age of onset greater than 30 years as having partial-onset seizures.

Reclassification 2: Re-classification of 1,164 individuals with generalised seizure and age at onset greater than 30 years and 574 individuals with missing seizure type to 'unclassified epilepsy type'

Appendix 15: Additional results of the IPD-NMA: time to 12 month and time to 6 month remission

Figure 34: Time to 12 month remission: All IPD-NMA results by epilepsy type

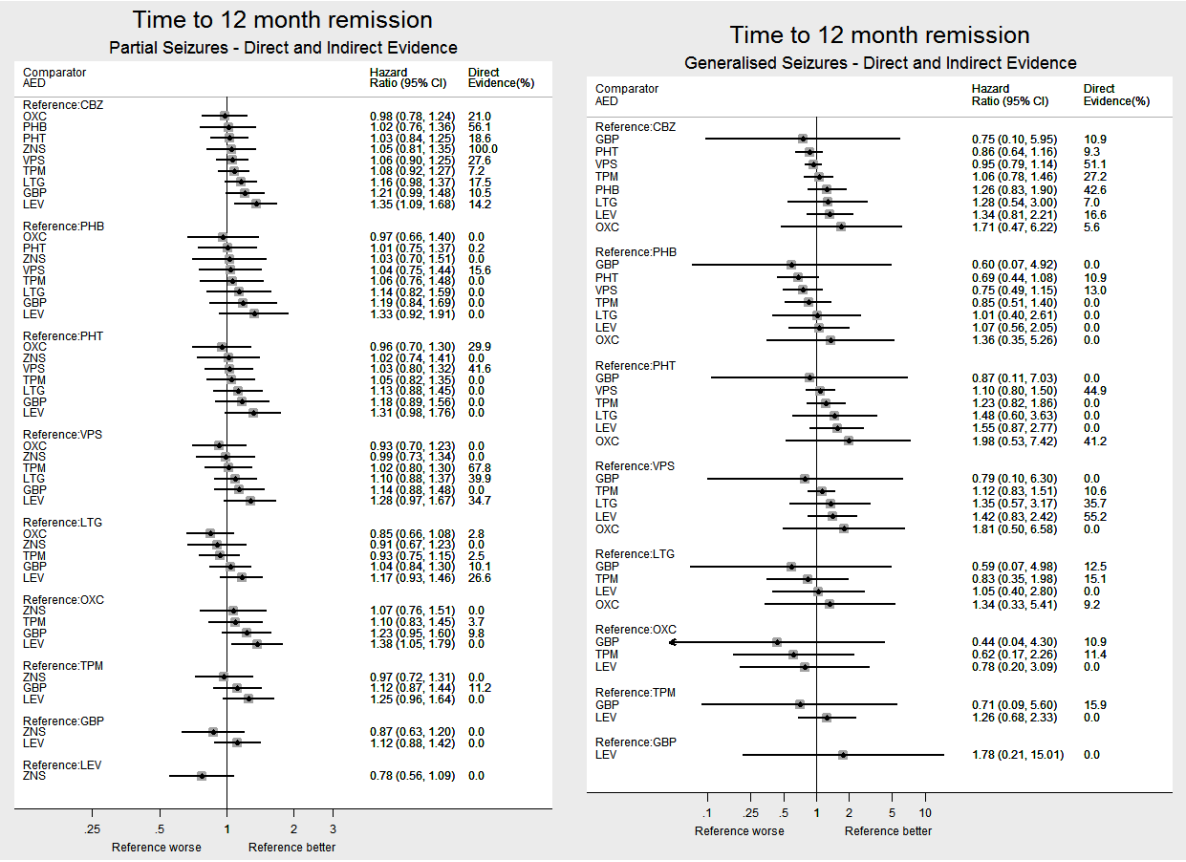
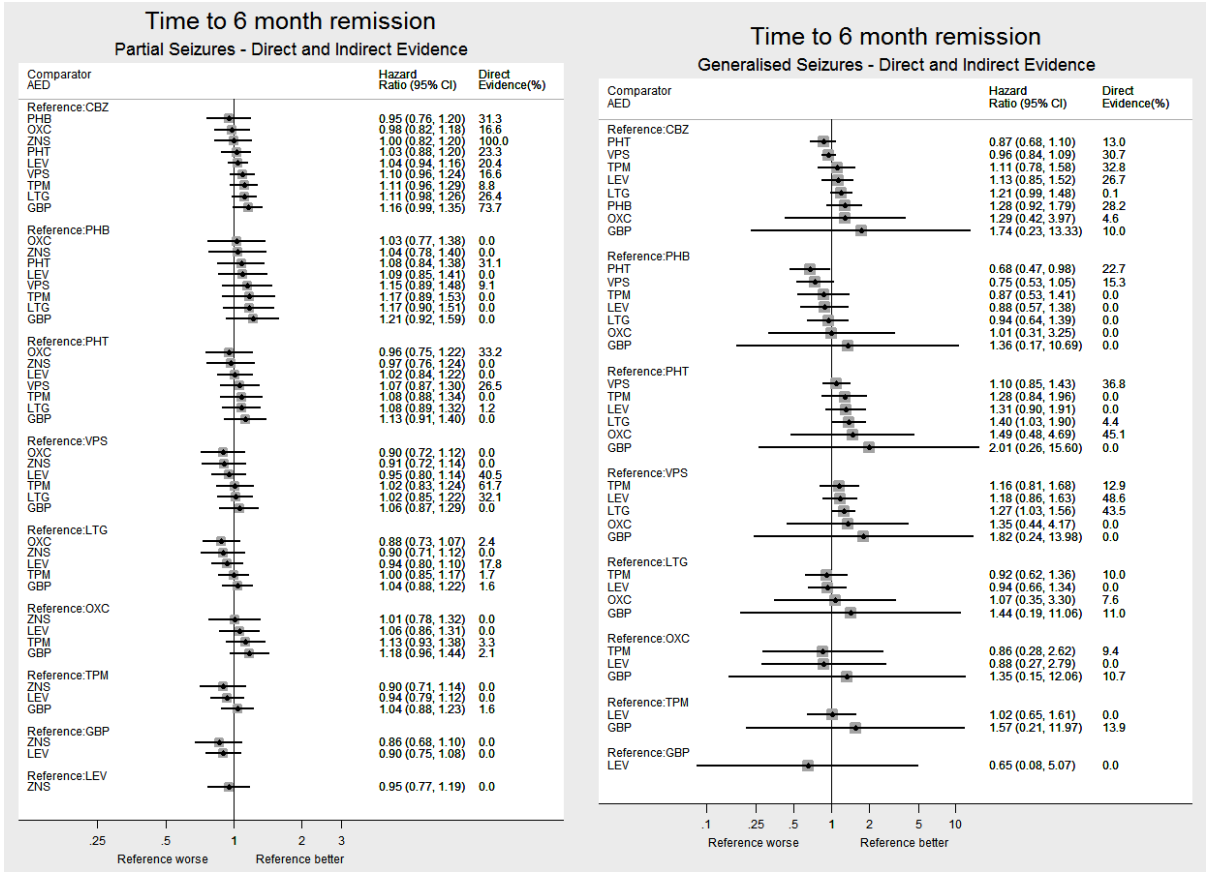


Figure 35: Time to 6 month remission: All IPD-NMA results by epilepsy type



Appendix 16: Investigation of inconsistency in Chapter 7

All figures present direct evidence (from pairwise meta-analysis), indirect evidence (from node-splitting) and NMA results.

Figure 36: Time-to-withdrawal of allocated treatment and time-to-first seizure (carbamazepine (CBZ) reference) for individuals with generalised seizures

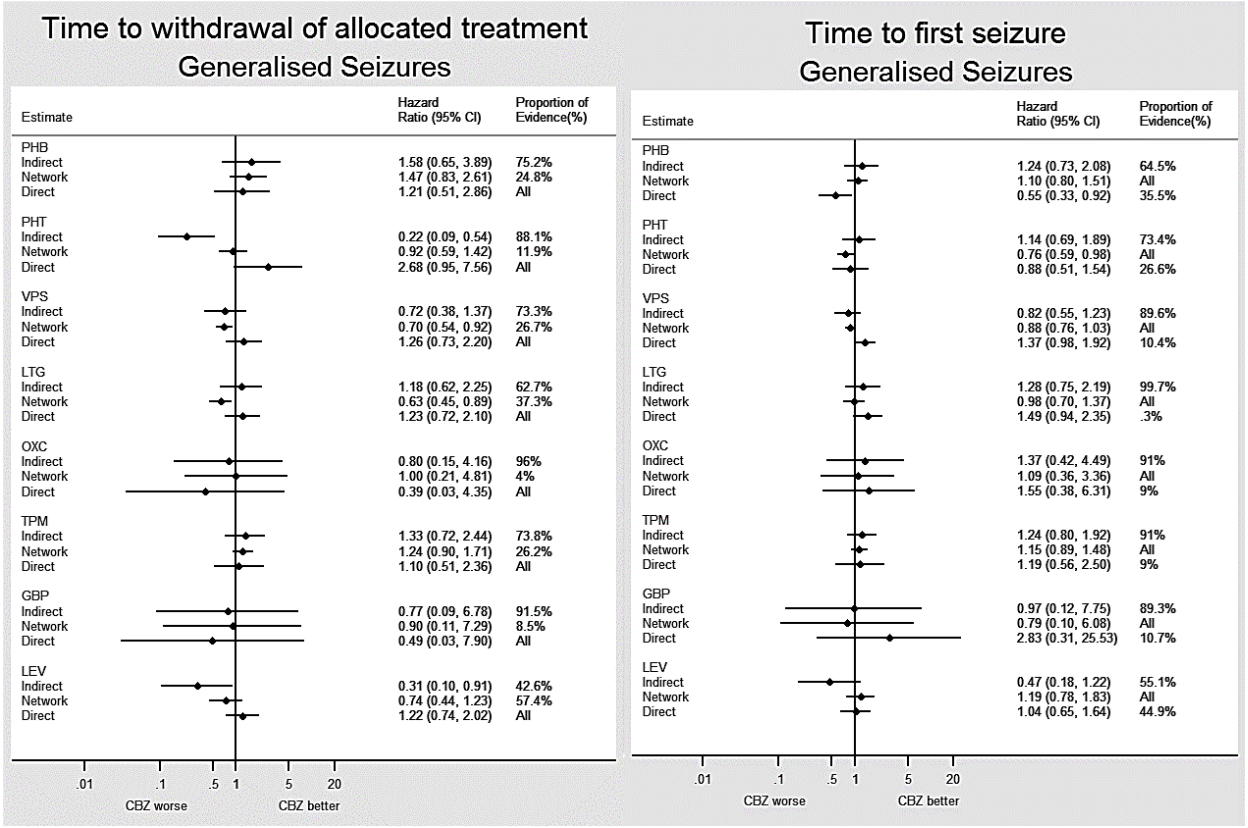


Figure 37: Time-to-withdrawal of allocated treatment and time-to-first seizure (gabapentin (GBP) reference) for individuals with generalised seizures

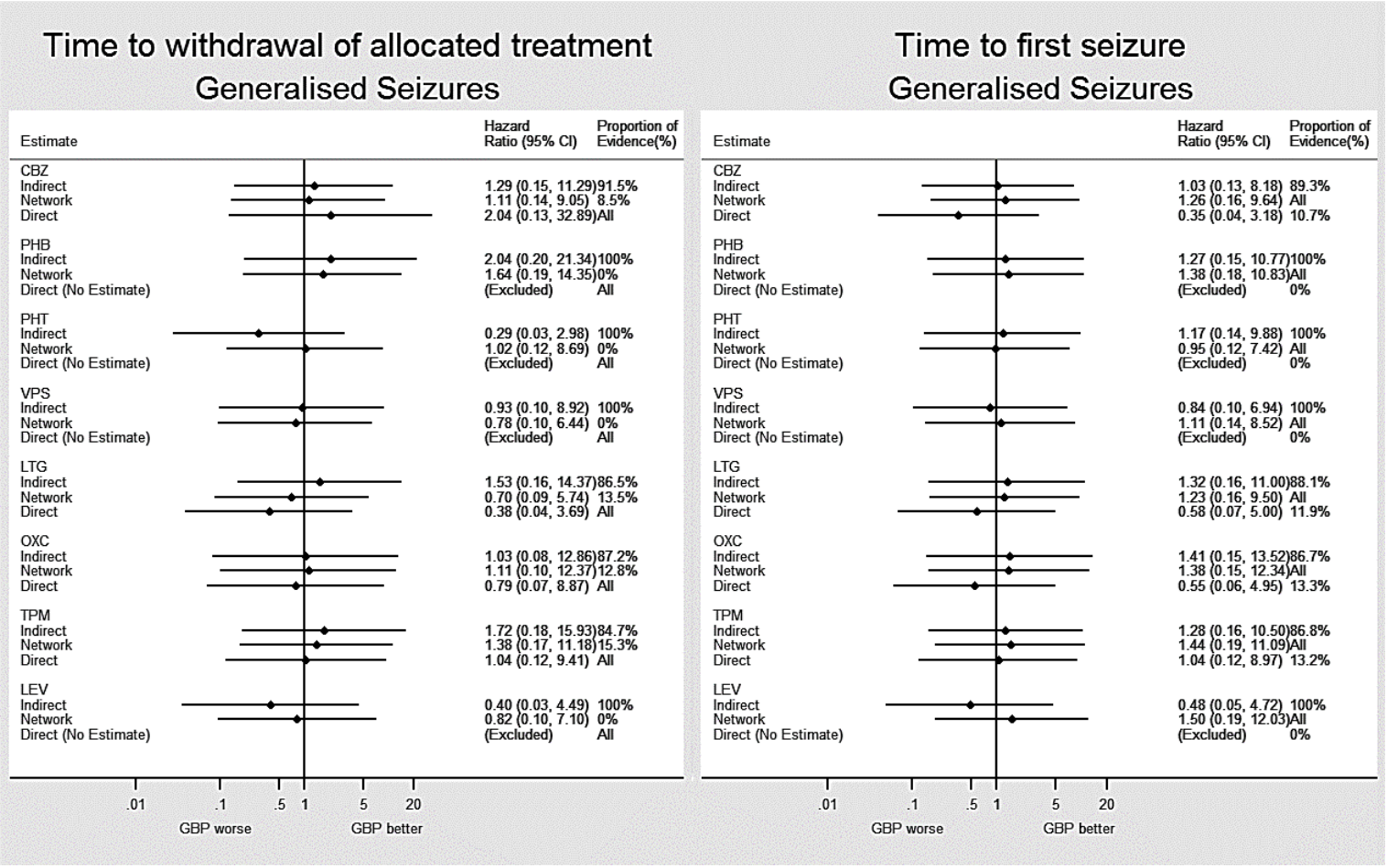


Figure 38: Time-to-withdrawal of allocated treatment and time-to-first seizure (levetiracetam (LEV) reference) for individuals with generalised seizures

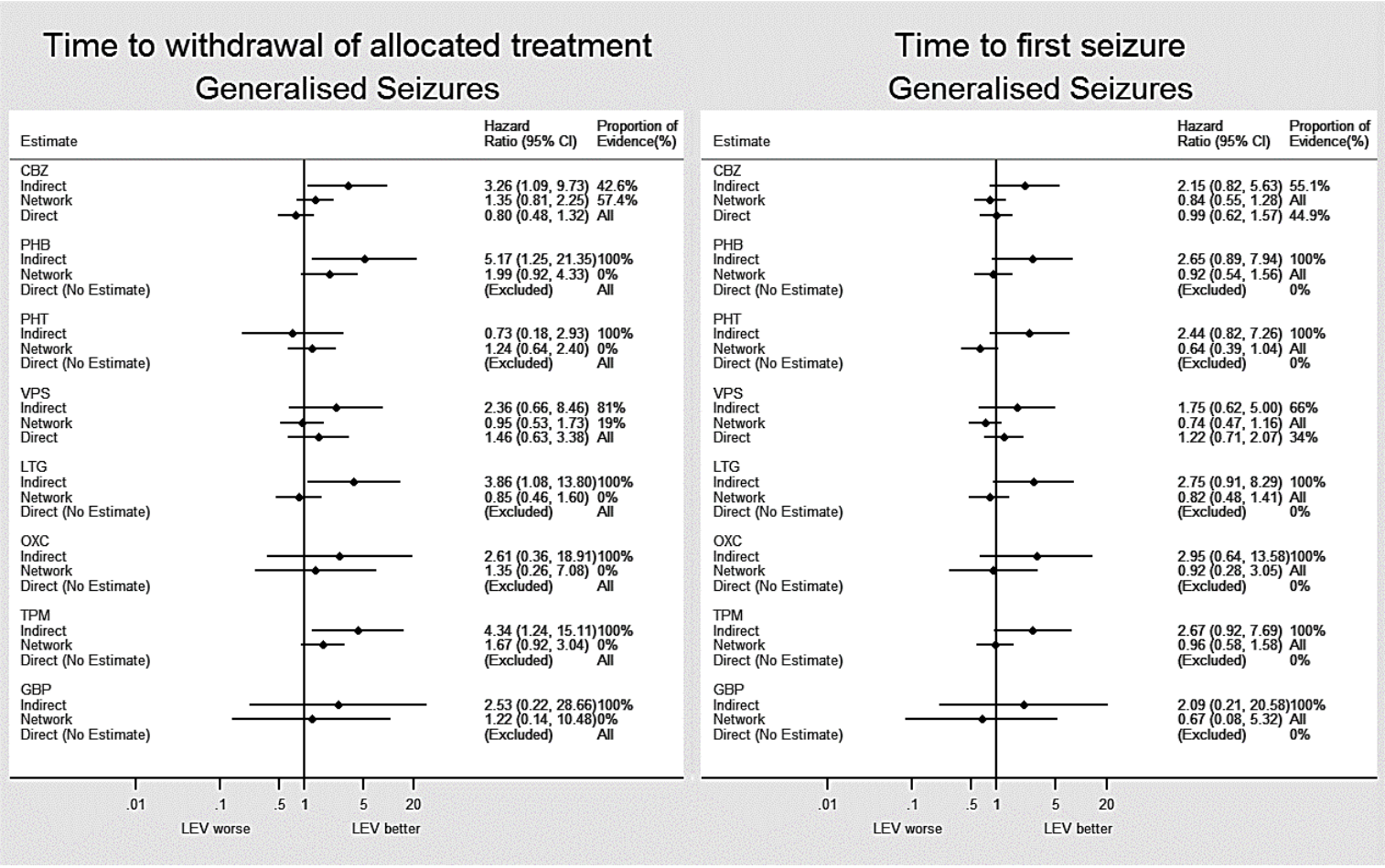


Figure 39: Time-to-withdrawal of allocated treatment and time-to-first seizure (lamotrigine (LTG) reference) for individuals with generalised seizures

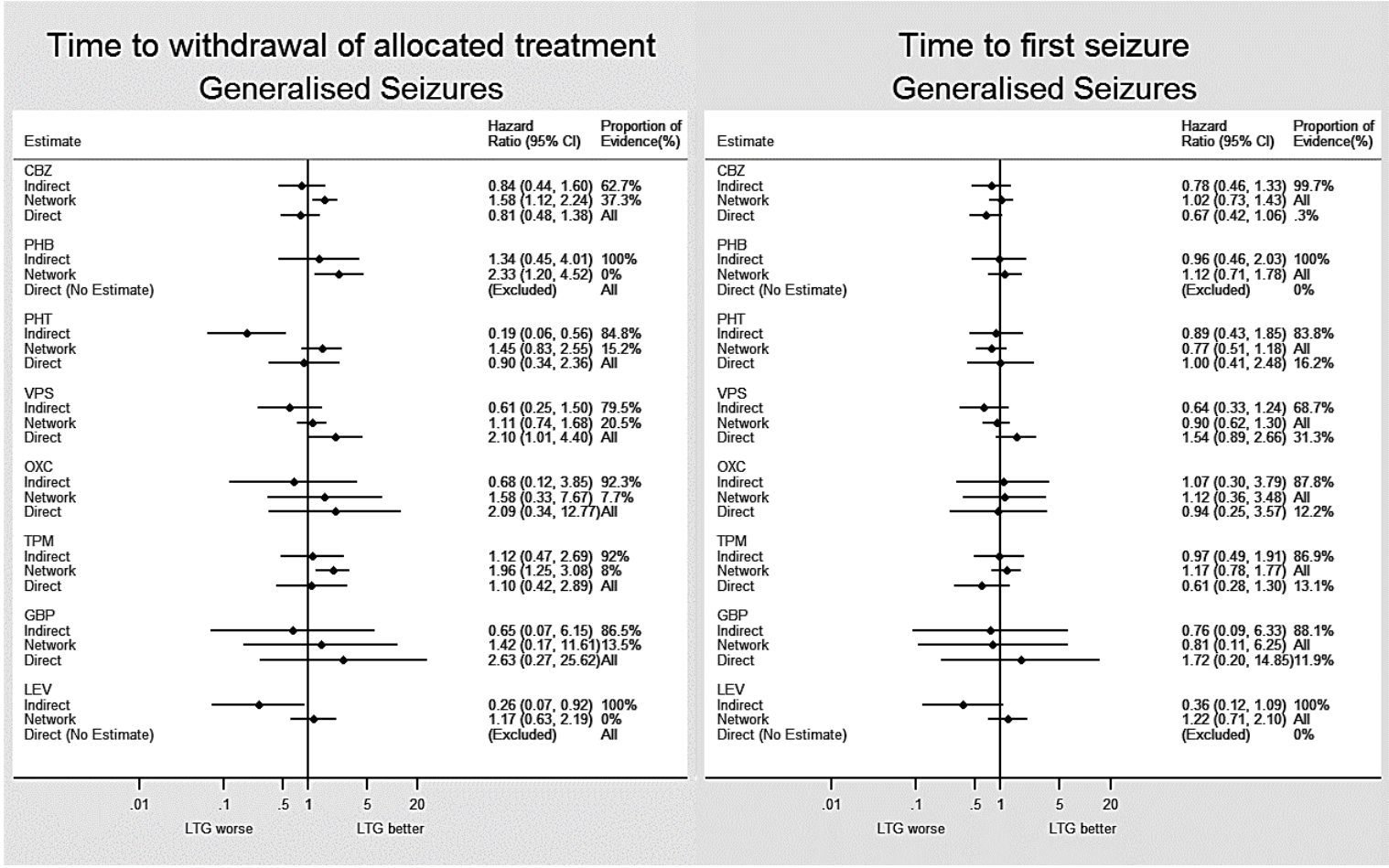


Figure 40: Time-to-withdrawal of allocated treatment and time-to-first seizure (oxcarbazepine (OXC) reference) for individuals with generalised seizure

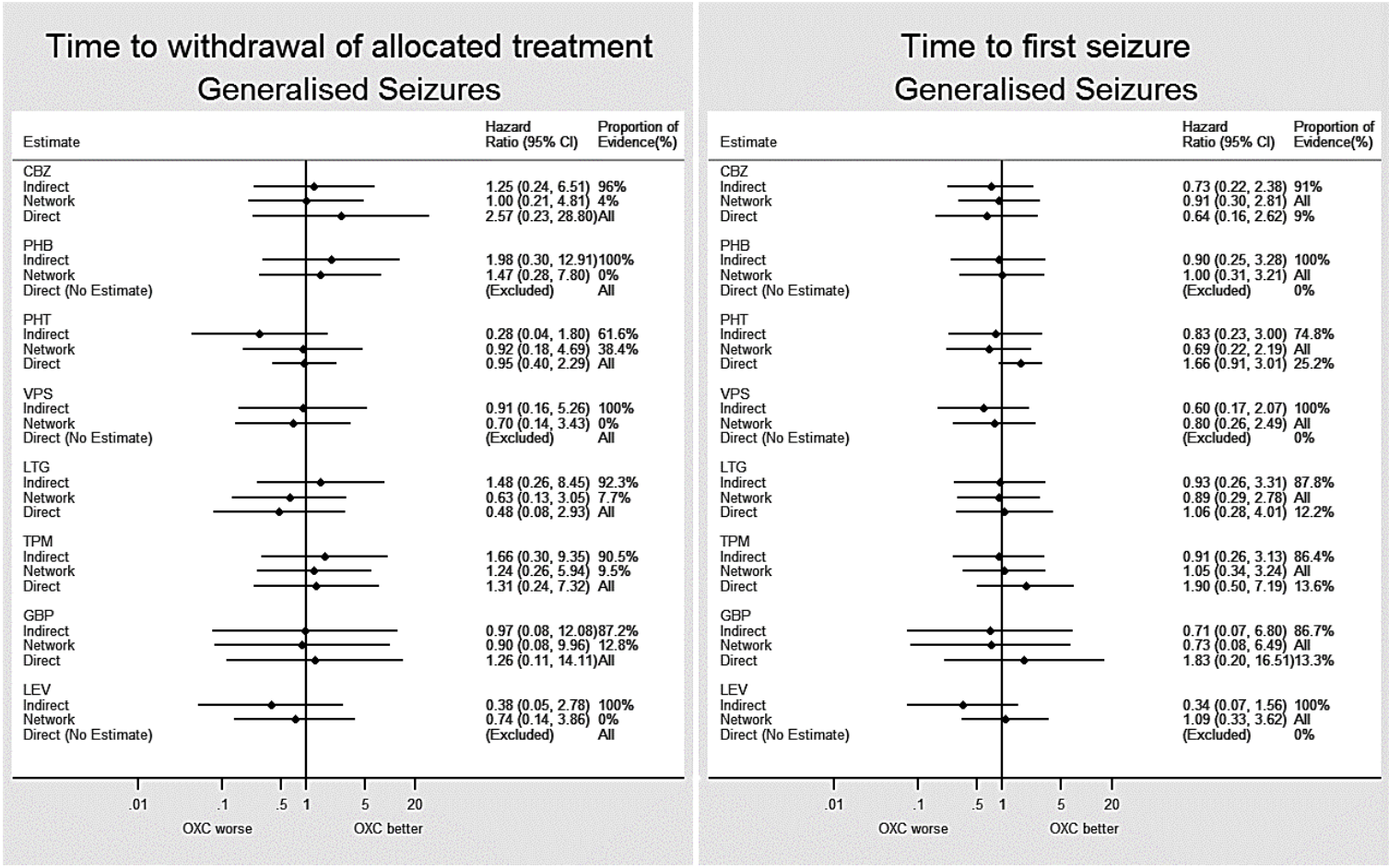


Figure 41: Time-to-withdrawal of allocated treatment and time-to-first seizure (phenobarbitone (PHB) reference) for individuals with generalised seizures

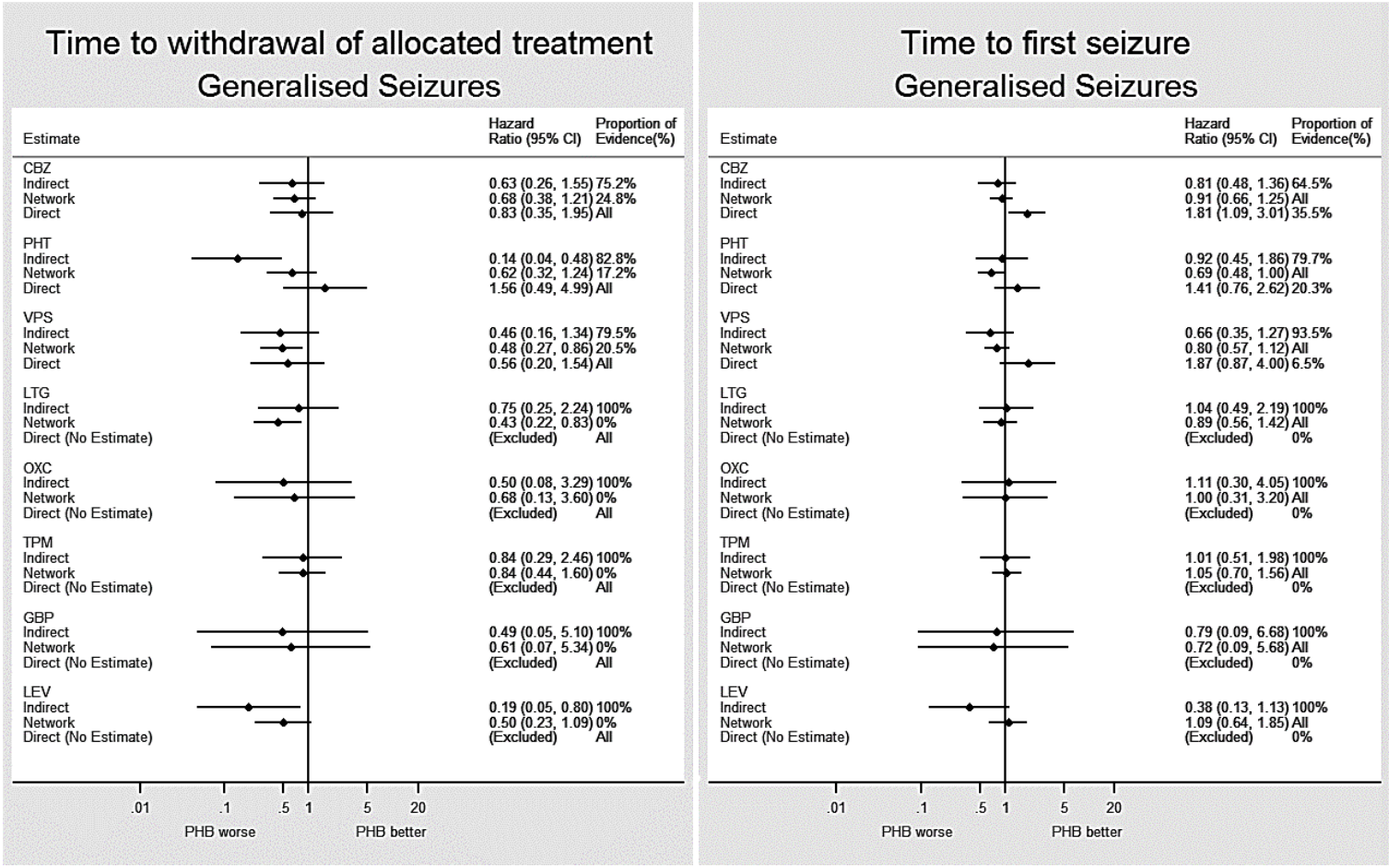


Figure 42: Time-to-withdrawal of allocated treatment and time-to-first seizure (phenytoin (PHT) reference) for individuals with generalised seizures

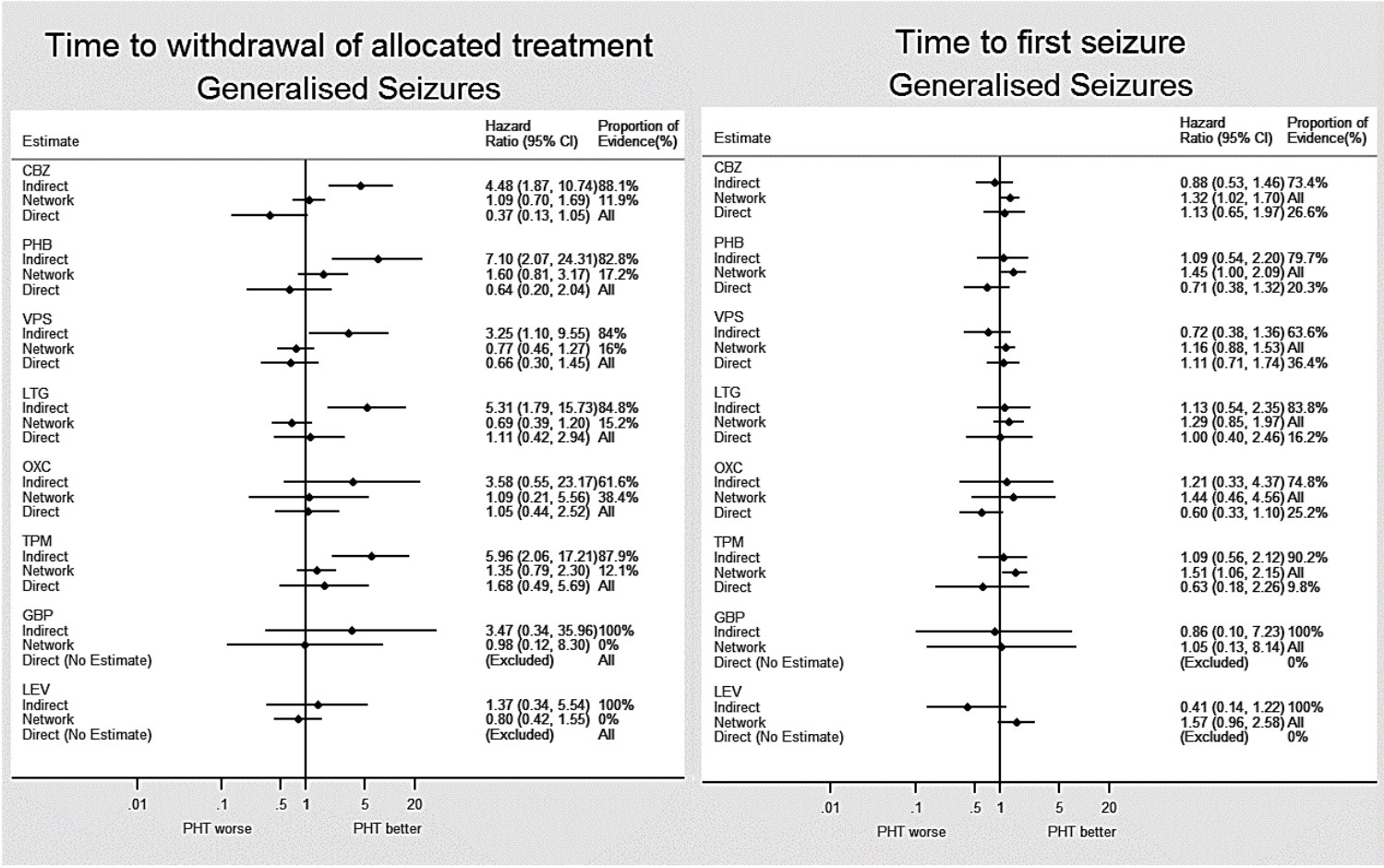


Figure 43: Time-to-withdrawal of allocated treatment and time-to-first seizure (topiramate (TPM) reference) for individuals with generalised seizures

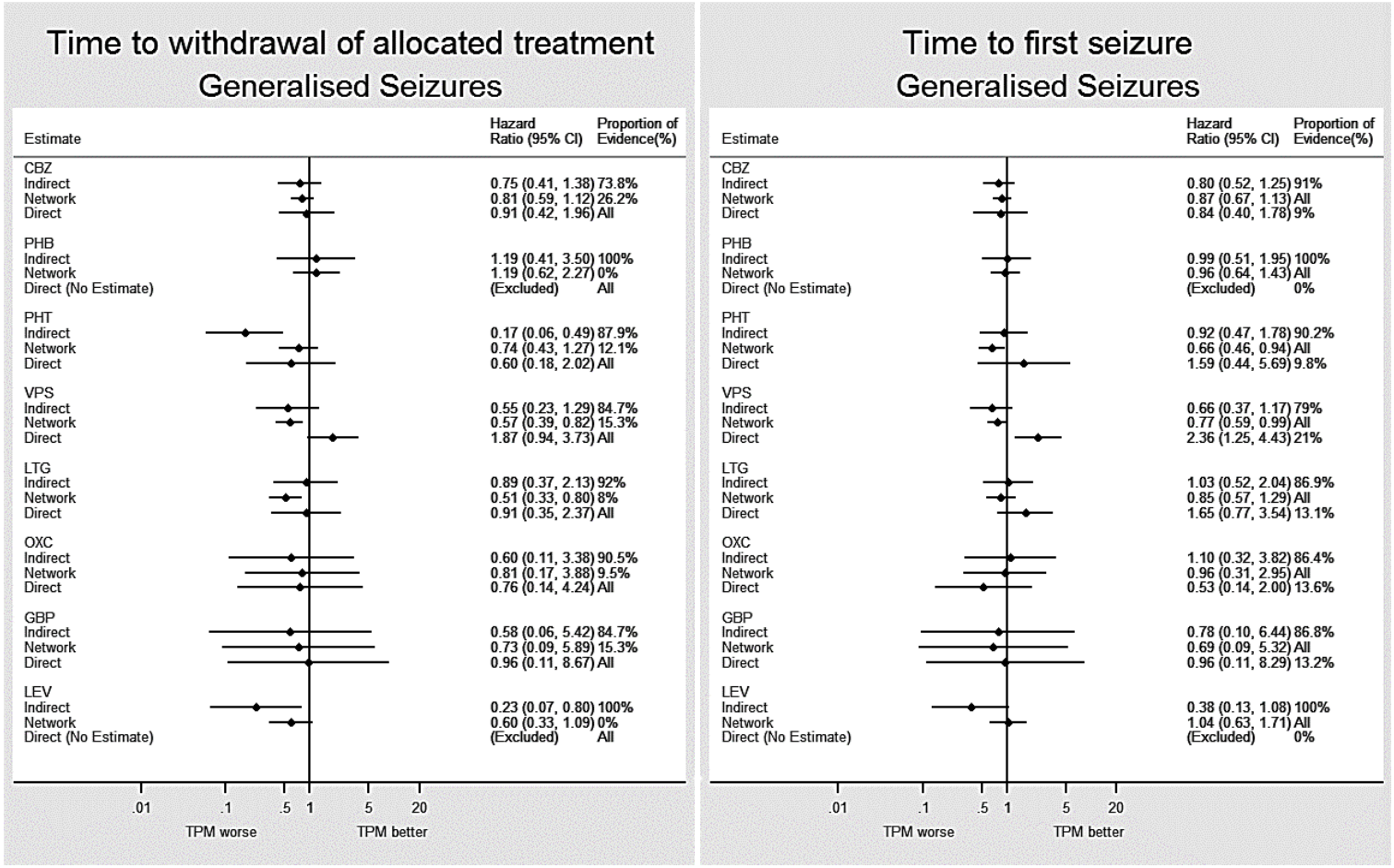


Figure 44: Time-to-withdrawal of allocated treatment and time-to-first seizure (sodium valproate (VPS) reference) for individuals with generalised seizures

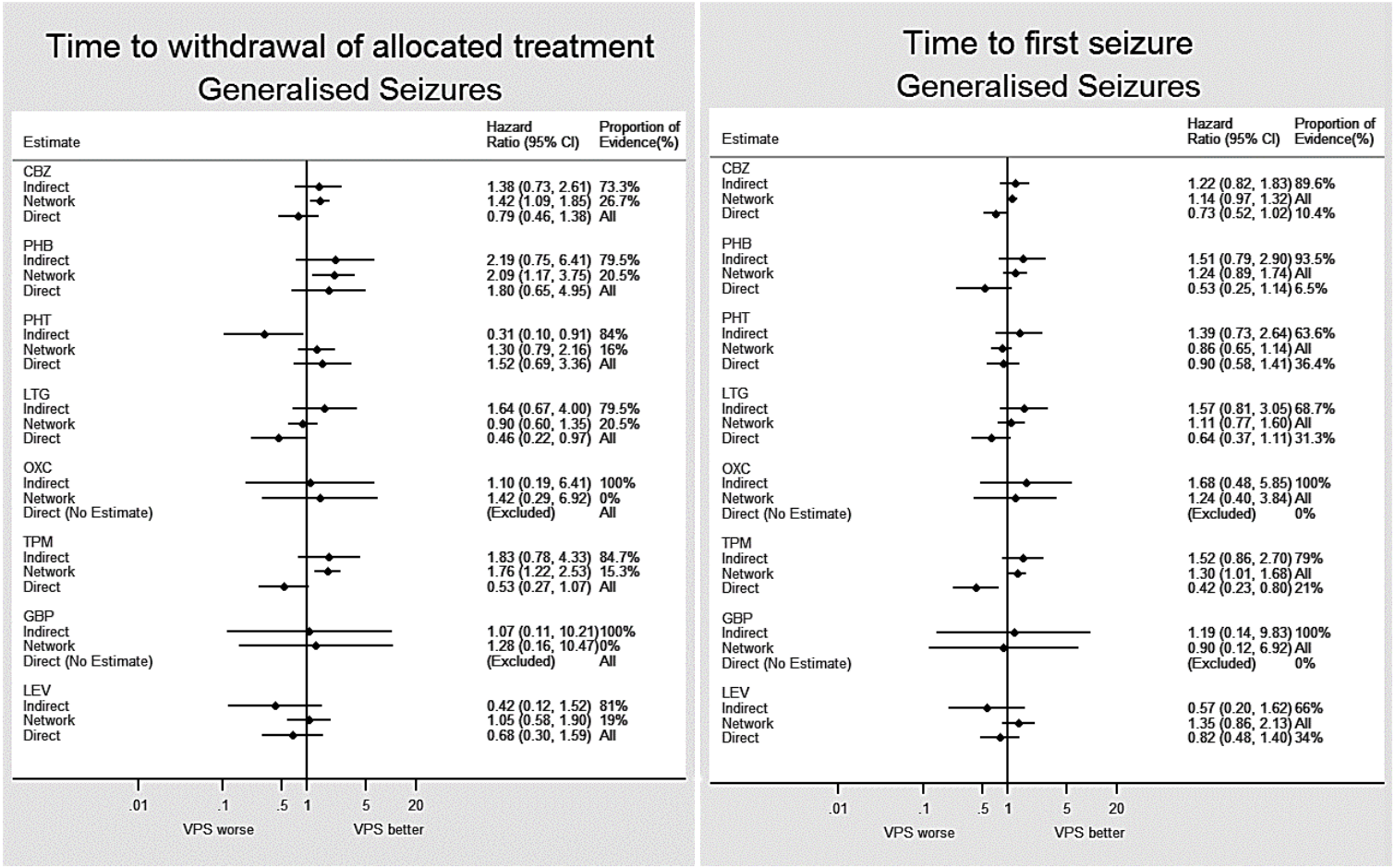


Figure 45: Time-to-withdrawal of allocated treatment and time-to-first seizure (carbamazepine (CBZ) reference) for individuals with partial seizures

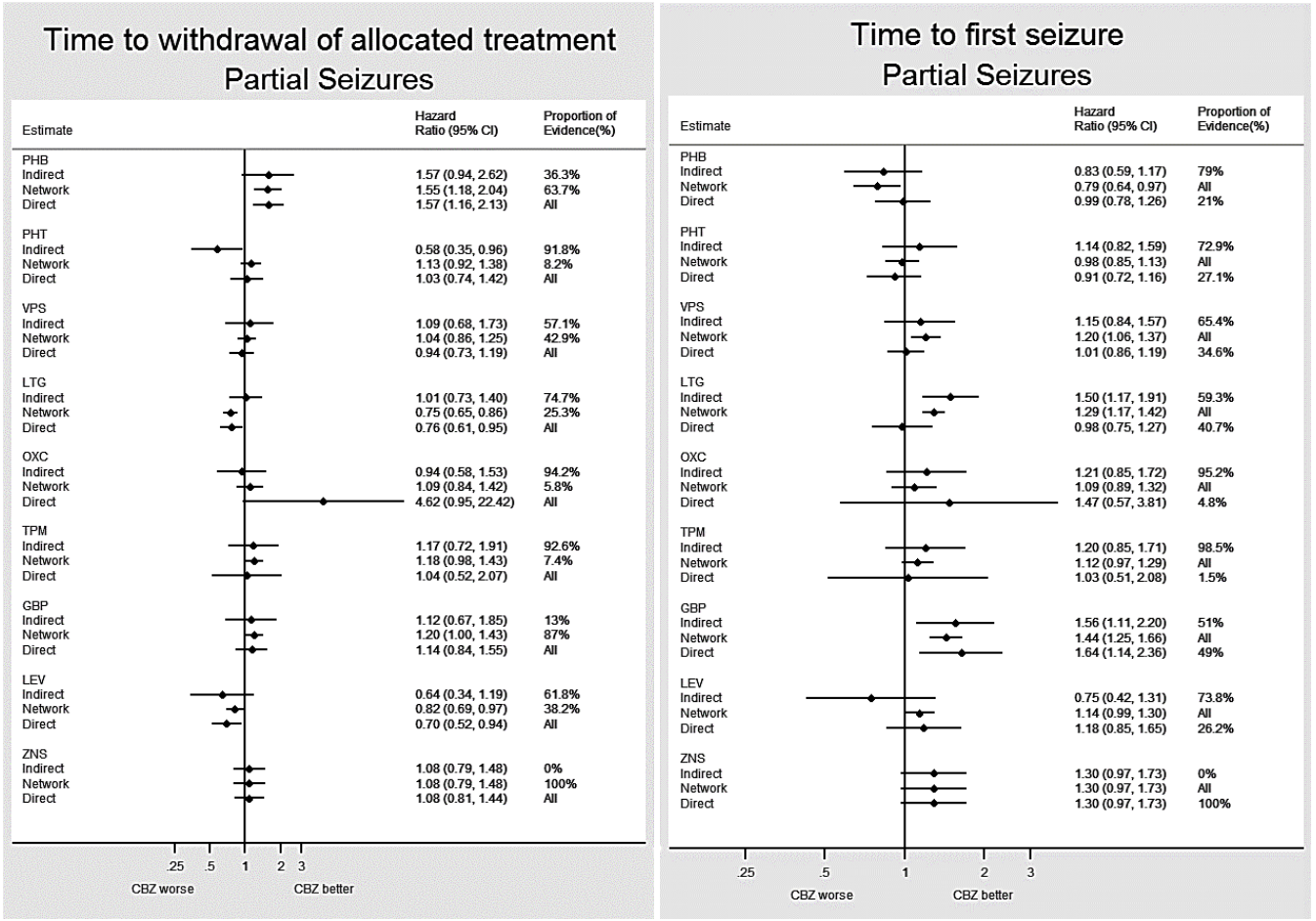


Figure 46: Time-to-withdrawal of allocated treatment and time-to-first seizure (gabapentin (GBP) reference) for individuals with partial seizures

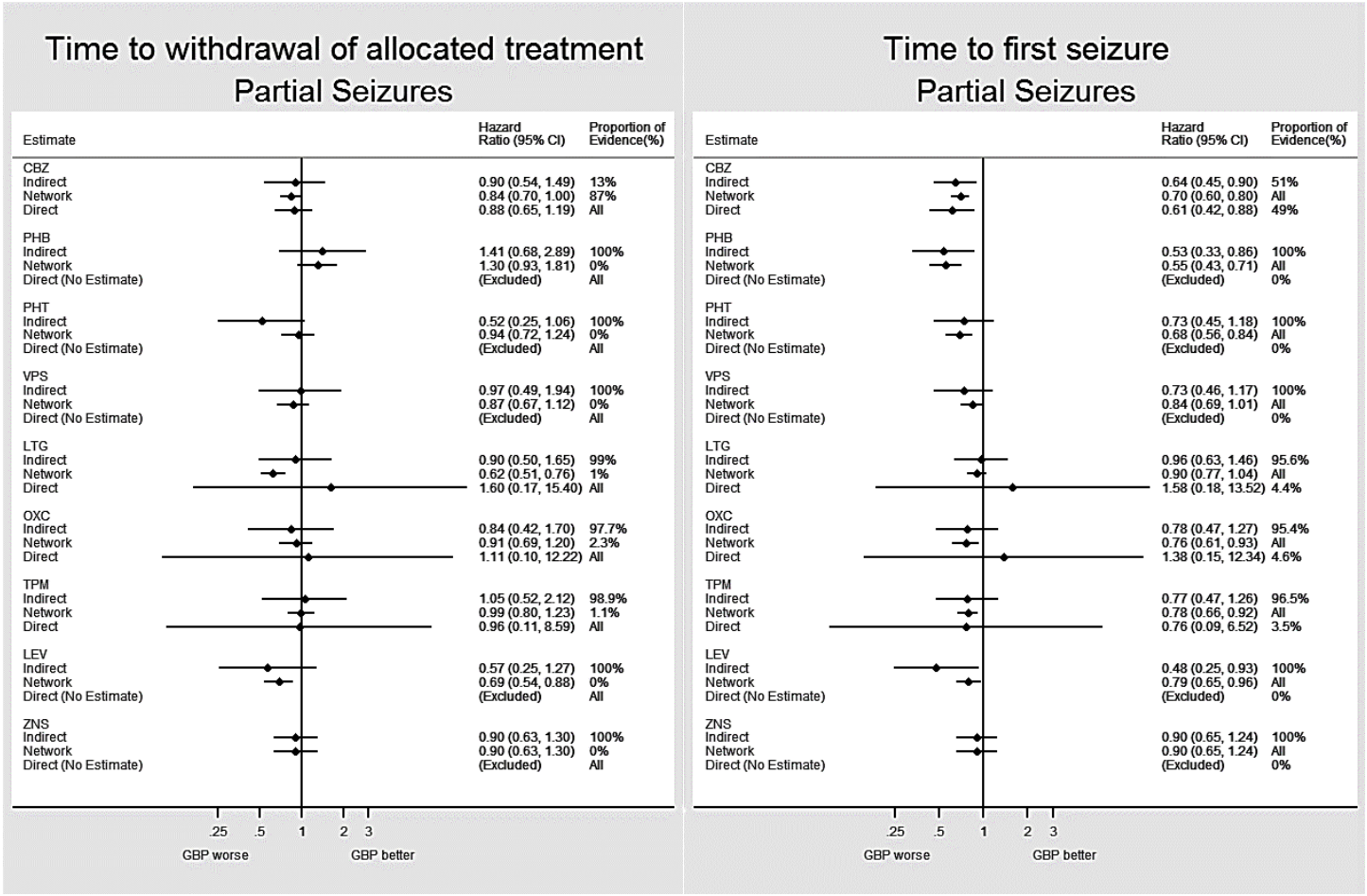


Figure 47: Time-to-withdrawal of allocated treatment and time-to-first seizure (levetiracetam (LEV) reference) for individuals with partial seizures

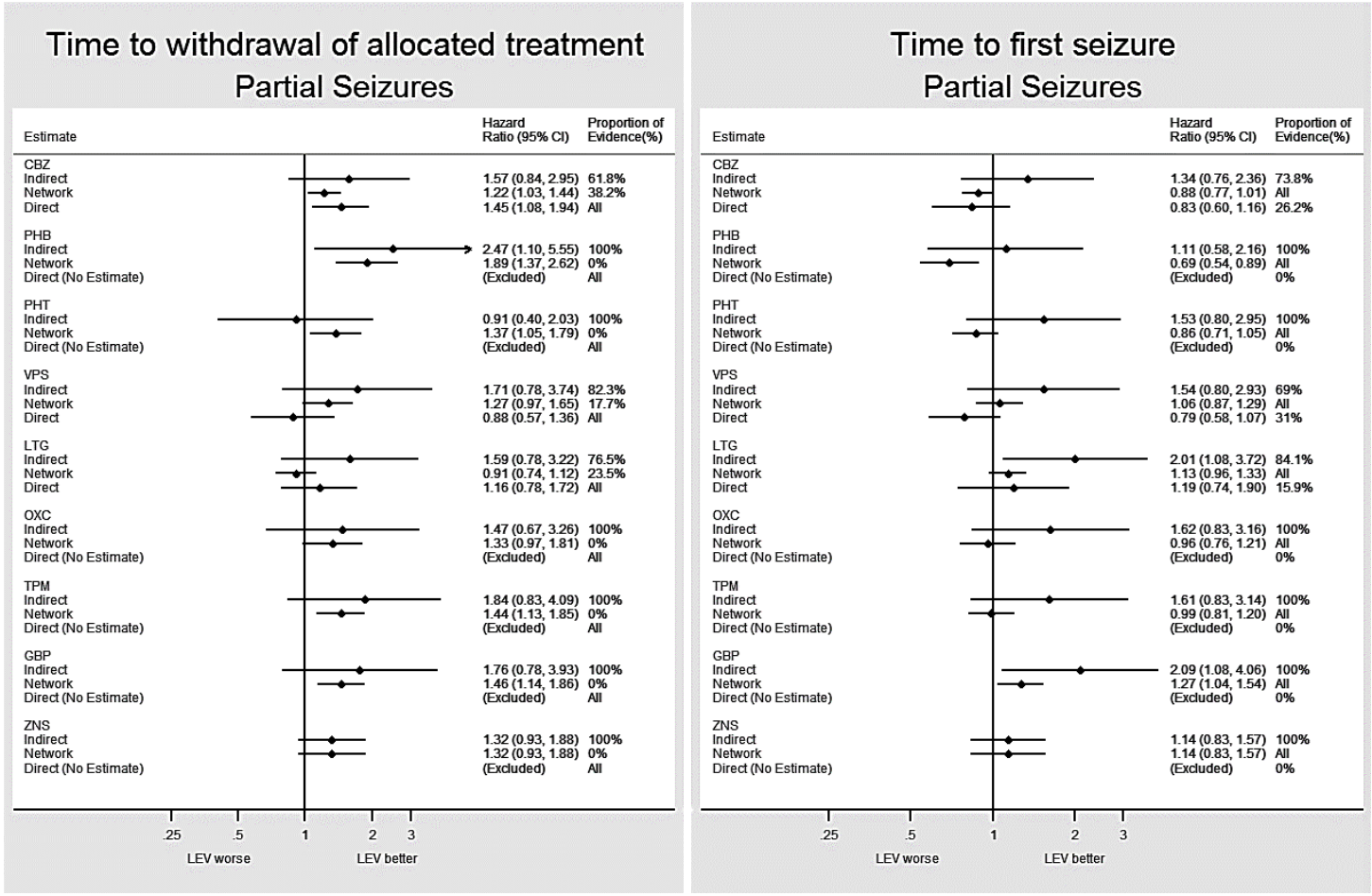


Figure 48: Time-to-withdrawal of allocated treatment and time-to-first seizure (lamotrigine (LTG) reference) for individuals with partial seizures

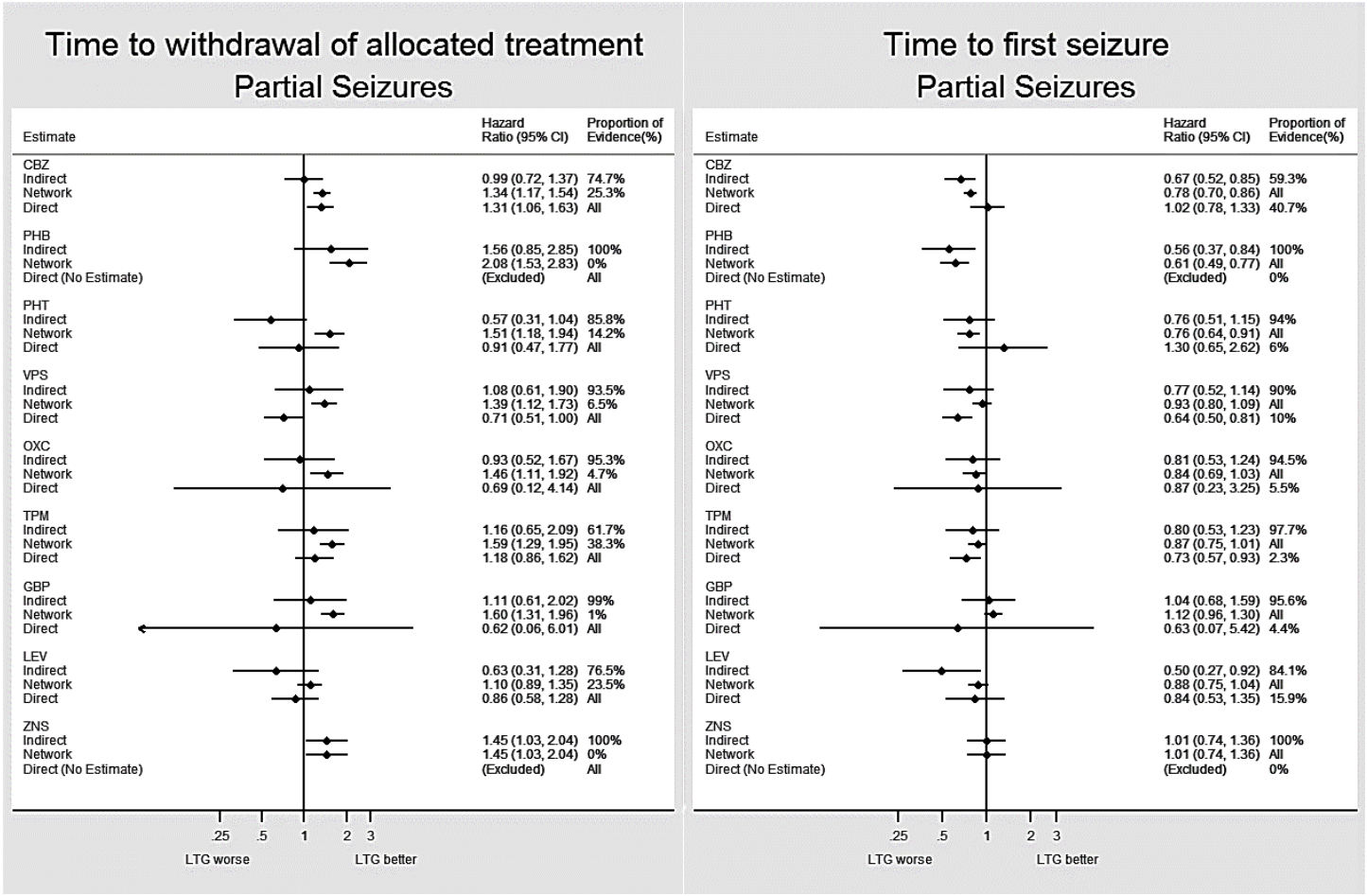


Figure 49: Time-to-withdrawal of allocated treatment and time-to-first seizure (oxcarbazepine (OXC) reference) for individuals with partial seizures

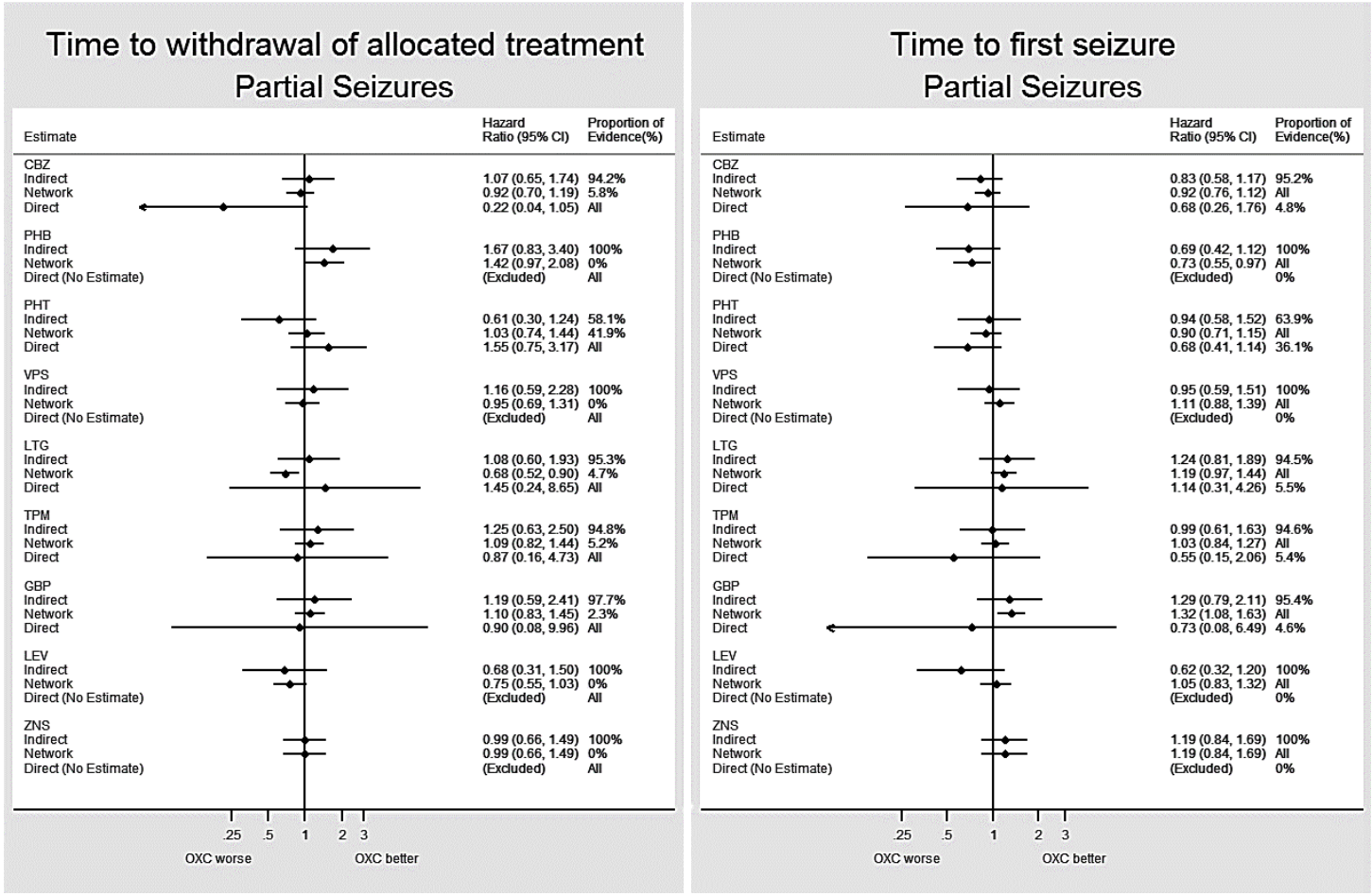


Figure 50: Time-to-withdrawal of allocated treatment and time-to-first seizure (phenobarbitone (PHB) reference) for individuals with partial seizures

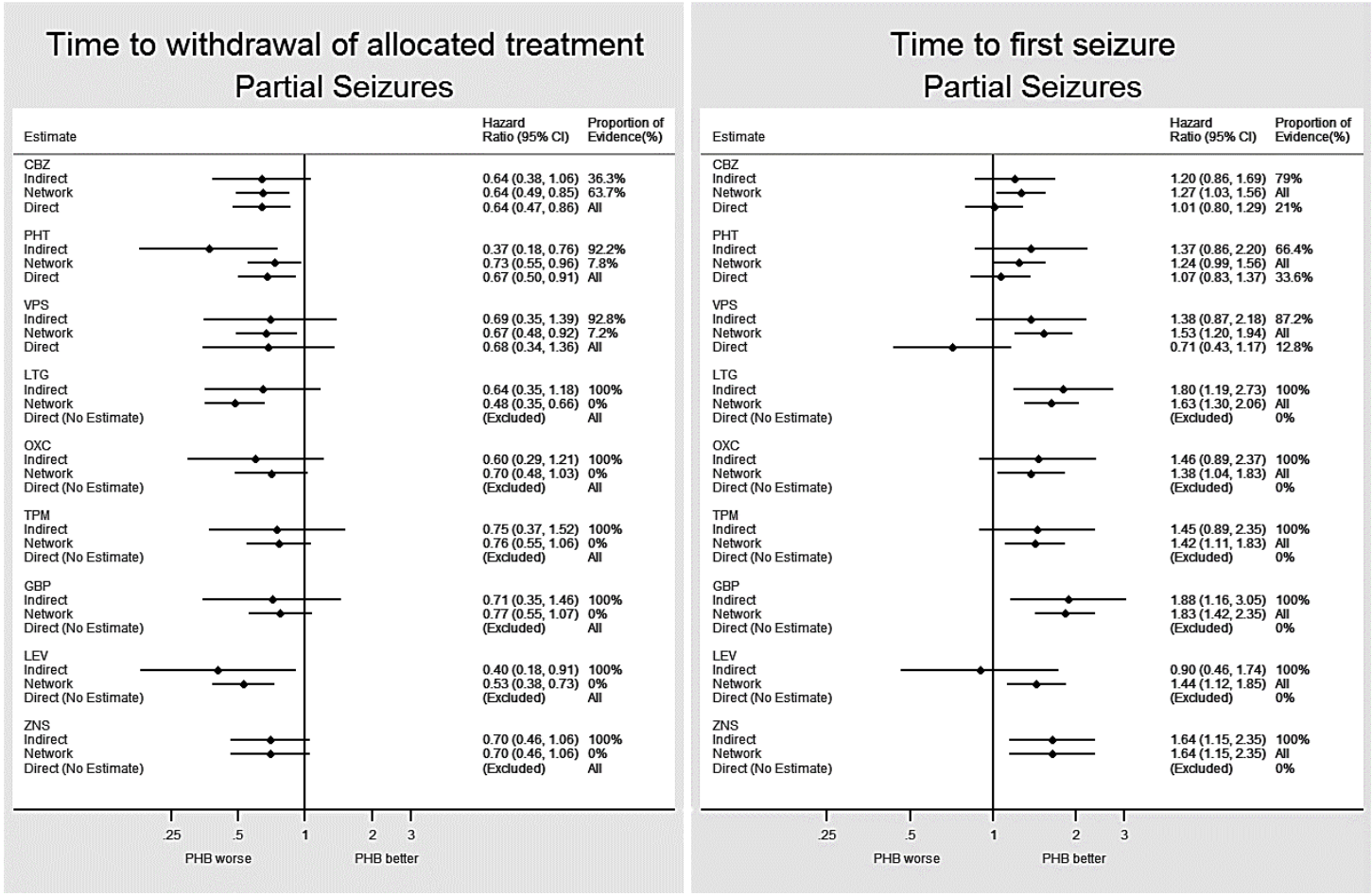


Figure 51: Time-to-withdrawal of allocated treatment and time-to-first seizure (phenytoin (PHT) reference) for individuals with partial seizures

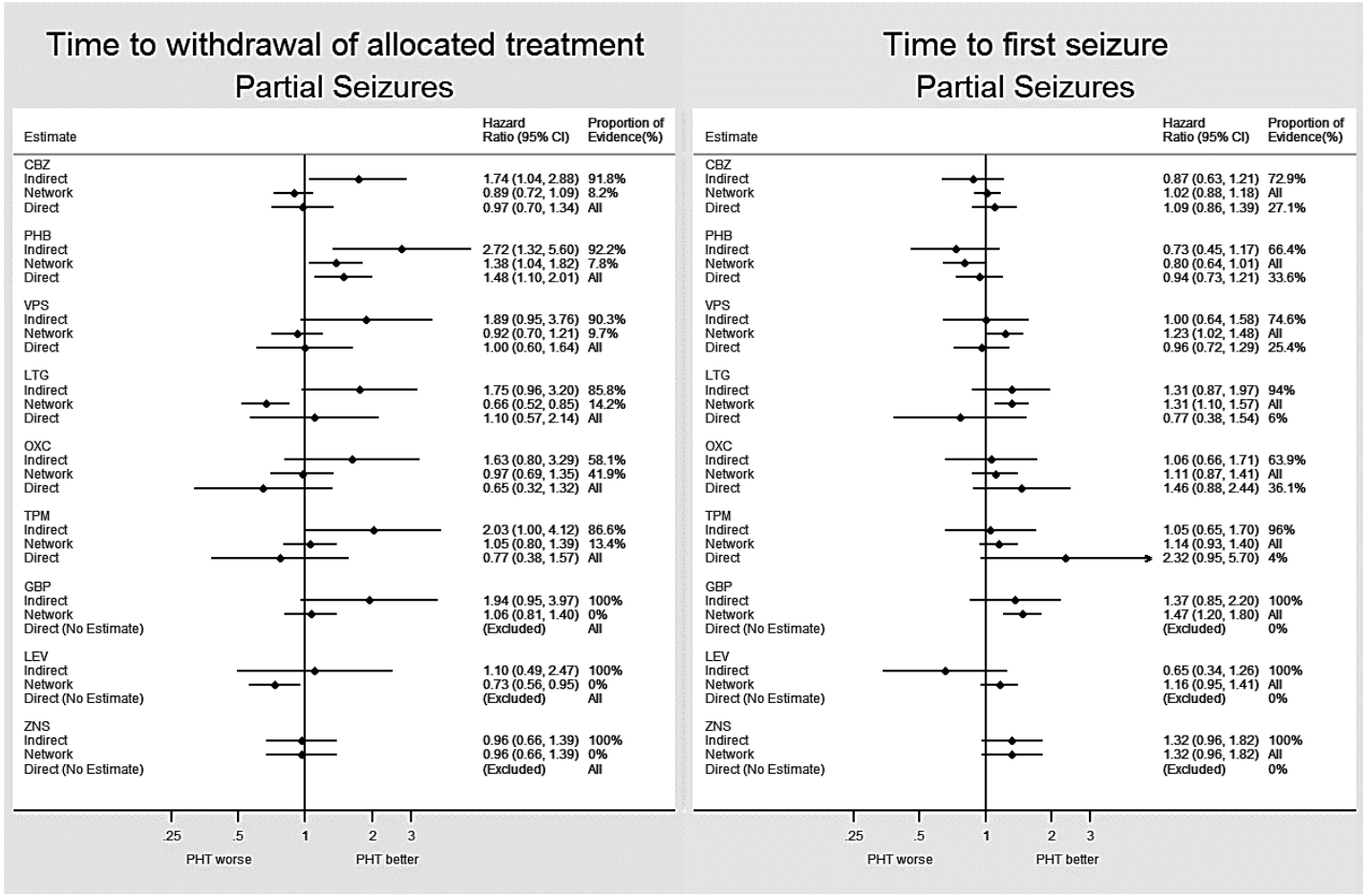


Figure 52: Time-to-withdrawal of allocated treatment and time-to-first seizure (topiramate (TPM) reference) for individuals with partial seizures

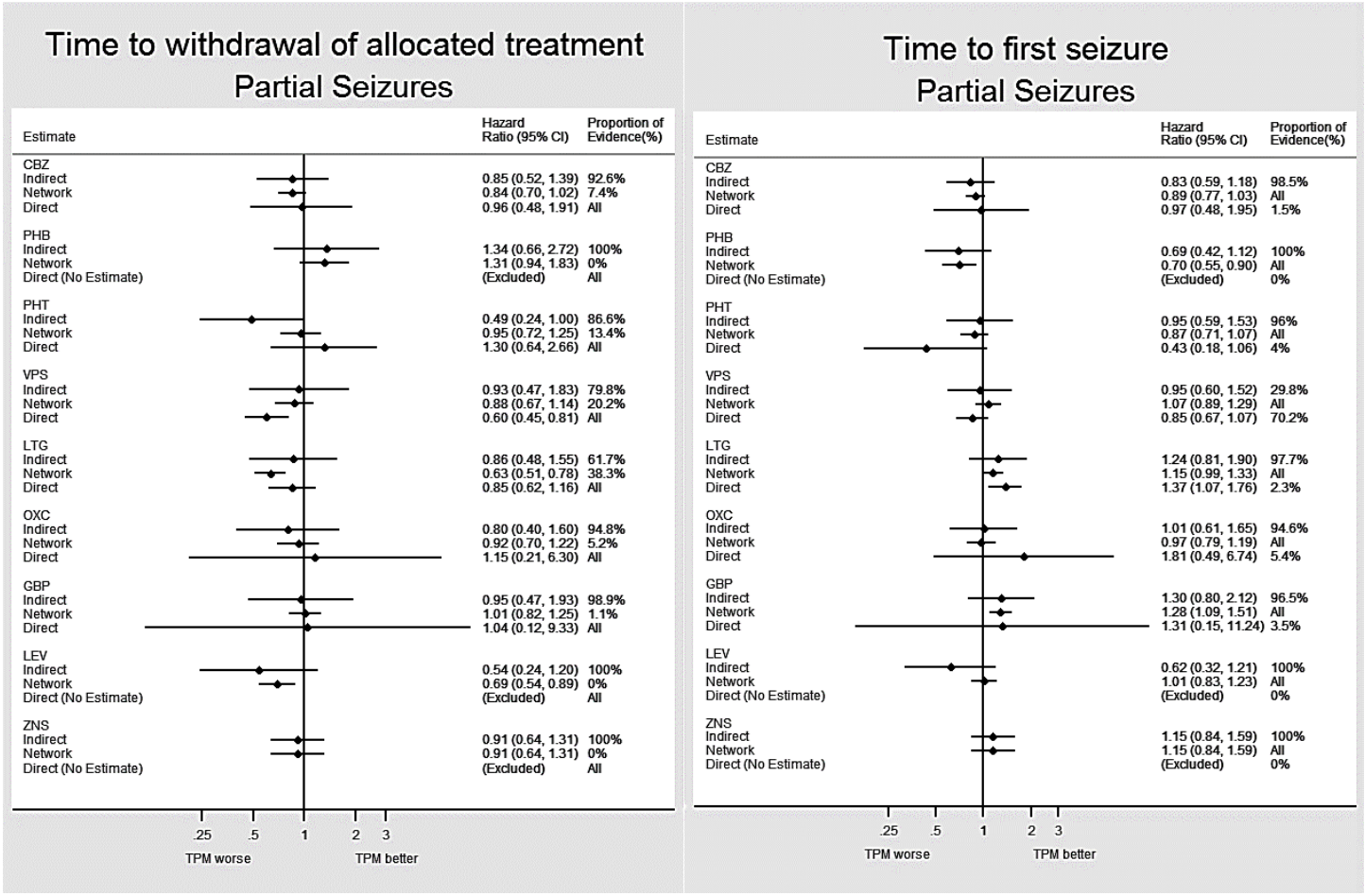


Figure 53: Time-to-withdrawal of allocated treatment and time-to-first seizure (sodium valproate (VPS) reference) for individuals with partial seizures

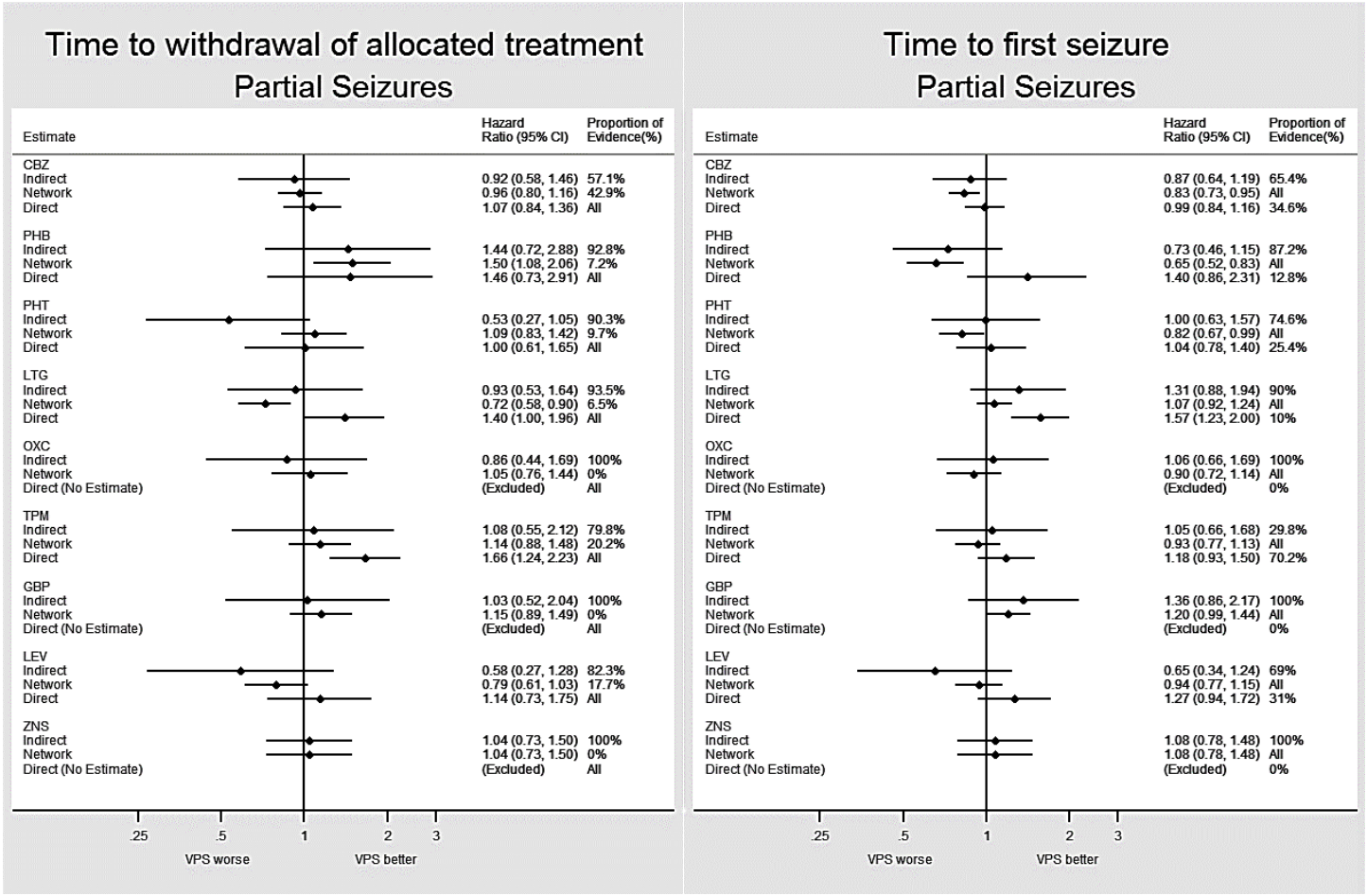
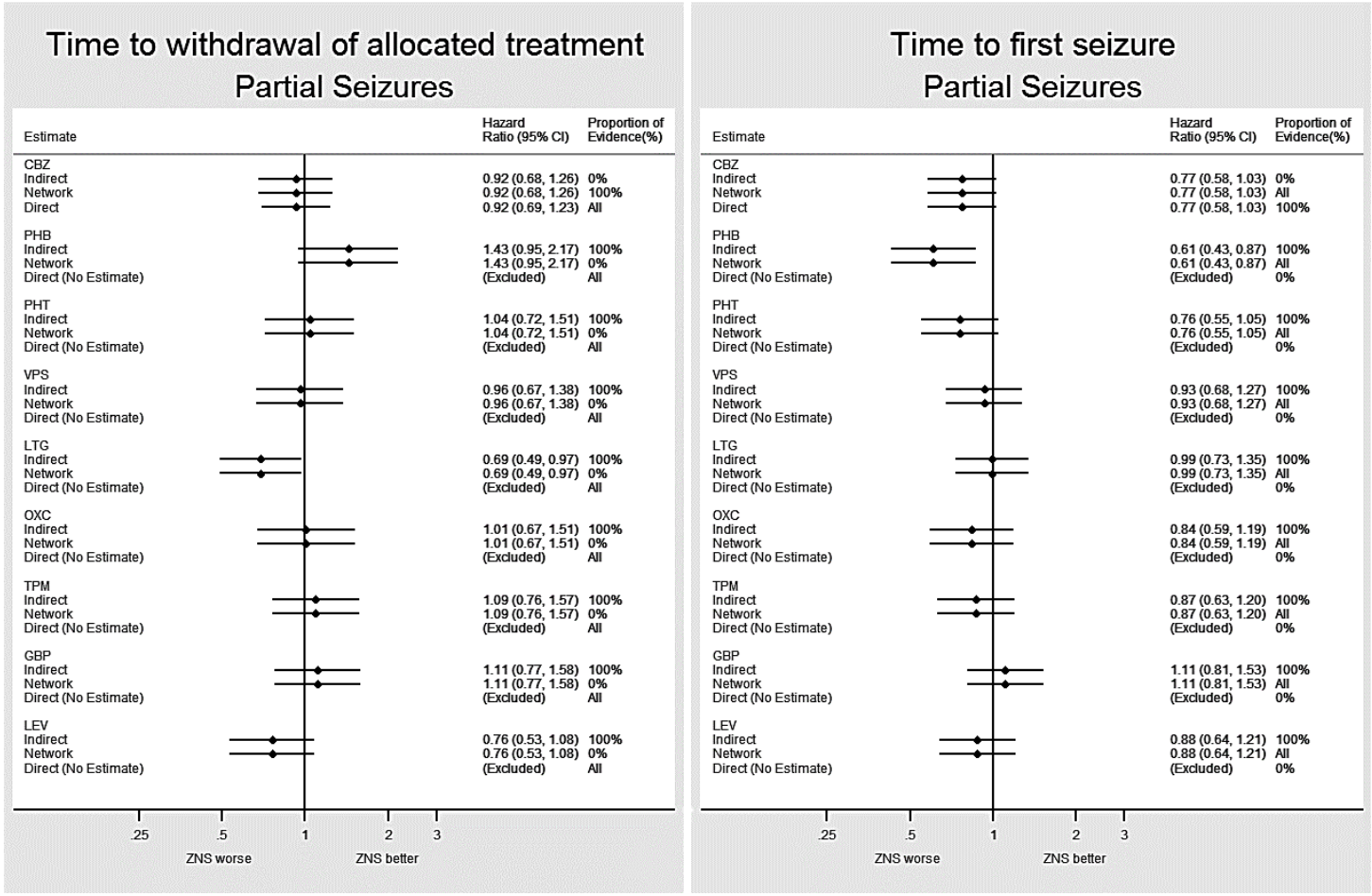


Figure 54: Time-to-withdrawal of allocated treatment and time-to-first seizure (zonisamide (ZNS) reference) for individuals with partial seizures



Appendix 17: Additional results from Chapter 7

Figures presented within this Appendix show NMA results for all pairwise comparisons for the additional analyses described in Chapter 8.

See Appendix 10 for references of the trials included in the Cochrane IPD-NMA and Chapter 5.2.1.3 for abbreviations of drugs.

Results shown on all figures are those from NMA (direct and indirect evidence combined).

Generalised tonic-clonic seizures with or without other seizure types is shortened to 'Generalised seizures' for brevity

Figure 55: NMA results of IPD reduced to summary statistics: individuals with partial seizures

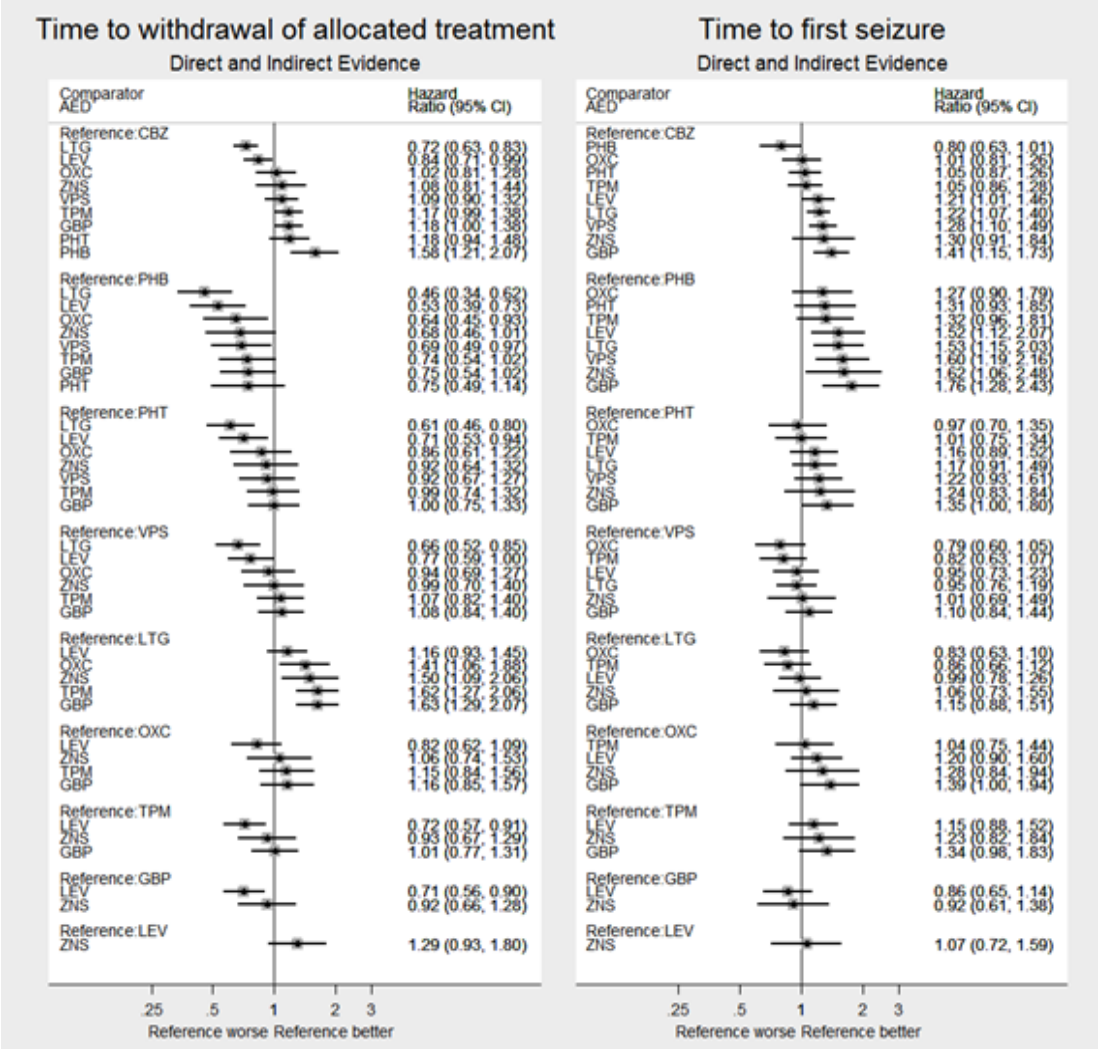


Figure 56: NMA results of IPD combined with AD: individuals with partial seizures

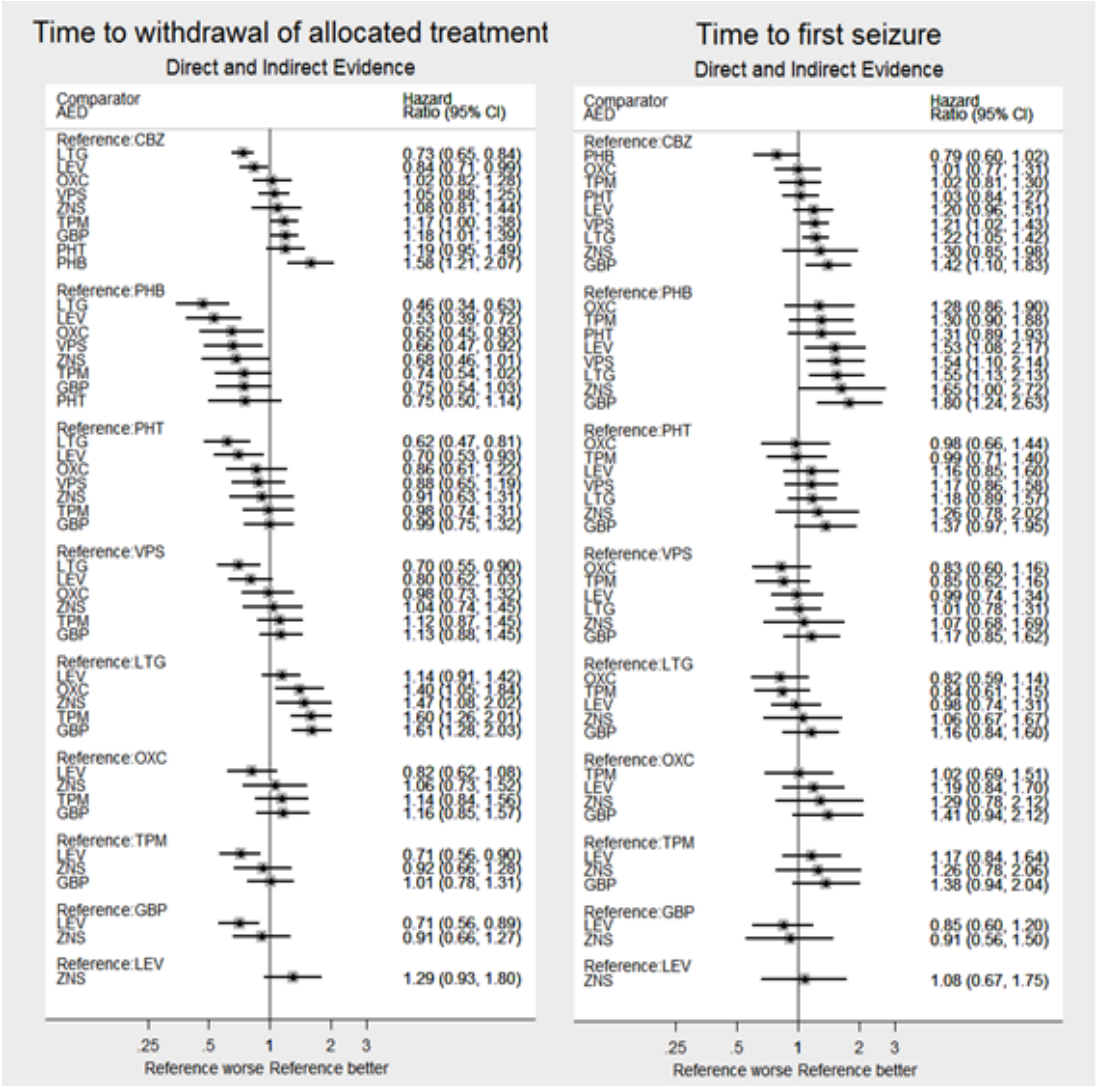


Figure 57: NMA results of IPD reduced to summary statistics: individuals with generalised seizures

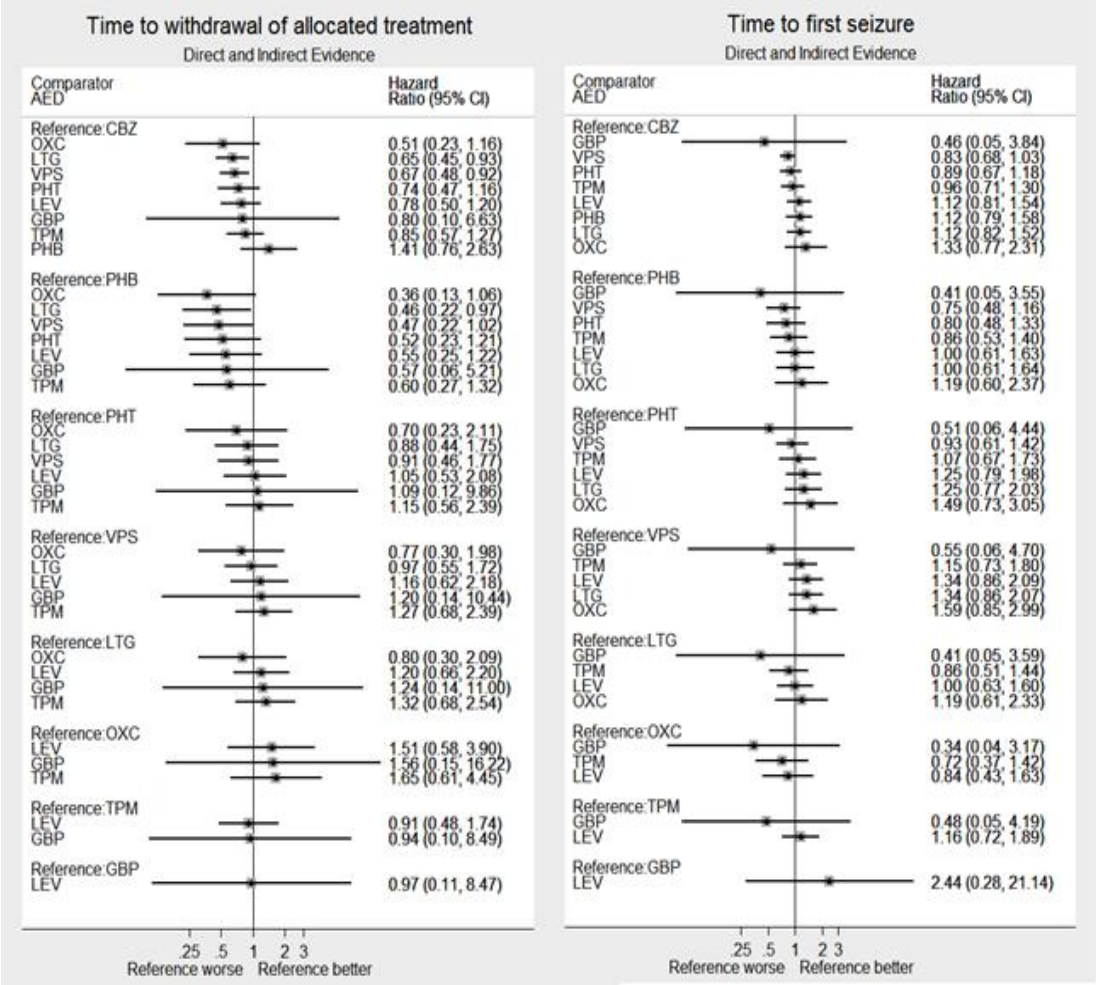


Figure 58: NMA results of IPD combined with AD: individuals with generalised seizures

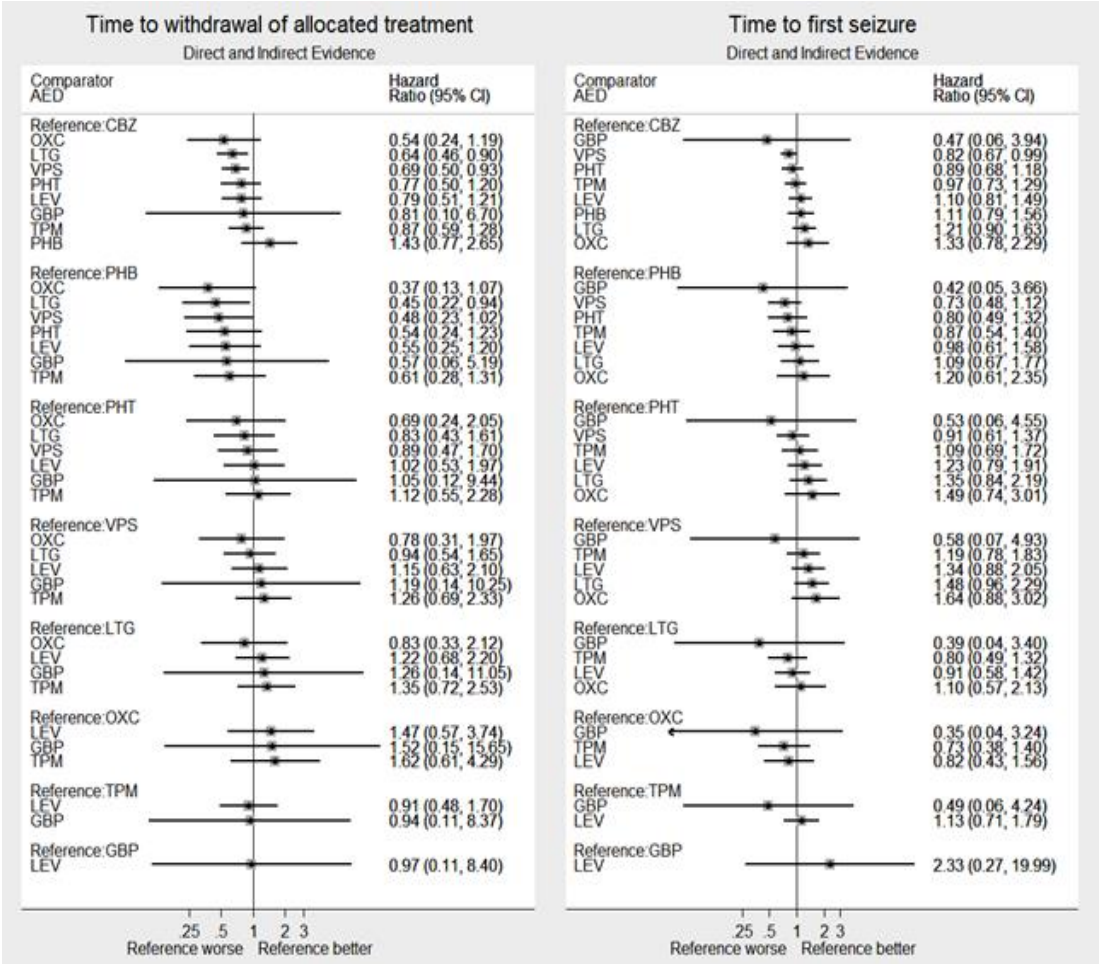


Figure 59: NMA results of IPD reduced to summary statistics for all individuals regardless of epilepsy type

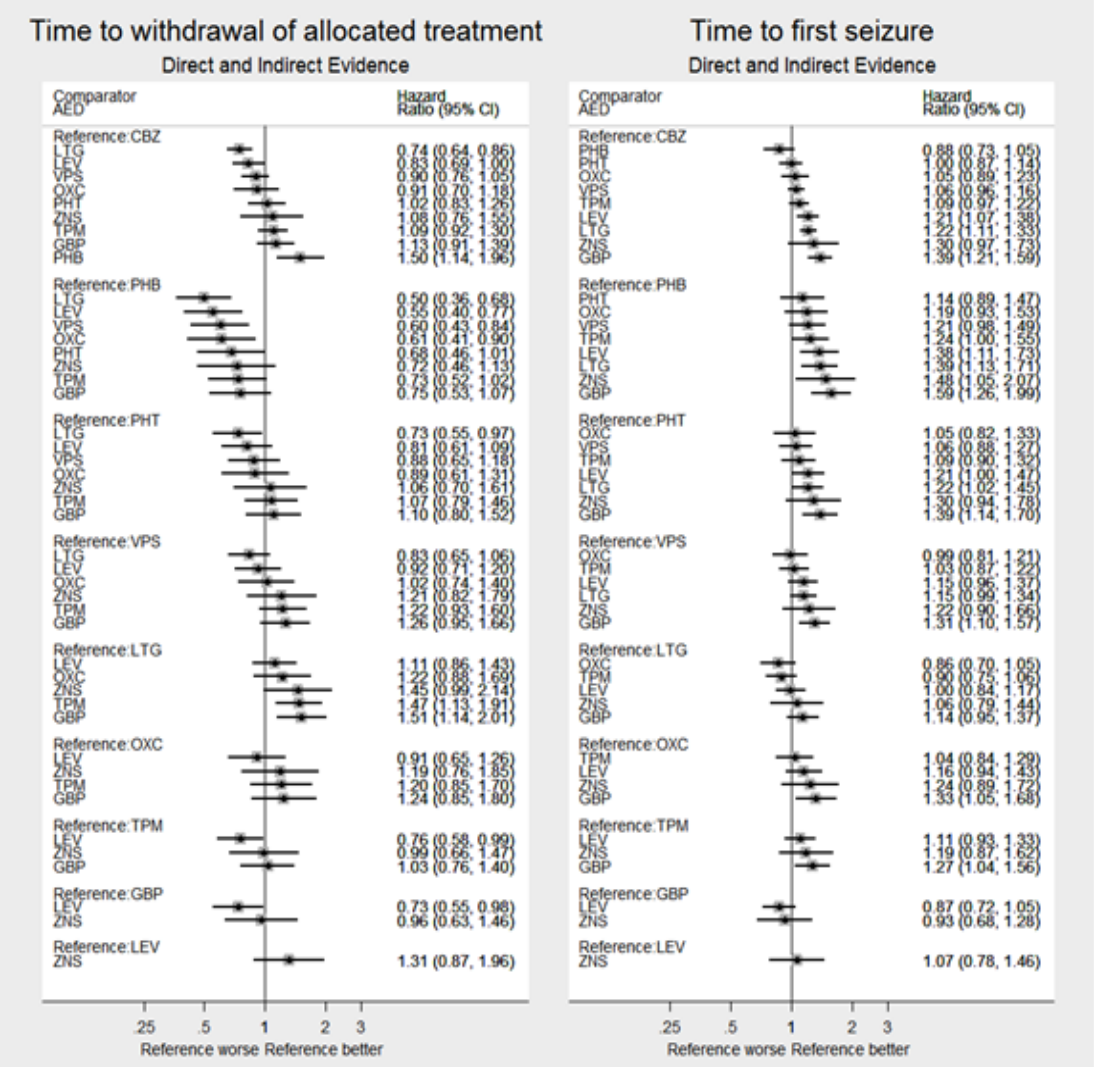
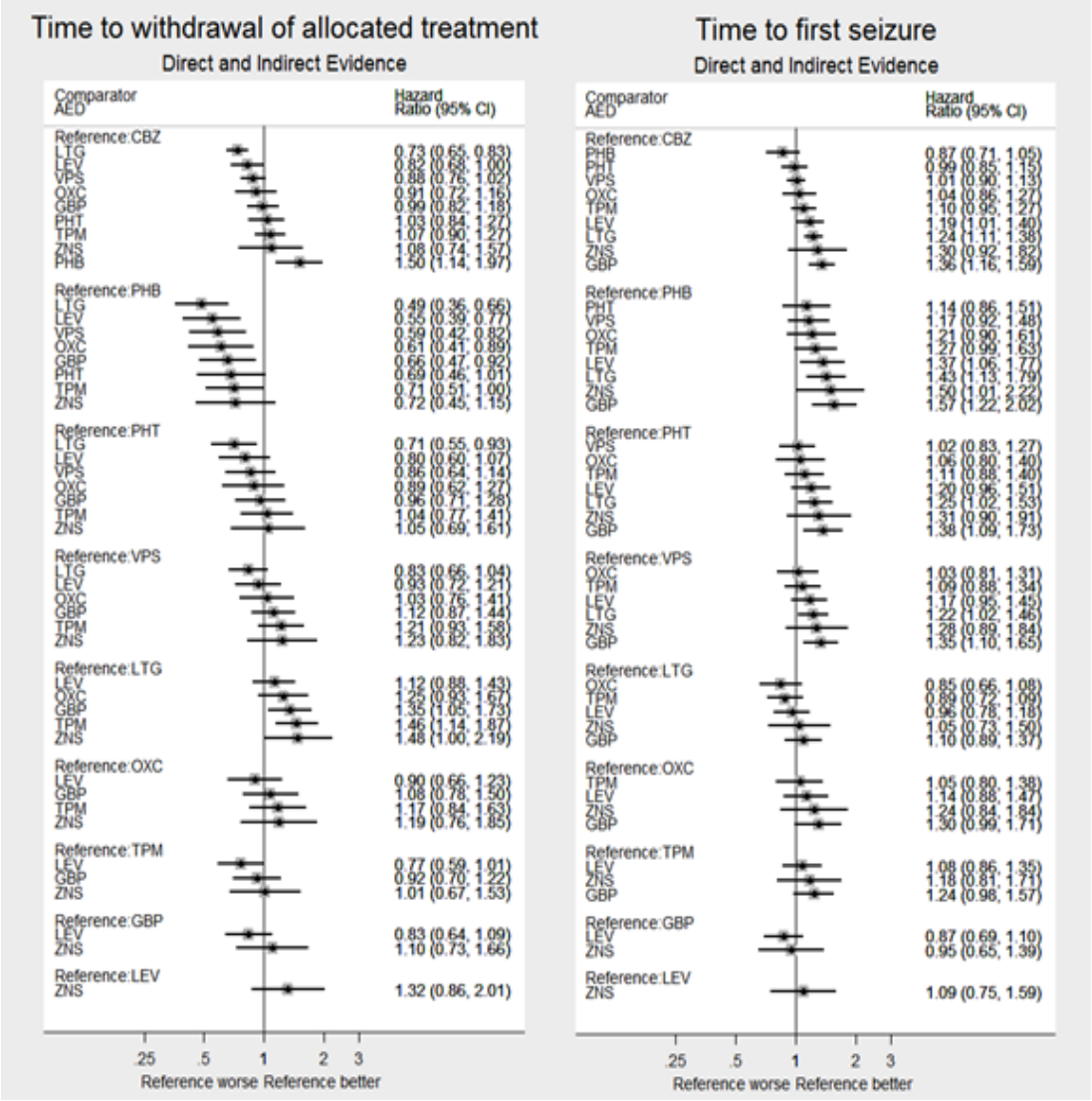


Figure 60: NMA results of IPD combined with AD for all individuals regardless of epilepsy type



Appendix 18: Published work from this thesis

A full list of publications and presentations of work in this thesis is provided at the start of this thesis. Locations of publications directly relating to this thesis are provided:

This text box is where the unabridged thesis included the following third party copyrighted material;

Nevitt SJ, Sudell M, Weston J, Tudur Smith C, Marson A. *Antiepileptic drug monotherapy for epilepsy: a network meta-analysis of individual participant data.* Cochrane Database of Systematic Reviews 2017, Issue 6, Art. No.: CD011412. DOI: 10.1002/14651858.CD011412.

doi: 10.1002/14651858.CD011412

This text box is where the unabridged thesis included the following third party copyrighted material;

Nevitt SJ, Marson AG, Davie B, Reynolds S, Williams L, Tudur Smith C. *Exploring changes over time and characteristics associated with data retrieval across individual participant data meta-analyses: systematic review*. BMJ 2017;357:j1390

doi: https://doi.org/10.1136/bmj.j1390

This text box is where the unabridged thesis included the following third party copyrighted material;

Nolan SJ, Sudell M, Weston J, Tudur Smith C, Marson A. *Antiepileptic drug monotherapy for epilepsy: an overview of systematic reviews and network meta-analysis (protocol).* Cochrane Database of Systematic Reviews 2014, Issue 12, Art. No.: CD011412. DOI: 10.1002/14651858.CD011412.

doi: 10.1002/14651858.CD011412